

CARTILAGE PATHOLOGY IN MPS VI:

disease modeling using iPSCs and
CRISPR/Cas9

Mike Broeders

The research described in this thesis was supported by Zeldzame Ziekten Fonds/WE Foundation, Metakids (project number 2018-082), and Stofwisselkracht.



The studies presented in this thesis were performed at the departments of Pediatrics and Clinical Genetics of the Erasmus University Medical Center, Rotterdam, The Netherlands.

ISBN: 978-94-6419-216-2

Author: Mike Broeders

Printed by: Gildeprint

Copyright © Mike Broeders, 2021, Rotterdam, The Netherlands

All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, without prior written permission of the author.

CARTILAGE PATHOLOGY IN MPS VI:

disease modeling using iPSCs and CRISPR/Cas9

KRAAKBEENPATHOLOGIE IN MPS VI:

ziekte modellering met behulp van iPSCs en CRISPR/Cas9

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus
Prof.dr. F.A. van der Duijn Schouten
and in accordance with the decision of the Doctorate Board.

The public defence shall be held on

June 1st at 15:30 hours

by

Mike Broeders
born in Tilburg

Doctoral Committee

Promotor Prof.dr. A.T. van der Ploeg

Other members Prof.dr. E.H.H.M. Rings
Prof.dr. J. Gribnau
Prof.dr. M.A. Tryfonidou

Copromotor Dr. W.W.M. Pijnappel

TABLE OF CONTENTS

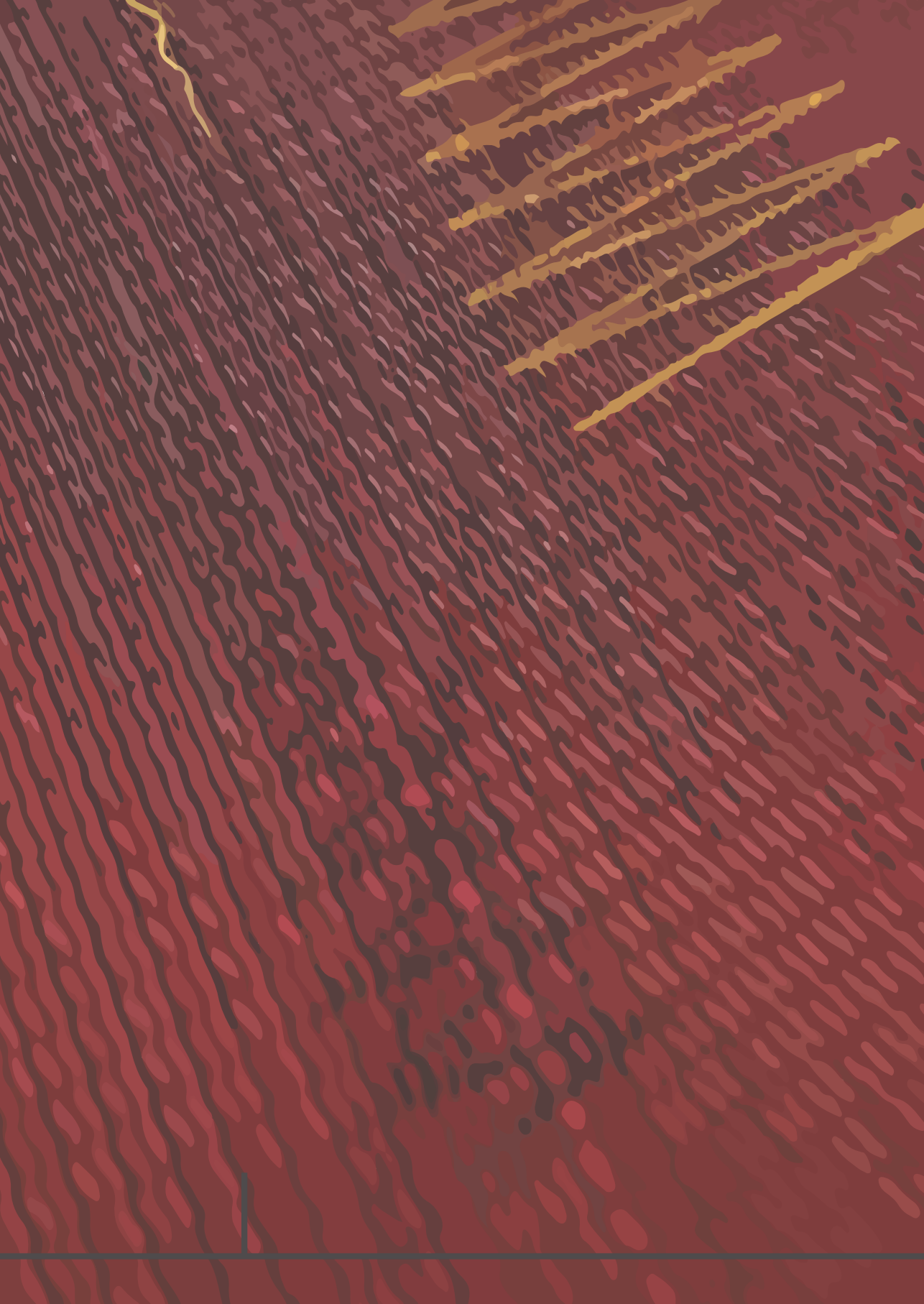
| | | |
|------------------------------|--|-----|
| List of abbreviations | | 6 |
| Chapter 1 | General introduction and outline of thesis | 11 |
| Chapter 2 | Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options | 49 |
| Chapter 3 | A generic assay to detect aberrant ARSB splicing and mRNA degradation for the molecular diagnosis of MPS VI | 109 |
| Chapter 4 | Sharpening the molecular scissors: advances in gene-editing technology | 129 |
| Chapter 5 | Ready for repair? Gene editing enters the clinic for the treatment of human disease | 151 |
| Chapter 6 | Large-Scale Expansion of Human iPSC-Derived Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies | 181 |
| Chapter 7 | CRISPR/Cas9-mediated gene editing in human induced pluripotent stem cells | 209 |
| Chapter 8 | Generation of a disease model for cartilage pathology in MPS VI using patient-derived and isogenic gene-corrected hiPSCs | 243 |
| Chapter 9 | General Discussion | 279 |
| Addendum | Summary | 292 |
| | Samenvatting | 294 |
| | Curriculum vitae | 296 |
| | PhD portfolio | 297 |
| | List of publications | 299 |
| | Acknowledgements | 300 |

List of abbreviations

| | | | |
|-------------|---|---------|--------------------------------------|
| 2'O-mePS | 2'O-methyl phosphorothioate | CRISPRi | CRISPR interference |
| AAV | Adeno associated virus | CS | Chondroitin sulphate |
| AC | Articular cartilage | dCas9 | Deficient Cas9 protein |
| ACT | Adoptive cell therapy | DCs | Dendritic cells |
| ADA | Adenosine deaminase | DMD | Duchenne muscular dystrophy |
| ADAR2 | Adenosine deaminase that acts on RNA type 2 | DS | Dermatan sulphate |
| ADA-SCID | ADA severe combined immunodeficiency | DSBs | Double-strand breaks |
| a-EJ | Alternative end joining | EBV | Epstein-Barr virus |
| AIDS | Acquired immunodeficiency syndrome | ECM | Extracellular matrix |
| ALL | Acute lymphoblastic leukemia | ER | Endoplasmic reticulum |
| AML | Acute myeloid leukemia | ERT | Enzyme replacement therapy |
| AON | Antisense oligonucleotide | ESC | Embryonic stem cell |
| ARSB | Arylsulfatase B | ESE | Exonic splicing enhancer |
| ARSB | N-Acetylgalactosamine-4-sulfatase | eSpCas9 | Enhanced specificity Cas9 |
| ART | Antiretroviral therapy | ESS | Exonic splicing silencer |
| BER | Base excision repair | EV | Empty targeting vector |
| BMD | Becker muscular dystrophy | evoCas9 | Evolved Cas9 |
| C4S | Chondroitin 4-sulfate | FACS | Fluorescence-activated cell sorting |
| C6S | Chondroitin 6-sulfate | FC | Fold change |
| CADD | Combined annotation-dependent depletion | FDR | False discovery rate |
| CAR | Chimeric antigen receptor | FIX | Clotting factor IX |
| CD19 | Cluster of differentiation 19 | FVIII | Clotting factor VIII |
| cDNA | Complementary DNA | GAA | Acid α -glucosidase |
| CGD | Chronic granulomatous disease | GAGs | Glycosaminoglycans |
| CHX | Cycloheximide | GalNAc | N-acetylgalactosamine |
| CI-M6PR | Cation-independent mannose 6-phosphate receptor | GlcA | Glucuronic acid |
| CLIP | Crosslinking and immunoprecipitation | GlcNAc | N-acetylglucosamine |
| CNS | Central nervous system | GNPs | Gold nanoparticles |
| CPM | Counts per million | GO | Gene ontology |
| CRISPR RNA | crRNA | GVHD | Graft versus host disease |
| CRISPR/Cas9 | Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9 | HA | Hyaluronic acid |
| | | HbA | Hemoglobin A |
| | | HbF | Fetal hemoglobin |
| | | HDR | Homology directed repair |
| | | HGMD | Human gene mutation database |
| | | hiPSCs | Human induced pluripotent stem cells |

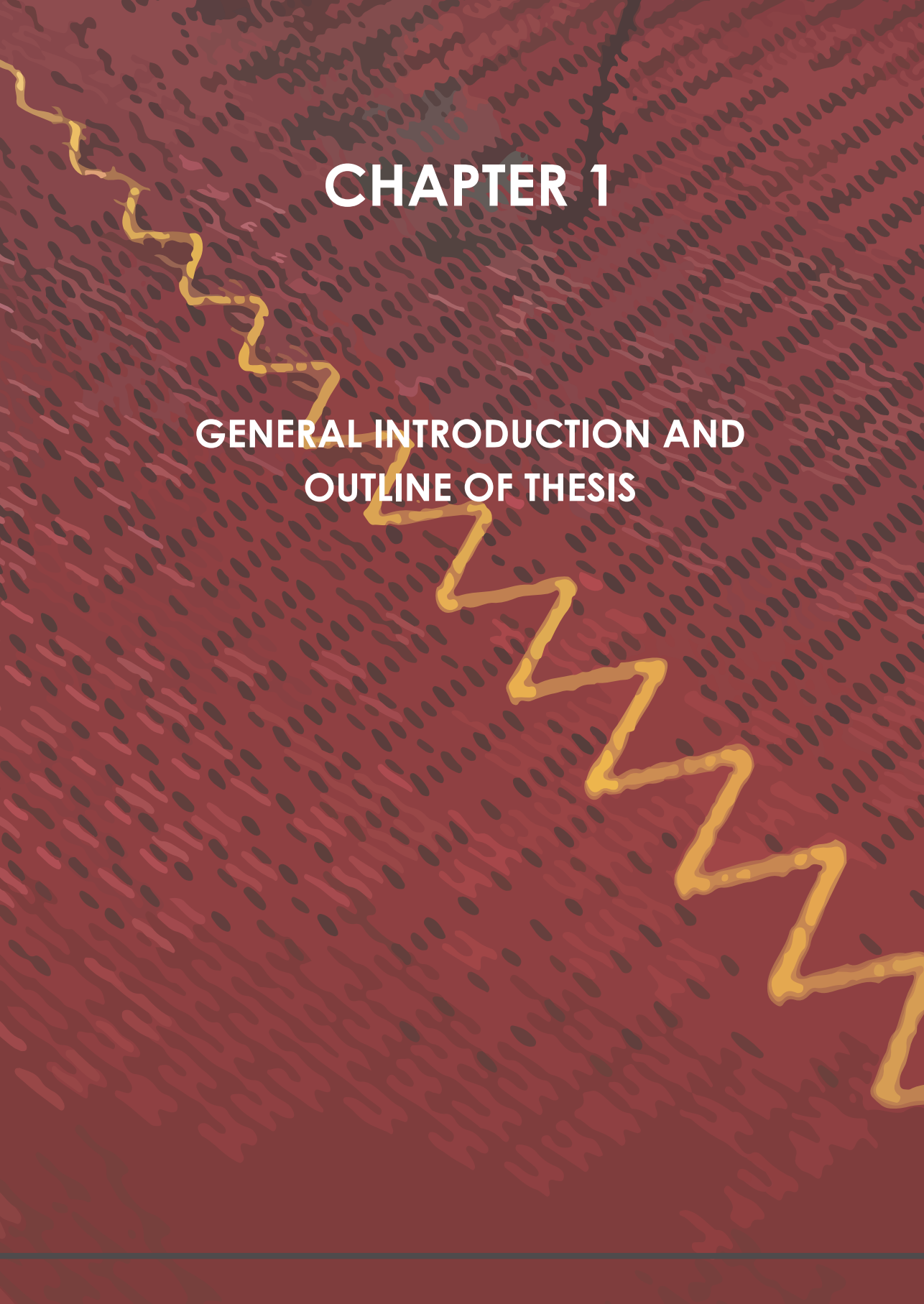
| | | | |
|----------|---|----------|---|
| HITI | Homology-independent targeted integration | MSC | Mesenchymal stem cells |
| HIV | Human immunodeficiency virus | MW | Molecular weight |
| HLA | Histocompatibility leukocyte antigen | NGD | No-go decay |
| hnRNPs | Heterogeneous nuclear ribonucleoproteins | NGS | Next-generation sequencing |
| HPFH | Hereditary persistence of HbF | NHEJ | Nonhomologous end joining |
| HPV | Human papillomavirus | NK | Natural killer |
| HS | Heparan sulphate | NMD | Nonsense-mediated decay |
| HSCs | Hematopoietic stem cells | NSD | Nonstop decay |
| HSCT | Hematopoietic stem cell transplantation | NTDT | Non transfusion-dependent thalassemia |
| HSPCs | Hematopoietic stem and progenitor cells | PAM | Protospacer-adjacent motif |
| HypaCas9 | Hyper-accurate Cas9 | PAR-CLIP | Photoactivatable-ribonucleoside enhanced CLIP |
| iCLIP | Individual-nucleotide resolution UV CLIP | PAS | Periodic acid–Schiff |
| IdoA | L-iduronic acid | PCR | Polymerase chain reaction |
| IDS | Iduronate-2-sulfatase | PCT | Pharmacological chaperone therapy |
| IDUA | A-L-iduronidase | pegRNA | Prime editing extended guide RNA |
| iPSCs | Induced pluripotent cells | PIDs | Primary immunodeficiency diseases |
| IRAEs | Immune-related adverse events | PITCh | Precise integration into target chromosome |
| ISS | Intronic splicing silencer | PKU | Phenylketonuria |
| KS | Keratan sulphate | PMO | Phosphorodiamidate morpholino oligomer |
| LCA | Leber's congenital amaurosis | PST | Protein substitution therapy |
| LNPs | Lipid nanoparticles | rAAVs | Recombinant adeno-associated viruses |
| LSDs | Lysosomal storage disorders | RNAi | RNA interference |
| M6P | Mannose-6-phosphate | RNA-seq | RNA sequencing |
| M6PR | M6P receptor | RNP | Ribonucleoprotein |
| MEFs | Mouse embryonic fibroblasts | RRM | RNA recognition motif |
| MHC | Major histocompatibility complex | RS | Recursive splicing |
| MHC | Myosin heavy chain | RSV | Rous sarcoma virus |
| miRNA | Microrna | SCD | Sickle cell disease |
| MMEJ | Microhomology-mediated end joining | SF3B | Splicing factor 3B |
| MMR | Mismatch repair | sgRNA | Single guide RNA |
| MNs | Meganucleases | shRNA | Short hairpin RNA |
| MOE | 2'O-methoxyethyl | SMA | Spinal muscular atrophy |
| MPS | Mucopolysaccharidoses | SNP | Single-nucleotide polymorphism |
| mRNA | Messenger RNA | snRNA | Small nuclear RNA |

| | |
|----------|---|
| snRNP | Small nuclear ribonuclear protein |
| SNVs | Single-nucleotide variants |
| sQTL | Splicing quantitative trait locus |
| SR | Serine/arginine-rich |
| SREs | Splicing regulatory elements |
| SRSF | SR splicing factor |
| SRT | Substrate reduction therapy |
| SSBs | Single-strand breaks |
| SSOs | Splice-switching oligonucleotides |
| TALENs | Transcription activator-like effector nucleases |
| TDT | Transfusion-dependent thalassemia |
| TILs | Tumor-infiltrating lymphocytes |
| TLR4 | Toll-like receptor 4 |
| TRAC | T cell receptor α constant |
| tracrRNA | Trans-activating CRISPR RNA |
| tTCR | Transgenic T cell receptor |
| uGAGs | Urinary gags |
| VCN | Vector copy number |
| vg/dg | Genome/diploid genome |
| WES | Whole exome sequencing |
| WGS | Whole genome sequencing |
| ZFN | Zinc finger nucleases |



CHAPTER 1

GENERAL INTRODUCTION AND OUTLINE OF THESIS



Lysosomal storage disorders

Lysosomes are organelles with a highly specialized role in the degradation and recycling of a wide range of macromolecules, including proteins, glycogen, sphingolipids and glycosaminoglycans (GAGs) [1, 2]. Transport of substrates to lysosomes is governed by the nature of the substrate itself and occurs through a variety of endocytic mechanisms. These include phagocytosis, macropinocytosis, clathrin-mediated endocytosis, caveolin-mediated endocytosis, and clathrin- and caveolin-independent endocytosis [2]. In addition, cells are able to transport their cytoplasmic intracellular materials to the lysosome through autophagy. The degradation of substrates in lysosomes is managed through a network of approximately 70 acidic hydrolases that interact with substrates in the lysosomal lumen. These lysosomal enzymes are primarily synthesized in the endoplasmic reticulum (ER) and travel to the *cis*-Golgi compartment where mannose-6-phosphate (M6P) residues are added [3]. Most lysosomal enzymes are transported to the lysosome via the M6P groups which interact with the M6P receptor (M6PR) [4].

Lysosomes act as a waste disposal system and are essential regulators of cellular homeostasis. In addition, advances in the understanding of lysosomal function have shown that lysosomes play a vital role in nutrient and amino acid sensing, and in cell growth and cell catabolism via recycling of degradation products to the cell [5, 6]. Lysosomal enzymes are therefore essential for the regulation of the complex regulatory elements involved in cellular homeostasis that are initiated within the lysosome.

Disease-associated variants in the enzymes involved in complex stepwise degradation processes in the lysosome lead to rare and severe genetic diseases known as lysosomal storage disorders (LSDs). LSDs are generally inherited as recessive traits with a cumulative incidence of approximately 1 in 5,000 live births [7-10]. Disease-associated variants in lysosomal enzymes result in their deficiency or absence, which leads to progressive lysosomal accumulation of undegraded substrate that ultimately leads to cellular and organ dysfunction [11]. LSDs are classified based on the deficient enzyme and on the accumulated substrate.

LSDs are multisystem disorders that can feature a wide range of clinical symptoms and although clinical features can vary amongst the disorders, they all share a progressive disease course. The rate of the disease progression is variable between and within LSDs. Symptoms can develop at a wide range of ages, starting in utero or in newborns, or become evident in adulthood.

Mucopolysaccharidosis

Within the LSDs, Mucopolysaccharidoses (MPS) are monogenic autosomal recessive diseases caused by deficiencies in lysosomal enzymes involved in the degradation of GAGs. GAGs are essential components in connective tissues and extracellular matrix and are continuously renewed throughout the whole body. The GAG family includes dermatan sulphate (DS), heparan sulphate (HS), keratan sulphate (KS), chondroitin sulphate (CS) and hyaluronan. All these GAGs have their own specific degradation pathways. GAGs can be degraded extracellularly or are transported to the lysosome via endocytosis where they are degraded by enzymes through a multistep process. Deficiency of any of the enzymes involved in the GAG degradation process leads to different types of MPS. The partially degraded GAG molecules are stored in lysosomes, but are also excreted into urine, a hallmark that is often used for diagnostic purposes. Currently, there are 11 known enzymes that catalyze the degradation of GAGs, and their dysfunction leads to 7 distinct MPS types (**Table 1**) [12].

Disease-associated variants in the *IDUA* gene cause MPS I and lead to the accumulation of HS and DS [13, 14]. There are three MPS I subtypes based on the disease severity: the severe Hurler form, the intermediate Hurler-Scheie form and the mild Scheie form [15]. Symptoms of MPS I include organomegaly, skeletal abnormalities and corneal clouding. Neurodegeneration occurs in Hurler, but not in Scheie [15]. MPS II – the only X-linked MPS disease – is caused by disease-associated variants in the *IDS* gene [16]. MPS II has historically been classified into the two subtypes attenuated and severe [17, 18]. More recently the two sub-types neuropathic and non-neuropathic are proposed, based on CNS involvement [19]. Incomplete GAG degradation leads to accumulation of HS and DS, which causes severe skeletal abnormalities, organomegaly and neurodegeneration [20-22]. MPS III is subdivided into 4 separate types, each with a different enzyme deficiency: A, B, C and D, caused by Heparan sulfamidase, N-acetylglucosaminidase, Heparan- α -glucosaminide N-acetyltransferase, and N-acetylglucosamine 6-sulfatase, respectively [23-27]. All subtypes feature substantial HS accumulation in organs and in the central nervous system (CNS), leading to severe diseases with progressive neurodegeneration [28-33]. MPS IV is categorized into two subtypes; A and B, caused by disease-associated variants in *GALNS* and *GLB1* respectively [34, 35]. Both subtypes show accumulation of KS, while MPS IVA also shows accumulation of chondroitin 6-sulfate (C6S). Patients with both subtypes develop severe skeletal abnormalities and organomegaly [36]. MPS IVB patients generally exhibit milder phenotypes than MPS IVA patients. There is no indication for CNS involvement in both MPS IVA and MPS IVB [37]. MPS VII patients have disease-associated variants in

the *GUSB* gene, resulting in accumulation of HS, DS and chondroitin 4-sulfate (C4S) and C6S [38]. This leads to severe organomegaly and skeletal abnormalities as well as neurodegeneration [39-42]. The extremely rare MPS IX form is caused by disease-associated variants in *HYAL1*, which leads to accumulation of hyaluronan [43, 44]. Patients present joint problems and short stature, but not with visceral or neurological involvement [43, 45]. Lastly, and the focus of this thesis, MPS VI is primarily a skeletal and joint disease caused by DS and C4S accumulation [46, 47].

| MPS type | Eponym | Enzyme deficiency | Gene | Storage product | CNS Pathology | Bone/ Cartilage Pathology |
|----------|-----------------------------------|---|---------------|------------------|----------------------------------|----------------------------|
| I | Hurler Hurler/Scheie Scheie | α -L-iduronidase | <i>IDUA</i> | HS, DS | Severe Mild to Moderate No | Severe Moderate Mild |
| II | Hunter | Iduronate-2-sulfatase | <i>IDS</i> | HS, DS | No to Severe | Moderate |
| III-A | Sanfilippo type A | heparan N-sulfatase | <i>SGSH</i> | HS | Severe | Mild |
| III-B | Sanfilippo type B | α -N-acetylglucos-aminidase | <i>NAGLU</i> | HS | Severe | Mild |
| III-C | Sanfilippo type C | Acetyl-CoA: α -glucosaminide-N-acetyltransferase | <i>HGSNAT</i> | HS | Severe | Mild |
| III-D | Sanfilippo type D | N-acetylglucosamine-6-sulfatase | <i>GNS</i> | HS | Severe | Mild |
| IV-A | Morquio type A | galactose-6-sulfatase | <i>GALNS</i> | KS, C6S | Moderate | Severe |
| IV-B | Morquio type B | β -galactosidase | <i>GLB1</i> | KS, | Moderate | Severe |
| VI | Matoteaux-Lamy | N-acetylgalactosamine-4-sulfatase | <i>ARSB</i> | DS, C4S | No | Moderate to Severe |
| VII | Sly | β -glucuronidase | <i>GUSB</i> | HS, DS, C4S, C6S | Mild | Mild to Moderate |
| IX | Natowicz syndrome | Hyaluronidase | <i>HYAL1</i> | Hyaluronan | No | Moderate* |

Table 1. Enzyme deficiencies and storage products in MPS. C4S: chondroitin 4-sulphate, C6S: chondroitin 6-sulphate, DS: dermatan sulphate, HS: heparan sulphate, KS: keratan sulphate, *based on a very small patient population.

Mucopolysaccharidosis VI

MPS VI (OMIM #253200), also known as Maroteaux-Lamy disease, is an autosomal recessive disease characterized by deficiency of the *ARSB* enzyme. *ARSB* deficiency is caused by disease-associated variants in the *ARSB* (arylsulfatase B, also called

N-acetylgalactosamine-4-sulfatase) gene [46]. This impairs degradation of DS and C4S, which leads to progressive accumulation of these GAGs in the lysosome and ultimately to progressive cellular and organ deterioration [47]. The incidence of MPS VI has been estimated between 1 in 43,261 and 1 in 1,505,160 live births [48], but has been shown to be more prevalent in some regions such as Saudi Arabia and certain parts of Brazil, influenced by higher degrees of consanguinity [49]. MPS VI patients present a clinical spectrum ranging from mildly affected patients with a slow disease progression to severely affected patients with a rapid disease progression [50, 51].

Disease progression in MPS VI

MPS VI patients present without obvious symptoms at birth with relatively normal early childhood milestones, regardless of phenotypic severity [52]. Signs and symptoms of MPS VI often begin to become apparent during early childhood. Changes in general appearance include coarse facial features such as a broad noses, thick lips and protruding tongue, large head circumference, and restricted mobility of the hand, shoulder and elbow joints [17, 21]. In addition, MPS VI patients are prone to ear infections and progressive hearing loss caused by conductive and sensorineural deficits. GAG deposition in the respiratory system can lead to upper respiratory infections, severe airway obstruction and narrowing of the trachea. These airway changes often lead to sleep apnea and further progress to reduced pulmonary function and eventually require ventilator support [17]. This is additionally compromised by changes in the thoracic skeleton, causing severe skeletal abnormalities that ultimately affect the cardio-respiratory system. GAG accumulation in the heart, liver and spleen result in cardiomyopathy, cardiac valve dysplasia and enlargement of abdominal organs (hepatosplenomegaly). Accumulation of GAGs in the cornea often leads to corneal clouding, which comes with significant vision loss. Although not typically described as a neurological disease, there is an indication for neurological involvement, but not as evident and well established as in other mucopolysaccharidoses such as I and II. MPS VI patients can develop hip dysplasia, short stature, joint pain and arthropathy with restrictions in their range of motion due to dysostosis multiplex and irregular ossification of the joints [51, 53]. Hands can become claw-like and aggravated by the frequent occurrence of carpal tunnel syndrome, whilst ribs have an unusual thickened appearance. Without treatment, death usually occurs in late childhood or adolescence for severely affected MPS VI patients. Milder affected MPS VI patients can survive into adulthood, but with a significantly reduced life expectancy. The most prominent causes of death include heart disease and airway obstruction.

Molecular biology of ARSB

ARSB is a 533 amino acid protein synthesized in the ER. After translocation of the polypeptide to the lumen of the ER, a 39 amino acid long signal peptide is cleaved off. The 55.8 kDa ARSB polypeptide acquires the mannose 6-phosphate lysosomal targeting signal. A precursor ARSB protein is formed during transport through the ER and the Golgi complex with a molecular mass of 66 kDa, which is subsequently processed to 57 kDa. Further maturation of the protein results in an amino-terminal 43 kDa, a central 7k Da species and a carboxyl-terminal 8 kDa species within the lysosome. Variants in the ARSB gene can affect any aspect of these processes, from synthesis to transport and maturation, as well as the catalytic activity. The human ARSB gene contains 8 exons and 7 introns, and is exceptionally large, spanning almost 209 kb with a coding sequence of 5,3kb. Over 200 disease-associated variants have been reported for the ARSB gene, which are spread out over the entire gene [49].

Glycosaminoglycans

The GAG family is a group of complex macromolecules involved in a large variety of important biological processes. Based on the different disaccharide units forming the GAGs, four main groups can be distinguished: heparan sulfate, chondroitin sulfate/ dermatan sulfate, keratan sulfate and hyaluronan. In proteoglycans, hyaluronan acts as a backbone to which link and core proteins are bound. HS, CS/DS or KS extend perpendicular from the core protein in a feather- like structure (**Figure 1**). GAGs can form different proteoglycans depending on the core protein. E.g. for aggrecan, a major component of articular cartilage, the protein core is encoded by the ACAN gene. Over 40 different protein cores have been identified, all with their own GAG pattern and function [54, 55]. GAGs are renewed continuously; this involves a step-wise process of synthesis and degradation, the latter occurs via a series of lysosomal enzymes.

Heparan sulfate

Heparan sulfate is an important component of both the extracellular matrix and the plasma membrane [56]. Due to the large number of potential modifications on HS chains, HS occurs in highly diverse structural forms that make these very information rich and that are key to their protein-binding and regulatory properties [57, 58]. HS interactions are crucial in cellular and micro-environmental regulation, and in maintaining cellular functions [59]. HS has also been shown to be an essential component for normal development and physiology by interacting with growth

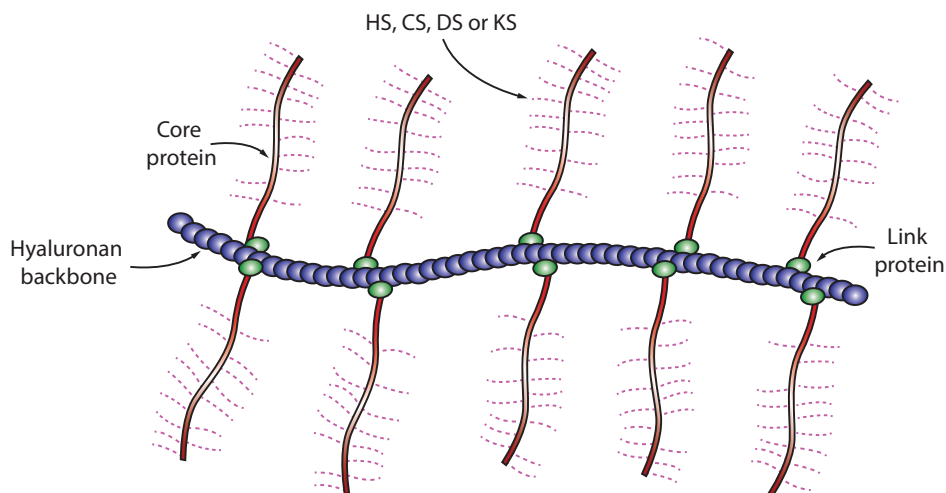


Figure 1. Schematic representation of a proteoglycan. Link (green) and core (red) proteins are connected to the backbone Hyaluronan (Blue). GAGs (pink) are covalently linked to the core protein in a feather like structure. HS: heparan sulfate, CS: chondroitin sulfate, DS: dermatan sulfate, KS: keratin sulfate.

factors, enzymes, extracellular matrix proteins and proteins on the surface of pathogens [60, 61].

HS polysaccharides are composed of repeating disaccharide units, either glucuronic acid (GlcA) or L-iduronic acid (IdoA) bound to N-acetylglucosamine (GlcNAc), which can be further modified. HS features minimally modified N-acetylated regions primarily made up of GlcA-GlcNAc repeats, which separate the highly sulphated domains – also known as S domains - from the rest of the chain. The S domains are generally 3-8 disaccharides in length and can be modified by the addition of N- and O-sulphate groups and by epimerisation of GlcA to IdoA [57, 62].

Large HS complexes are hydrolyzed by endo- β -glucuronidase that initiate the degradation (**Figure 2**). HS degradation is then followed by desulphation of the terminal iduronic acid-2-sulphate by Iduronate-2-sulfatase, making the substrate suitable for α -iduronidase. After hydrolysis, the remaining residue is converted into an N-glucosamine by Heparan N-sulphatase-mediated removal of a sulphate. The membrane-bound N-acetyltransferase catalyses the N-acetylation of the remaining amino group, leaving an N-acetylglucosamine. The terminal GlcNAc residue is enzymatically cleaved by α -N-Acetylglucosaminidase followed by cleavage of a glucuronic acid by β -glucuronidase. The remaining sulphates are removed by N-Acetylglucosamine-6-sulfatase and the final product undergoes a final hydrolysis by α -N-Acetylglucosaminidase [63].

Chondroitin sulfates & Dermatan sulfate

Chondroitin sulfates are key components in the ECM of hyaline cartilage. In addition, chondroitin sulfate plays an important role in neural signal transduction and is found in various other biological tissues such as skin, blood vessels and bone [64–67]. In chondroitin sulfates, the disaccharide unit contains N-acetylgalactosamine (GalNAc) and GlcA [62]. Chondroitin sulfates can be classified in different groups based on the position of sulfur groups that are present on the chondroitin sulfate backbone. The most prevalent C4S and C6S are sulfated on position 4 and 6 respectively [64, 68, 69]. Dermatan sulfate (DS) is a key component involved in maintaining structure and integrity in many tissues in the animal and human body such as the cornea, the heart valve, bone and cartilage [65, 66, 70, 71]. Moreover, DS is involved in biological processes such as cell division, cell signaling, wound repair and inflammation [72, 73]. Dermatan sulfate has a similar desulphation pattern as chondroitin sulfate with GalNAc and GlcA as the disaccharide units. However, unlike CS, DS can also contain IdoA residues instead of GlcA [62]. The level of complexity of the DS molecule varies greatly due to a variety of factors, such as the total length of the DS chain, the presence and location of IdoA residues, the level of sulphation and the alternatives for core proteins [70]. This provides DS with high diversity in binding affinity and controls protein interactions.

DS and C4S are related GAG molecules with a similar structure and follow the same degradation pathway, with an extra step for DS to remove the IdoA units (**Figure 3**). DS degradation is initiated by the desulphation of the terminal iduronic acid by iduronic acid-2-sulphatase, followed by another desulphation catalyzed by α -iduronidase to remove all iduronic acids. Hereafter, this substrate and C4S follow the same degradation pathway. The terminal GalNAc-4-SO₄ can be processed via two pathways. The first hydrolysis of the GalNAc-4-SO₄ is mediated by ARSB, followed by further hydrolysis of the terminal hexosamines by β -N-Acetylhexosaminidase A or B. In the second pathway, the GalNAc-4-SO₄ unit is removed by β -N-acetylhexosaminidase A followed by ARSB sulfatase cleavage. Both pathways result in a substrate that is further cleaved by β -Glucuronidase to remove the β -glucuronic acid unit from the GalNAc-4-SO₄ unit. The GalNAc-4-SO₄ unit is further cleaved by ARSB [63].

Keratan sulfate

Keratan sulfate is divided into three subtypes KS I, KS II, and KS III based on the linkage to the core protein [74, 75]. KS I is linked to an asparagine residue and found in the cornea, where it maintains the even spacing of type I collagen fibrils to allow the passage of light [74, 76]. KS II is linked to a serine or threonine residue and is found in cartilage [77, 78]. Here, KS II is linked to the large proteoglycan aggrecan

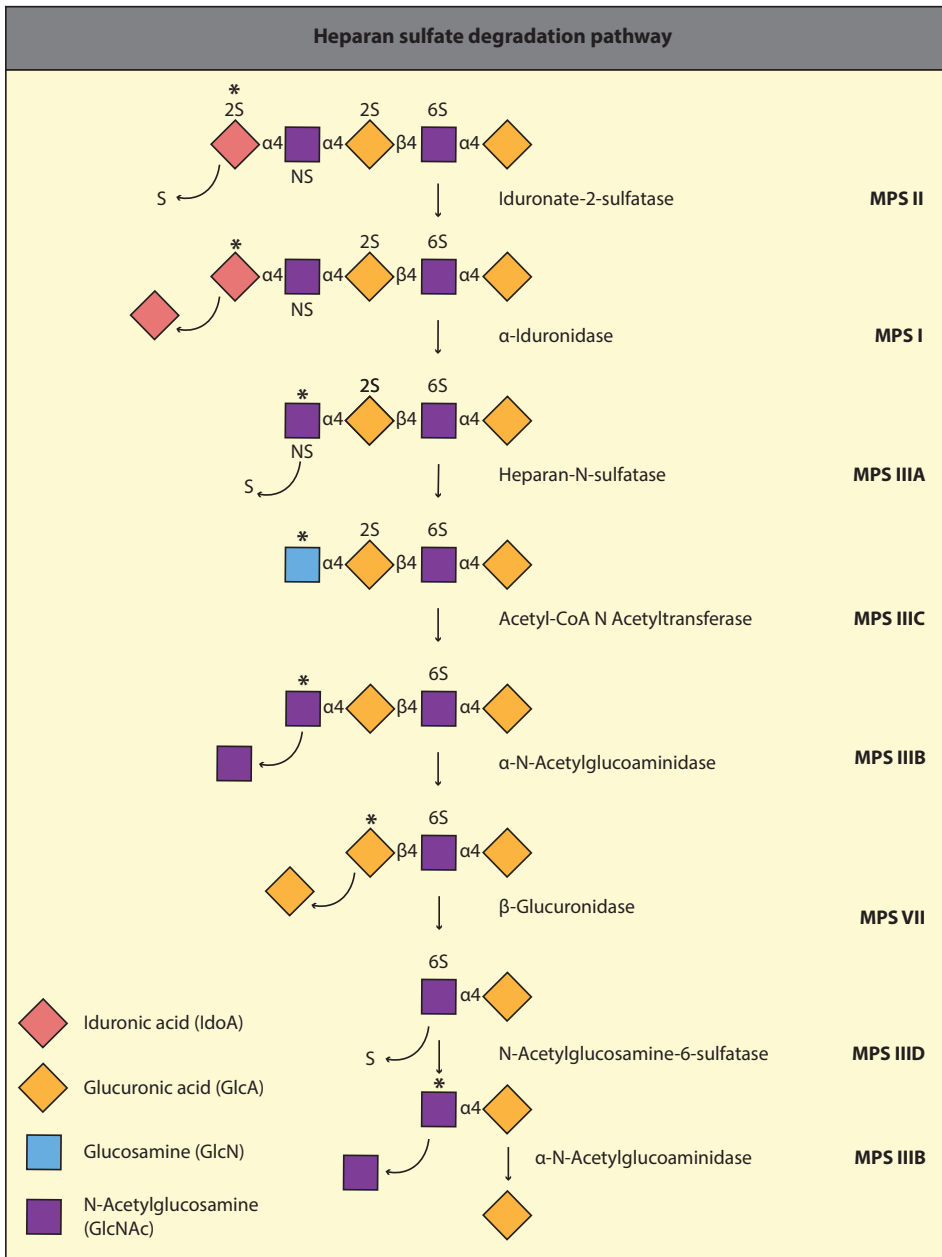


Figure 2. Heparan sulphate degradation pathway. Heparan sulphate is degraded by series of lysosomal enzymes within the lysosomes. Deficiencies of each enzyme results in specific MPS disorders, as illustrated on the right side.

and provides the properties to joints to bear significant loads and to counteract the compressive forces that occur during movement. KS III is linked to the core protein through a mannose-serine and is found in the brain, where it plays an important role in nervous system development and metabolism [79, 80].

KS is a heavily sulfated chain of repeating units of N-acetyllactosamines which can reach up to a length of 50 disaccharides and contains a mixture of nonsulfated, monosulfated and disulfated disaccharide units [62]. KS degradation is initiated with a desulphation by Galactose 6-sulfatase, followed by β -galactosidase digestion (**Figure 4**). The terminal GlcNAc-6-SO₄ is desulphated by N-Acetylglucosamine-6-sulfatase followed by an elimination of the GlcNAc group by β -N-Acetylhexosaminidase A or B. Alternatively, the GlcNAc-6-SO₄ group can be cleaved by β -N-Acetylhexosaminidase A directly. Both processes result in formation of the same residue that is processed by β -Galactosidase, and finally the remaining monosaccharide is desulphated by N-Acetylglucosamine-6-sulfatase [63].

Hyaluronan

Hyaluronic acid (HA) or hyaluronan is the largest and most predominant GAG, and the only GAG that is not sulfated [81]. HA is an essential component of the extracellular matrix of connective, epithelial and brain tissue. HA can retain water up to 100 times its weight, control water homeostasis and is able to self-aggregate or act as a backbone in proteoglycans [82-84]. In cartilage, these characteristics help to form an elastic tissue that absorbs energy and revert back to its original shape upon release of mechanical stress. HA plays a crucial biological role in wound healing in the skin, with an important role from early inflammatory activation up to the reepithelization process [81, 85-87].

HA is a polymer of repeating GlcA and GlcNAc saccharides that forms molecules of up to 25,000 disaccharide repeats in length [81]. Degradation of hyaluronan starts on the cell surface with cleavage of HA by Hyaluronidase-2 (**Figure 5**). The resulting fragments of approximately 50 disaccharides are internalized and transported to the lysosome. The fragments are degraded into tetra- and disaccharides by Hyaluronidase-1. A series of cleavages by β -Glucuronidase and β -N-Acetylhexosaminidase to remove the GlcA and GlcNAc, respectively, result in the complete breakdown of hyaluronan into monosaccharides [63].

Bone and Cartilage development

Cartilage can be classified in three types; elastic cartilage, hyaline cartilage and fibrocartilage, which differ in the composition of their extracellular matrix (ECM). Chondrocytes are the primary cells in cartilage, and due to the high ratio of matrix to cell volume in cartilage, occupy only up to 5% of the cartilage volume.

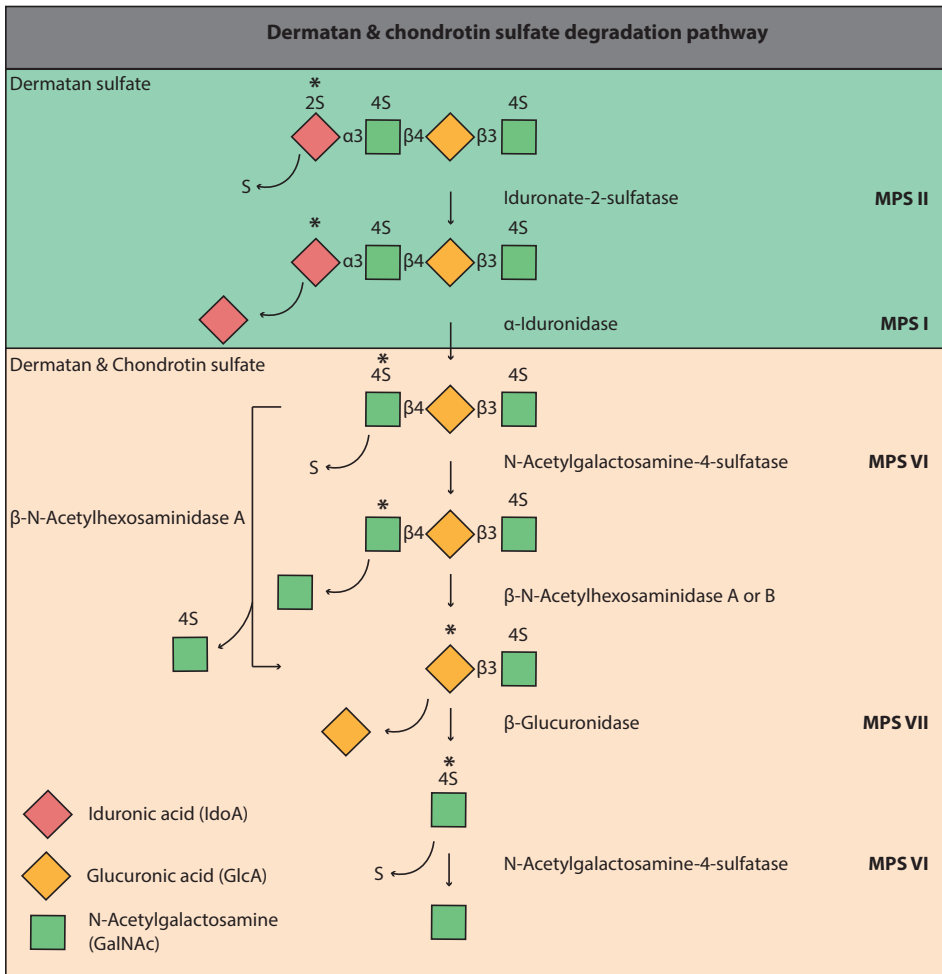


Figure 3. Chondroitin & dermatan sulphate degradation pathway. Chondroitin & dermatan sulphate are degraded by series of lysosomal enzymes within the lysosomes. Deficiencies in each enzyme results in various MPS disorders, as illustrated on the right side. Dermatan sulphate requires an additional step to remove the IdoA units. The subsequent substrate and chondroitin sulphate follow the same degradation pathway.

Chondrocytes are highly metabolically active, and synthesize and degrade the ECM components that compose cartilage. The ECM in articular cartilage (AC) is mainly made up of collagen II fibers, GAGs and proteoglycans produced by chondrocytes. The negatively charged glycans in the cartilage retain water and form a network with the collagen fibers that provide the tissue with elasticity and shock absorbing qualities, and facilitate movement in the joints. AC is an avascular tissue and chondrocytes rely on diffusion for their nutrients [88].

During development, long bones are formed through endochondral ossification, a process where cartilage tissue is replaced by bone tissue. Growth in long bones is achieved through endochondral ossification of the epiphyseal plate – or growth plate. Four distinct populations of cell types can be identified in bone: osteoprogenitor cells, osteoblasts, osteocytes and osteoclasts. Osteoprogenitors are stem cells that invade the tissue during endochondral ossification and are precursors of osteoblasts. Osteoblasts are mononucleated cells involved in bone development and can further differentiate into osteocytes, which are mature bone cells associated with the maintenance of bone. Osteoclasts are large, multinuclear cells of hematopoietic origin, involved in the resorption of bone.

Although many regulatory pathways are involved in controlling bone and cartilage histogenesis, Sox9 and Runx2 are known as the master regulatory genes. Sox9 regulates cartilage formation and directly controls transcriptional activation of other genes in its regulatory network such as Sox5, Sox6, Col2, Col9, Acan and Col11. Runx2 regulates bone formation and directly controls transcriptional activation of other genes in its regulatory network such as SP7, Col10, Ihh, Mmp13, Alpl, Spp1, Sparc and Bglap [89].

Bone and Cartilage development in Mucopolysaccharidosis

In all types of MPS one or more types of GAGs accumulate, all of which are crucial during the development, growth and maintenance of bone and cartilage. Dysregulation of these processes due to the accumulation of GAGs leads to the bone and cartilage disease, which are hallmarks of all types of MPS. Symptoms include joint problems leading to immobility and malformation of bones. Lysosomal deposition of GAGs in the chondrocytes leads to stiff joints, contractures and poor mobility, and ultimately manifests as a degenerative joint disease. Due to the hip being one of the primarily affected joints, MPS VI patients become wheelchair bound at early age. In addition, the range of motion of joints as well as hand function is decreased [51, 53].

Three main mechanisms contribute to the bone and cartilage pathology in MPS; dysregulation of signaling pathways, inflammation and the impact of mechanical forces.

The composition and metabolism of the extracellular matrix of cartilage is regulated by the degradation and de novo synthesis of its components, which are mainly mediated by osteocytes and chondrocytes in bone and cartilage, respectively. Intralysosomal GAG storage followed by cellular dysfunction provokes a cascade of pathological processes that directly affects the homeostasis of the ECM. GAG accumulation interferes with normal growth and maintenance of cartilage, which in turn leads to short stature and skeletal malformations. GAGs have

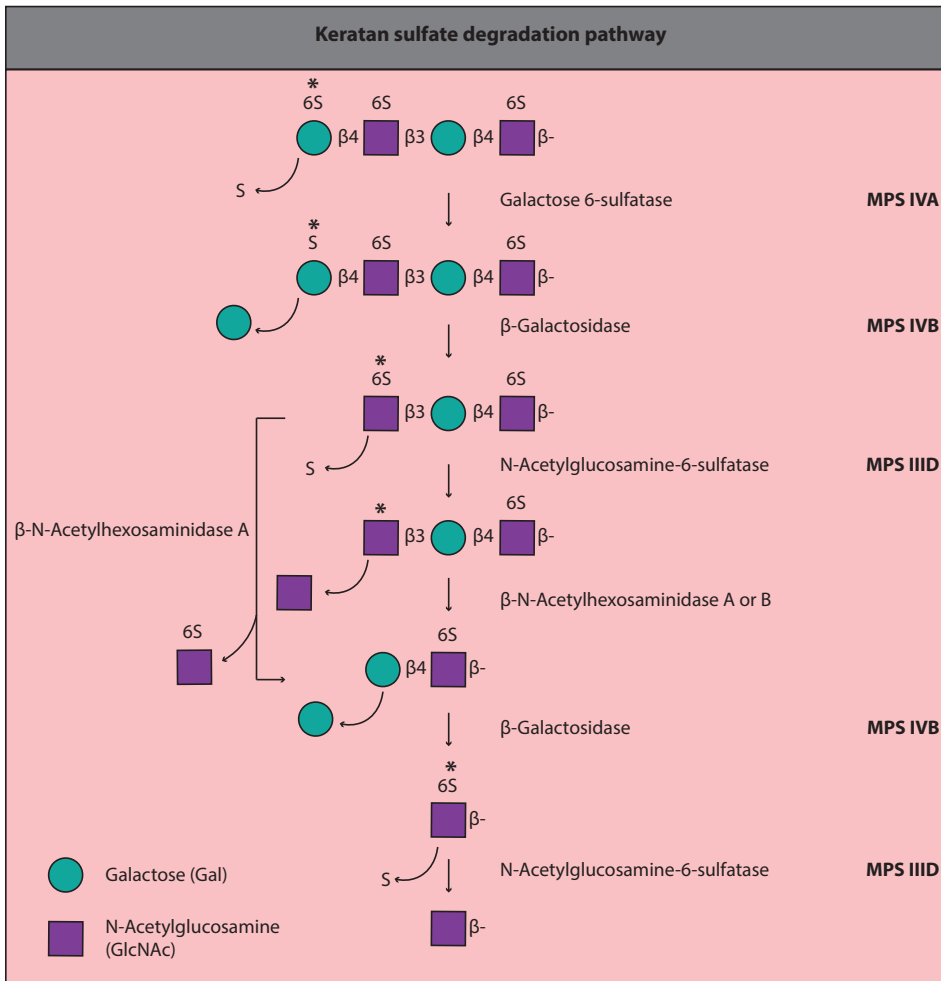


Figure 4. Keratan sulphate degradation pathway. Keratan sulphate is degraded by series of lysosomal enzymes within the lysosomes. Deficiencies of each enzyme results in various MPS disorders, as illustrated on the right side.

an important regulatory function in endochondral ossification, which is disturbed by GAG accumulation in MPS; partially degraded and undegraded GAGs interact with several growth factors, such as BMPs, TGF- β , Wnt, FGFs and the FGF receptor, all crucial factors in the formation of bone and cartilage [90].

GAG accumulation can lead to the activation of the Toll-like receptor 4 (TLR4) pathway through its interaction with GAG fragments [91, 92]. The activation of this pathway in osteoblasts, osteocytes, osteoclasts, and chondrocytes has been associated with inflammation. The inactivation of the TLR4 pathway has improved

the pathology in an MPS VII model, further highlighting the importance of the TLR4 pathway in MPS disease progression [93]. In addition, the deterioration of the joints leads to pain and further inflammation of these joints. TLR4 activation results in the secretion of pro-inflammatory cytokines such as TNF- α , IL-1 β and nitric oxide into the ECM, and bears a striking resemblance to the secretion in degenerative and inflammatory diseases such as osteo- and rheumatoid arthritis.

As the cartilage tissue degenerates, friction between the bone tissue leads to its malformation and degeneration. This changes the patients' posture, leading to further deterioration due to the alteration of the weight bearing forces in the joints and bones [51].

The processes of dysregulation of signaling pathways, inflammation and problems with mechanical loading are tightly linked and influence each other greatly. Hence, a more in-depth analysis into the interplay of these processes would greatly deepen our understanding of the events leading up to the bone and cartilage pathology.

Although all types of MPS share bone and cartilage pathology, the severity and presentation can differ between and within each MPS type. The difference between different MPS types in the involvement of bone and cartilage disease can be explained by the difference in GAG and metabolites accumulation. GAGs have an important function in cartilage biology but the prevalence and turnover of the different types of GAGs in the cartilage differ. In addition, GAG metabolites have variable binding affinities and binding partners, and can influence various key pathways involved in chondrogenesis [90].

In MPS I and MPS II, the accumulation of HS and DS leads to short stature, joint contracture, malformation of flat bones and dysostosis multiplex. The severity of the bone and cartilage pathology in MPS I depends on the subtype, with a mild involvement in Scheie to a more severe form in Hurler. Although similar products accumulate in MPS I and II, the skeletal involvement is generally less severe in MPS I. The reasons for these differences are poorly understood but is likely due to the different levels of GAG accumulation. In addition, the ratio of HS and DS accumulation may differ between the two types of MPS, which might influence the pathology. The bone and cartilage pathology are the mildest of all MPS types in patients with any of the four types of MPS III. In MPS III, the sole accumulation of HS leads to widening and thickening of bones at older age. No defects are observed in the growth of these patients as stature is not affected. Accumulation of HS is observed in chondrocytes, but not as prominently as in MPS I. The accumulation of CS and KS lead to severe bone and cartilage pathology in MPS IVA patients. Similar features are observed as in the severe MPS I patients. In addition, the accumulation of these GAGs in ligaments and adjacent connective tissue lead to hypermobile

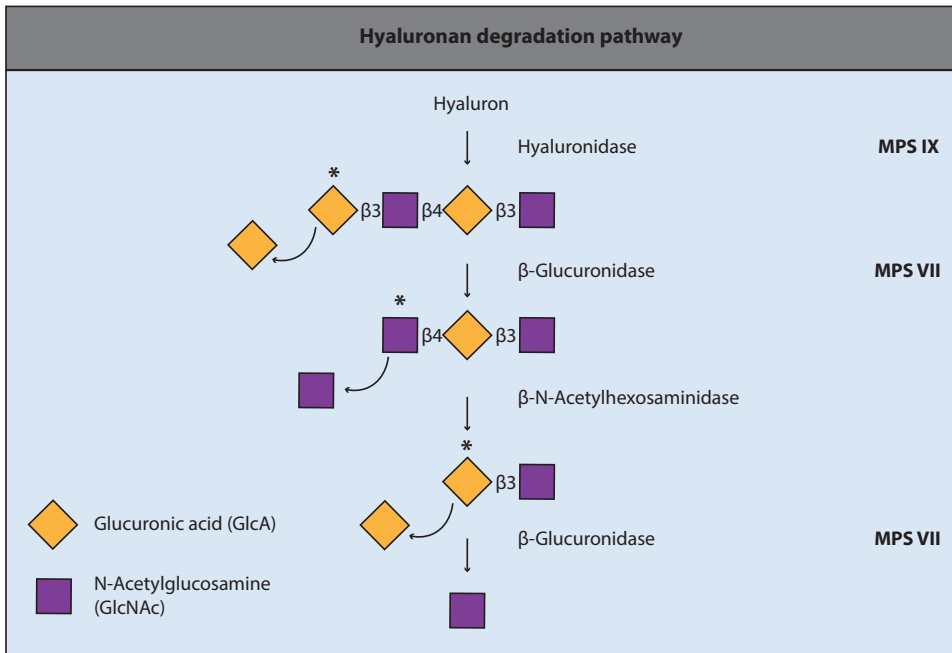


Figure 5. Hyaluronan degradation pathway. Hyaluronan is degraded by series of lysosomal enzymes within the lysosomes. Deficiencies in each enzyme results in various MPS disorders, as illustrated on the right side.

joints, a unique trait within the MPS disorders. Accumulation in the cartilage disturbs ossification of cartilage at the proximal lateral portion of the tibia and leads to fragmentation of the capital femoral epiphyses. MPS IV-B Patients usually present similar but milder clinical features compared to those associated in MPS IV-A. MPS VI is characterized by DS and CS accumulation, and bone and cartilage pathology severity varies greatly between patients. In the rapidly progressive phenotype, MPS VI leads to severe bone and cartilage pathology at a young age. In the slowly progressive phenotype, symptoms are milder and present later in life. Symptoms include the fragmentation of the capital femoral epiphyses, short stature, joint contracture, and dysostosis multiplex. Irrespective of the disease progression, hip disease is one of the most prominent features of MPS VI pathology. Due to the small patient population, little is known about bone and cartilage disease in MPS VII and IX. In MPS VII, skeletal abnormalities are often present due to the accumulation of DS, HS and CS. The bone and cartilage pathology generally seem to lead to similar symptoms and to a similar range of severity as seen in MPS I. MPS IX is the rarest form of MPS, but all patients seem to present some form of bone and cartilage pathology such as a mild short stature.

The difference in bone and cartilage pathology within the MPS types can

be explained by factors such as the severity of the disease-associated variant, the presence of a disease-modifying factor and a difference in the immune system between patients. Not all disease-associated variants have the same effect on the enzyme activity. Although all these variants lead to the typical enzyme deficiency seen in patients, the levels of residual enzyme activity can differ between variants and a genotype-phenotype relationship has been suggested for multiple types of MPS [19, 94-96]. For some types of MPS, the biochemical assays to measure the enzyme activity fail to detect any difference between the variants, most likely due to the sensitivity of these assays. For instance, in MPS II there is a strong genotype-phenotype relationship for the involvement of brain pathology, but enzyme activity assays failed to detect any difference in enzyme activity between the groups [19].

Modifying factors can influence the severity and clinical penetrance of a disease, as seen in other diseases [97-100]. Although no modifying factors have been reported to influence the clinical pathology in MPS patients so far, one cannot rule out the possible involvement of these factors. The identification of modifying factors is challenging, especially with the limited number of patients available for analysis. Although these factors might provide some insight in the difference in disease progression, they are not widely studied.

In a similar manner to the modifying factors, differences in immune responses between patients can play a vital role in the disease progression, as inflammation appears to play an important role in the establishment and disease progression in bone and cartilage.

Splicing and human disorders

Determining the disease-associated variant in MPS VI is a crucial part of diagnosis and genetic counseling. Although the identification of variants has become easier with the introduction of more advanced sequencing techniques, functional analysis of the variants is scarce and identifying the disease-associated variant can be challenging. In addition, intronic variants are often not considered although they can have a profound effect on splicing. Although the basic principles of splicing are well known, predicting the effect of a variant on the splicing outcome is still challenging due to its complex regulation. This ultimately leads to an underestimation of the percentage of splicing variants. Understanding the effect of variants on splicing is not only important for diagnostic purposes but it can be a target for the development of novel therapies such as antisense oligonucleotides and gene therapy. An in-depth review on the different types of alternative splicing, their involvement in genetic disease and possible therapeutic strategies to modulate alternative splicing can be found in **chapter 2**.

Therapy for MPS

The current therapy for most types of MPS consists of palliative care and management of clinical symptoms in the form of enzyme replacement therapy (ERT). In ERT, purified recombinant disease-specific enzyme is administered intravenously on a weekly or bi-weekly basis to prevent accumulation of substrate [17]. Lysosomal enzymes circulating in blood are endocytosed and trafficked to the lysosome via the M6P receptors on the surface of target cells [101]. ERT has been approved for MPS I (2003), MPS II (2006), MPS IVA (2015), MPS VI (2005) and MPS VII (2017).

ERT has significant positive effects on numerous parameters such as liver and spleen size, urinary GAGs and the overall quality of life [102, 103]. However, the effects of ERT on the cornea, heart valves and central nervous system are limited or absent [104]. Although ERT ameliorates bone and joint problems in various animal models if started very early, little improvement of bone and cartilage disease is observed in humans [51, 93, 105]. Moreover, due to the foreign nature of the infused enzyme, patients treated with ERT can develop neutralizing antibodies that reduce the efficacy of the therapy [106, 107].

Allogeneic hematopoietic stem cell transplantation (HSCT) is an alternative option as the standard of care for some MPS types. HSCT is a well-established stem cell therapy for MPS that uses Hematopoietic stem cells (HSCs) to reconstitute the entire lympho-haematopoietic system in the patient [108]. The therapeutic efficacy varies depending on the type of MPS, age of the patient, clinical severity, and disease stage.

Multiple patients across the MPS types have undergone HSCT with limited efficacy; although higher circulation enzyme levels led to decreased symptoms such as hepatosplenomegaly and to improved joint mobility, the therapeutic effect on the skeletal problems and brain phenotype was limited [109]. Moreover, graft versus host disease (GVHD) has been reported for 36% of all MPS VI patients that received allogeneic HSCT [109].

Although allogeneic HSCT offers a one-time intervention and is more cost efficient than ERT, it remains a controversial therapy due the limited effect on crucial targets such as the skeletal problems and the brain. Due to the limited efficacy, limited availability of donors, GVHD, mortality risk and other complications associated with allogeneic HSCT, ERT remains mostly preferred above allogeneic HSCT.

Both intravenous ERT and allogeneic HSCT have limited effects on the skeletal problems and the central nervous system, due to the poor vascularization of the tissue or due to the inability of the enzyme to cross the blood brain barrier, respectively. Therefore, the development of novel therapies to treat these symptoms is warranted.

In a pre-clinical setting, intra-articular injection of ERT reduced storage material in articular cartilage of MPS VI cats [110]. However, 1 month after injection re-accumulation occurred in the surface of the cartilage and within 2 months, GAG accumulation recurred throughout the full thickness of the articular cartilage [110]. Monthly intra-articular injections showed significant improvement after 10 months of treatment with no adverse events and negligible anti rhARSB antibody titers [111]. No outward signs of inflammation following intra-articular injection were observed in the injected joints. Although the feasibility of intra-articular injection of ERT in MPS patients is uncertain, these results do demonstrate that ARSB is able to penetrate the avascular articular cartilage if administered via injection.

Other treatment modalities for MPS disorders are currently being evaluated, including substrate reduction therapy (SRT) [112-114], anti-inflammatory therapy [115-117], stop codon read-through and pharmacological chaperone therapy (PCT) [118, 119]. Although these therapies offer new treatment modalities alone or in combination with ERT, they are not curative and/or only beneficial for a small number of patients.

Gene therapy

Gene therapy is designed to be a one-time and generic treatment with long-term benefit. In gene therapy, the missing or defective gene is introduced into patient cells *in vivo* or *ex vivo*, followed by the continuous release of this corrected enzyme into the body.

In vivo gene therapy

For *in vivo* gene therapy, the missing or deficient gene is usually transferred into the somatic cells of a patient using different viral vector systems, such as lentiviral and adeno associated virus (AAV)-based vectors. AAV vectors persist as episomal vectors and thus pose a low risk of insertional mutagenesis and genotoxicity. However, the efficiency of AAV vectors to deliver the enzyme drops over time due to their episomal nature, especially in dividing cells. A second transduction with the AAV virus is hindered due to the development of neutralizing antibodies after the first infection. In addition, a significant percentage of the population have already developed neutralizing antibodies against naturally-occurring AAV viruses and this pre-existing immunity to AAV renders patients not eligible for treatment. AAV entry is determined by the interaction of the virus capsid proteins and the receptors on the cell surface. AAV serotypes are characterized by their varying capsid proteins, all with different affinities to specific cell types, which gives this approach the potential to target the AAV virus to specific tissues.

MPS I

Promising results in relieving the brain pathology were obtained in a feline model of MPS I with intrathecally administered AAV9. Enzyme activity was detectable in parts of the brain and spinal cord, while no enzyme activity was detected in untreated MPS I cats. Administration of AAV9 improved CNS pathology and reduced GAG accumulation in peripheral tissues such as liver, spleen, heart, and lung [120]. Similar results using intrathecally administered AAV9 gene therapy were obtained in an immunologically tolerant MPS I canine model [121]. A phase I clinical trial using intravenously-administered AAV9 gene therapy designed to deliver a functional IDUA enzyme to the brain is currently ongoing [122].

MPS II

In an MPS II model mouse, intracisternal administration of AAV9 corrected neurological and systemic symptoms. Cross-correction occurred as GAG accumulation decreased in CNS lesions as well as in peripheral tissues such as liver, spleen heart, kidney, and lung in treated mice. In addition, treatment of MPS II mice corrected behavioral deficits and prolonged survival [123]. Hinderer et al. showed improved CNS pathology and long-term memory in MPS II mice after intracerebroventricular administration of AAV9 vector. Accumulation of lysosomal storage materials was decreased in the brain of treated MPS II mice and GAG accumulation was reduced in peripheral tissues such as the liver [124]. In a similar study, effectiveness of AAV9 was evaluated after intracerebroventricular injection in an MPS II mouse model. Low level of IDS activity in CNS lesions (7–40% of wild-type) was observed; however, GAG accumulation was reduced in brain, and neurologic deficits were improved [125].

Intravenous injection of AAV9 in MPS II mice led to an upregulation of circulating IDS with GAG reduction in peripheral tissues and the CNS. Survival normalized and the behavioral phenotype improved. Remarkably, GAG accumulation reduced and the survival rate normalized even with a treatment started at 9 months of age. A phase I/II clinical trial designed to treat the CNS effects in MPS II patients with intrathecally injected AAV9 gene therapy is currently ongoing [126].

MPS IIIA

Intramuscular administration of AAV1 or AAV8 resulted in low levels of circulating enzyme and did not achieve any therapeutic effect in an MPS IIIA mouse model. Conversely, intravenously-administered AAV8 led to normalized levels of circulating enzyme in female and a fourfold higher level in male mice. Lysosomal GAG accumulation was corrected in most tissues and the survival rate increased

in male mice. Interestingly, a 50% reduction of GAG accumulation in the brain in combination with a partial improvement of the brain pathology was observed in male mice, whereas no difference was observed in female mice due to the difference in AAV transduction between genders [127].

Improved therapeutic outcomes were observed in an MPS IIIA mouse model after intravenous administration of AAV2/8 carrying an engineered enzyme. The addition of a signal peptide and blood-brain barrier binding domain to increase enzyme secretion and blood-brain barrier transcytosis resulted in elevated enzyme levels in the brain and a normalized behavioral phenotype and improved brain pathology [128].

A phase I trial using intracisternal administered AAV10 was completed in 2014 [129, 130]. 4 MPS IIIA patients between 2 years and 8 months and 6 years of age were recruited and received intracerebral injections of AAV10 gene therapy via 12 needles over the course of 2 hours. Immunosuppressive treatment was started 15 days prior to and was maintained after intracerebral injection to reduce the risk of elimination of transduced cells. Data collected during the one-year follow-up showed good tolerance and no adverse effects related to the gene therapy or immunosuppressive treatment were reported. These results initiated a long-term follow-up study in 4 patients, which was completed in 2017 [131]. Patients tolerated the therapy well up to 30 months after injection. In addition, elevated enzymatic activity in the cerebrospinal fluid and an improvement in neuropsychological scores were reported 30 months after injection [132]. An updated version of the therapy was developed with a different promotor driving the transgene (CAG instead of mPGK), which resulted in a 3-fold increase in expression in the brain of MPS IIIA mice compared to the previous version. This gene therapy proved to be more effective in correcting the lysosomal pathology and reduced the inflammatory response in the mouse model. These results initiated a new phase II/III clinical trial with the improved AAV gene therapy, which is currently ongoing [133].

In a different study, a single intravenous administration of AAV9 in an MPS IIIA mouse model significantly reduced GAG accumulation throughout the central nervous system and peripheral tissues. The efficacy of the treatment was highest when the mice were treated at an early age. If treatment was initiated up to 3 months of age, the behavioral performance improved and survival rate normalized. Although GAG storage in the CNS was reduced with no further decline in the behavioral performance of older mice, their survival rate was increased but not completely normalized [134]. These results led to the initiation of two phase I/II clinical trials [135, 136].

MPS IIIB

Intracerebral administration of AAV2/5 resulted in improved biochemical and histological markers in an MPS IIIB dog model. Although the treatment did not result in the complete resolution of disease-related brain pathology, biochemical markers such as GAG levels were close to normal. The therapy was well tolerated when combined with immunosuppression [137]. These results led to the initiation of a phase I/II clinical trial, which has now been completed [138]. Results showed that intracerebral administration of AAV2/5 in combination with immunosuppressive treatment was well tolerated and resulted in upregulated circulating enzymes in the cerebrospinal fluid. All patients showed improvement of their neurodevelopmental symptoms and behavior over the follow-up period. The youngest patient close to normalized compared to unaffected children of the same age [139].

Intravenous administration of AAV9 achieved a high level of functional enzyme in the circulation with a reduction of GAG accumulation in peripheral and brain tissues in an MPS IIIB mouse model. Survival rate significantly increased and the behavioral phenotype improved after treatment [140]. In addition, most of the metabolomic abnormalities observed in MPS IIIB normalized in the serum of treated mice [141]. A phase I/II trials is currently ongoing based on this pre-clinical work [142].

MPS VI

In an MPS VI rat model, a comparison was made between intravenous and intramuscular administration of AAV with a constitutively active promotor [143]. Intravenous administration of AAV2/8 driven by a liver specific promotor was more efficacious than intramuscular administration of AAV2/1 gene therapy driven by either a muscle specific promotor or a constitutively active promotor in a feline MPS VI model. Intravenous administration resulted in elevated circulating enzymes in both the rat and feline model, however muscle transduction by intramuscular administration did not result in elevated circulating enzymes from the muscle. Although intramuscular administration of AAV with a constitutively active promotor resulted in elevated circulating enzymes, this was the result of transduction of extramuscular tissue. Intravenous administration resulted in elevated circulating enzymes in both the rat and feline model followed by a reduction of urinary GAG levels and GAG accumulation in peripheral tissues. An improvement in skeletal pathology was observed in intravenously-treated rats, while only a minor improvement was observed after intramuscular treatment. A follow-up study using intravenously-administered AAV2/8 gene therapy showed a similar improvement in skeletal pathology in the feline model [144]. These results lead to an ongoing phase I/II clinical trial based on AAV2/8 gene therapy encoding the ARSB enzyme via a

liver-specific thyroxine-binding globulin promoter [145].

Intravenous gammaretroviral gene therapy reduced the clinical manifestations in MPS VI cats. In a range of tissues, high levels of enzymes were detected with a normalization of GAG accumulation. Treated cats showed a reduction in aortic valve thickening and aortic dilatation and achieved higher body weights. In addition, articular cartilage erosion was reduced and longer appendicular skeleton lengths were reached after treatment. Although general bone disease was reduced, some aspects such as the abnormal carpal shape, bone lucency, cervical spine score and length of the cervical vertebral bone did not improve upon treatment [146].

MPS VII

A single AAV9 injection via the tail vein in an adult MPS VII mouse model led to the systemic delivery of the gene in brain and peripheral tissues, although the vector copy numbers in the CNS were lower compared to the numbers in the peripheral tissues. This treatment did not resolve the clearance of storage materials in several regions of the brain, nor did it improve the cognitive deficits. In addition, no change in survival rate was observed between treated and untreated MPS VII mice [147]. A comparison between intravenous injection and intrathecal injection of AAV9 gene therapy showed the latter to be more effective in the brain of an MPS VII canine model. Intrathecal injection significantly reduced GAGs in the CNS lesions, whereas intravenous injection was followed by an elevation of enzyme in the CNS, but no reduction of GAGs in the brain. A tail vein injection with a peptide-modified AAV2 showed positive effects on the disease pathology in MPS VII mice: cognitive deficits were reversed, GAG storage in the brain was reduced and survival was increased [147].

Neonatal retroviral treatment by intravenous injection resulted in long-term elevated enzyme levels in the circulation of MPS VII dogs [148, 149]. A complete or near-complete normalization of GAG accumulation was observed in all peripheral organs after treatment. GAG accumulation in the brain decreased after treatment, although not to WT levels [148]. The mean survival of treated dogs was increased from 0.4 years to 6.1 years [149]. Despite these positive clinical effects, enzyme analysis of the lumbar discs in treated dogs failed to detect any restoration of enzyme activity, and bone and spine disease was still present after treatment. The ventral vertebrae and ventral epiphyses remained hypoplastic and desiccation of the nucleus pulposus was observed in some disks.

Ex vivo gene therapy

For ex vivo gene therapy, recipient cells such as HSCs are isolated from the patient's

bone marrow or peripheral blood, and subsequently transduced *ex vivo* by viral vectors – most typically gammaretroviral or lentiviral - to introduce the genes encoding the lysosomal enzyme of interest. The transduced cells are transplanted into the recipient's body following a myeloablative pre-conditioning regimen, such as busulfan. Gammaretroviral or lentiviral systems can achieve a high transduction efficiency in a broad range of cells to enable long-term stable expression. These approaches are based on random integration of the transgene into the host genome, and high levels of circulating enzymes can be achieved when targeting HSCs. However, due to the random nature of viral integration, side-effects such as activation of oncogenes or inactivation of a tumor suppression gene can occur, potentially leading to tumorigenesis in the targeted cell. To reduce the risk, organizations such as the FDA recommend that the vector copy number (VCN) should remain below 5 copies per genome, and that a safe promoter that drives the transgene is selected.

MPS I

In a pre-clinical setting, lentiviral gene therapy showed a great therapeutic effect in an MPS I mouse model, albeit with a VCN of approximately 11 vector genome/diploid genome (vg/dg) in clonogenic progenitors [150]. After transduction, supraphysiological enzymes levels were achieved in the circulation as well as in brain and peripheral tissues. GAG accumulation was reduced to wild-type levels in peripheral tissues, and the number of enlarged lysosomes in brain and peripheral tissues was almost completely corrected. Skeletal abnormalities, such as the skull width and femur length, were resolved in a dose dependent manner. Since mice were treated at 8 weeks of age, it would be interesting to see whether the treatment can also reverse the pathology in older mice. A phase I/II clinical trial based on these findings is currently ongoing [151].

MPS II

In MPS II mice, lentiviral gene therapy increased enzyme levels and led to a reduction of GAGs in peripheral tissues. A slight increase of enzyme activity in the cerebrum improved the neurological outcome in treated mice, however the treatment did not completely normalize the neurological pathology. Unfortunately, the authors failed to provide the VCN of the transplanted cells. A lentiviral gene therapy strategy using an ApoEII peptide in MPS II mice showed an improved efficacy when compared to normal HSCT and gene therapy without the peptide [152]. Adding the ApoEII peptide did not negatively affect the *in vitro* intracellular enzyme activity and secretion, nor did it affect chimerism and VCN remained around 5 vg/dg. Enzyme levels in peripheral as well as in brain tissues were similar

between both gene therapy groups, whereas enzyme levels in plasma were more than three-fold higher with the ApoEII peptide compared to without the peptide. All treated groups showed a normalization of pathology markers such as heart failure markers and peripheral inflammation. All treated groups showed an improvement in skeletal abnormalities as the zygomatic arch, humerus and femur widths were normalized after treatment. Remarkably, only mice treated with the ApoEII peptide gene therapy showed a full normalization of GAG accumulation in brain tissue. The same pattern was seen for multiple other MPS II related symptoms such as behavior, astrogliosis, brain microglial activation and other brain-related pathology: only the ApoEII peptide gene therapy normalized the pathology completely.

MPS IIIA

A comparison between HSCT with donor cells originating from WT and transduced cells (donor or MPS IIIA) was made in an MPS IIIA mouse model, which showed an increased efficacy with lentiviral gene therapy. All methods resulted in elevated enzyme activity in the circulation with a close to total reduction of accumulated GAG in peripheral tissues. The survival of treated mice was significantly higher than the untreated group. With approximately 10% of enzyme activity in the brain compared to WT, mice treated with transduced donor cells showed improved neuropathology. However, with slightly lower enzymatic activity, transduced MPS IIIA cells were ineffective in correcting the neurological disease [153]. A further refinement improved the efficacy of the lentiviral gene therapy [154]. With supraphysiological levels of circulating enzymes and approximately 10% of the WT levels in the brain, GAG accumulation was normalized in peripheral as well as in brain tissue. In addition, the behavior and other neurological symptoms normalized after treatment. A pre-clinical safety and efficacy study further showed low genotoxicity and effective engraftment [155]. These results initiated an ongoing phase I/II clinical trial [156].

MPS IIIB

In a comparison between unmodified HSCT and HSCT with lentiviral gene therapy, the latter showed to be superior in treating the murine brain phenotype, although the VCN was high with approximately 7 vg/dg. *Ex vivo* lentiviral gene therapy in MPS IIIB mice showed supraphysiological levels of enzymes in peripheral tissues and the circulation, where conventional HSCT only achieved a fraction of those levels. In contrast to HSCT, an increased enzyme activity in the brain was observed with lentiviral gene therapy. Only gene therapy-treated mice showed normalization of GAG accumulation in all tissues with a correction in MPS IIIB-related neurological disease and behavior [157]. The addition of a tag designed to increase uptake into

the brain did not result in an improved phenotypical correction of the brain.

MPS VII

MPS VII mice treated with lentiviral gene therapy showed a reduction of GAG accumulation in a wide range of tissues. Mice treated at birth or at 7 weeks showed a clearance of GAGs in peripheral tissues. In addition, mice treated at birth showed a minor clearance of GAGs in brain tissue. Therapeutic intervention resulted in a correction of bone pathology in both groups: an improvement in bone surface density, cortical bone thickness, vertebral and femoral bone mineral volume was observed. However, chondrogenic pathology was unresponsive to the treatment: growth plate heights, lumbar and femoral bone lengths did not improve [158]. In a similar study using lentiviral gene therapy in MPS VII mice, bone disease improved and GAG accumulation in peripheral tissues decreased after treatment. However, cartilage pathology did not improve after treatment, and no reduction in GAG levels was observed in this tissue [159].

Gene editing

Viral based gene therapies offer a great therapeutic potential but also have drawbacks that limit their potential. Lentiviral gene therapy comes with concerns regarding insertional mutagenesis due to the random integration of the transgene. The efficacy of non-integrating gene therapies such as AAV based gene therapy drops over time due their episomal nature. In addition, existing neutralizing antibodies can reduce the therapeutic potential. Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR associated protein 9 (CRISPR/Cas9) introduced a novel system to develop gene therapy strategies and combining a cell-based therapy with CRISPR/Cas9 gene editing would offer a precise strategy with continuous expression and with reduced risk of random integration. Potential gene editing strategies include gene correction, insertion of a cDNA transgene in a safe harbor, gene knock-out and manipulation of gene-expression regulatory elements such as promoter activity or splicing. Depending on the cell target and gene editing efficiency, gene editing can in principle be developed into an *ex vivo* or *in vivo* therapy, although careful considerations need to be taken, especially for *in vivo* applications [160, 161]. In addition, gene editing offers a tool to generate and/or correct cell lines for the development of *in vitro* model systems to develop new therapies and disease modeling. Various technological advances in gene editing have been made recently, an overview of these advances and the obstacles still to overcome can be found in **chapter 4**. Although most work using gene editing is pre-clinical, advances have been made in the clinical implementation, an overview of these advances of gene editing can be found in **chapter 5**.

Aims and scope

Aim

Several *ex vivo* and *in vivo* gene therapies show promising pre-clinical results in reducing GAG accumulation in peripheral tissues and correction of abnormalities in the central nervous system for different types of MPS. Although joint and skeletal disease in MPS disorders is well documented and contribute significantly to the burden of disease experience by patients, most studies show little to no data on improvement of skeletal abnormalities. Treatment of bone and especially cartilage tissue is challenging and novel technologies should be considered to treat these severe skeletal problems. For the development of treatment options, human patient derived model systems that reflect the cartilage pathology in MPS VI are essential. It will also be important to establish the molecular consequences and the impact of genetic variants in MPS VI. This is important for genetic counseling and decision making on treatment with ERT, but also for future clinical trials to test novel therapies. The aim of this thesis is to improve the diagnosis of MPS VI, to generate an *in vitro* disease model for MPS VI based on induced pluripotent cells (iPSCs), and to apply this model for the investigation of molecular mechanisms in cartilage pathology in MPS and for the development of novel therapies to treat cartilage. An inventory of technology and clinical applications of gene editing will be made as this method plays a central role in the research described in this thesis.

Scope

In **chapter 2**, an overview is provided based on the latest literature on alternative splicing in genetic diseases, improved diagnostics based on splicing, and novel treatment options. Improving the molecular diagnosis of MPS VI with a focus on aberrant splicing is presented in **chapter 3**. We developed a novel splicing assay that resulted identification and characterization of several aberrant splicing events for ARSB.

Chapter 4 provides a literature review on gene editing with the focus on CRISPR/Cas9-mediated gene editing with an overview of the latest technological advances and the challenges to their clinical implementation. A further description of the preclinical developments and clinical trials aimed to develop treatment options for human disease using gene editing is given in **chapter 5**.

A generic gene correction method was developed using CRISPR/Cas9 in iPSCs and applied to the modeling of the lysosomal storage disorder Pompe disease in **chapter 6**. A detailed protocol for several forms of gene editing in human iPSCs is presented in **chapter 7**. In **chapter 8**, this gene editing pipeline was applied to generate isogenic iPSC lines from MPS VI patients, followed by a chondrogenic

differentiation of these cells in a serum-free differentiation protocol to investigate the molecular mechanisms involved in the cartilage pathology in MPS VI.

References

1. Settembre, C., et al., A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO J*, 2012. **31**(5): p. 1095-108.
2. Parenti, G., G. Andria, and A. Ballabio, Lysosomal storage diseases: from pathophysiology to therapy. *Annu Rev Med*, 2015. **66**: p. 471-86.
3. Braulke, T. and J.S. Bonifacino, Sorting of lysosomal proteins. *Biochim Biophys Acta*, 2009. **1793**(4): p. 605-14.
4. Reczek, D., et al., LIMP-2 is a receptor for lysosomal mannose-6-phosphate-independent targeting of beta-glucocerebrosidase. *Cell*, 2007. **131**(4): p. 770-83.
5. Zoncu, R., et al., mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+)-ATPase. *Science*, 2011. **334**(6056): p. 678-83.
6. Sancak, Y., et al., Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell*, 2010. **141**(2): p. 290-303.
7. Poorthuis, B.J., et al., The frequency of lysosomal storage diseases in The Netherlands. *Hum Genet*, 1999. **105**(1-2): p. 151-6.
8. Fletcher, J.M., Screening for lysosomal storage disorders—a clinical perspective. *J Inherit Metab Dis*, 2006. **29**(2-3): p. 405-8.
9. Jeyakumar, M., et al., Storage solutions: treating lysosomal disorders of the brain. *Nat Rev Neurosci*, 2005. **6**(9): p. 713-25.
10. Fuller, M., P.J. Meikle, and J.J. Hopwood, Epidemiology of lysosomal storage diseases: an overview. 2006.
11. Ballabio, A. and V. Gieselmann, Lysosomal disorders: from storage to cellular damage. *Biochim Biophys Acta*, 2009. **1793**(4): p. 684-96.
12. Neufeld, E.F. and J. Muenzer, The mucopolysaccharidoses. 8th edn ed. *The metabolic and molecular bases of inherited disease*, ed. B.A. Scriver CR, Sly WS, and e.C.B. Valle D, Kinzler KW, Vogelstein B, assoc. eds. 2001, New York: McGraw-Hill.
13. Scott, H.S., et al., Molecular genetics of mucopolysaccharidosis type I: diagnostic, clinical, and biological implications. *Hum Mutat*, 1995. **6**(4): p. 288-302.
14. Scott, H.S., et al., Chromosomal localization of the human alpha-L-iduronidase gene (IDUA) to 4p16.3. *Am J Hum Genet*, 1990. **47**(5): p. 802-7.
15. Vijay, S. and J.E. Wraith, Clinical presentation and follow-up of patients with the attenuated phenotype of mucopolysaccharidosis type I. *Acta Paediatr*, 2005. **94**(7): p. 872-7.
16. Hopwood, J.J., et al., Molecular basis of mucopolysaccharidosis type II: mutations in the iduronate-2-sulphatase gene. *Hum Mutat*, 1993. **2**(6): p. 435-42.
17. Wraith, J.E., et al., Mucopolysaccharidosis type II (Hunter syndrome): a clinical review and recommendations for treatment in the era of enzyme replacement therapy. *Eur*

- J Pediatr*, 2008. **167**(3): p. 267-77.
18. Burton, B.K. and R. Giugliani, Diagnosing Hunter syndrome in pediatric practice: practical considerations and common pitfalls. *Eur J Pediatr*, 2012. **171**(4): p. 631-9.
 19. Vollebregt, A.A.M., et al., Genotype-phenotype relationship in mucopolysaccharidosis II: predictive power of IDS variants for the neuronopathic phenotype. *Dev Med Child Neurol*, 2017.
 20. Kampmann, C., et al., Prevalence and characterization of cardiac involvement in Hunter syndrome. *J Pediatr*, 2011. **159**(2): p. 327-31 e2.
 21. Young, I.D. and P.S. Harper, The natural history of the severe form of Hunter's syndrome: a study based on 52 cases. *Dev Med Child Neurol*, 1983. **25**(4): p. 481-9.
 22. Holt, J., M.D. Poe, and M.L. Escolar, Early clinical markers of central nervous system involvement in mucopolysaccharidosis type II. *J Pediatr*, 2011. **159**(2): p. 320-6 e2.
 23. Scott, H.S., et al., Cloning of the sulphamidase gene and identification of mutations in Sanfilippo A syndrome. *Nat Genet*, 1995. **11**(4): p. 465-7.
 24. Zhao, H.G., et al., The molecular basis of Sanfilippo syndrome type B. *Proc Natl Acad Sci U S A*, 1996. **93**(12): p. 6101-5.
 25. Fan, X., et al., Identification of the gene encoding the enzyme deficient in mucopolysaccharidosis IIIC (Sanfilippo disease type C). *Am J Hum Genet*, 2006. **79**(4): p. 738-44.
 26. Hrebicek, M., et al., Mutations in *TMEM76** cause mucopolysaccharidosis IIIC (Sanfilippo C syndrome). *Am J Hum Genet*, 2006. **79**(5): p. 807-19.
 27. Robertson, D.A., et al., Chromosomal localization of the gene for human glucosamine-6-sulphatase to 12q14. *Hum Genet*, 1988. **79**(2): p. 175-8.
 28. Malm, G. and J.E. Mansson, Mucopolysaccharidosis type III (Sanfilippo disease) in Sweden: clinical presentation of 22 children diagnosed during a 30-year period. *Acta Paediatr*, 2010. **99**(8): p. 1253-7.
 29. Truxal, K.V., et al., A prospective one-year natural history study of mucopolysaccharidosis types IIIA and IIIB: Implications for clinical trial design. *Mol Genet Metab*, 2016. **119**(3): p. 239-248.
 30. Shapiro, E.G., et al., A Prospective Natural History Study of Mucopolysaccharidosis Type IIIA. *J Pediatr*, 2016. **170**: p. 278-87 e1-4.
 31. White, K.K., et al., Musculoskeletal manifestations of Sanfilippo Syndrome (mucopolysaccharidosis type III). *J Pediatr Orthop*, 2011. **31**(5): p. 594-8.
 32. Meyer, A., et al., Scoring evaluation of the natural course of mucopolysaccharidosis type IIIA (Sanfilippo syndrome type A). *Pediatrics*, 2007. **120**(5): p. e1255-61.
 33. Bax, M.C. and G.A. Colville, Behaviour in mucopolysaccharide disorders. *Arch Dis Child*, 1995. **73**(1): p. 77-81.
 34. Baker, E., et al., The morquio A syndrome (mucopolysaccharidosis IVA) gene maps to 16q24.3. *Am J Hum Genet*, 1993. **52**(1): p. 96-8.

35. Suzuki Y, et al., β -galactosidase deficiency (β -galactosidosis): GM1 gangliosidosis and Morquio B disease. *The Online Metabolic and Molecular Bases of Inherited Disease (OMMBID)*, ed. B.A. Valle D, Vogelstein B, Kinzler KW, Antonarakis SE, Ballabio A, Gibson K, Mitchell G, eds. 2014, New York: McGraw-Hill.
36. Sawamoto, K., et al., Mucopolysaccharidosis IVA: Diagnosis, Treatment, and Management. *Int J Mol Sci*, 2020. **21**(4).
37. K., H. and N. H., Mucopolysaccharidosis Type IVB: Clinical Features, Biochemistry, Diagnosis, Genetics and Treatment. *Mucopolysaccharidoses Update vol.1. Vol. Vol 1*. 2018, New York: Nova Science Publishers. pp. 273-283.
38. Schwartz, C.E., et al., Deletion mapping of plasminogen activator inhibitor, type I (PLANH1) and beta-glucuronidase (GUSB) in 7q21----q22. *Cytogenet Cell Genet*, 1991. **56**(3-4): p. 152-3.
39. de Kremer, R.D., et al., Mucopolysaccharidosis type VII (beta-glucuronidase deficiency): a chronic variant with an oligosymptomatic severe skeletal dysplasia. *Am J Med Genet*, 1992. **44**(2): p. 145-52.
40. Sly, W.S., et al., Beta glucuronidase deficiency: report of clinical, radiologic, and biochemical features of a new mucopolysaccharidosis. *J Pediatr*, 1973. **82**(2): p. 249-57.
41. Lee, J.E., et al., Beta-glucuronidase deficiency. A heterogeneous mucopolysaccharidosis. *Am J Dis Child*, 1985. **139**(1): p. 57-9.
42. Beaudet, A.L., et al., Variation in the phenotypic expression of beta-glucuronidase deficiency. *J Pediatr*, 1975. **86**(3): p. 388-94.
43. Natowicz, M.R., et al., Clinical and biochemical manifestations of hyaluronidase deficiency. *N Engl J Med*, 1996. **335**(14): p. 1029-33.
44. Csoka, A.B., et al., The hyaluronidase gene HYAL1 maps to chromosome 3p21.2-p21.3 in human and 9F1-F2 in mouse, a conserved candidate tumor suppressor locus. *Genomics*, 1998. **48**(1): p. 63-70.
45. Imundo, L., et al., A complete deficiency of Hyaluronoglucosaminidase 1 (HYAL1) presenting as familial juvenile idiopathic arthritis. *J Inherit Metab Dis*, 2011. **34**(5): p. 1013-22.
46. Fidzianska, E., et al., Assignment of the gene for human arylsulfatase B, ARSB, to chromosome region 5p11----5qter. *Cytogenet Cell Genet*, 1984. **38**(2): p. 150-1.
47. Litjens, T. and J.J. Hopwood, Mucopolysaccharidosis type VI: Structural and clinical implications of mutations in N-acetylgalactosamine-4-sulfatase. *Hum Mutat*, 2001. **18**(4): p. 282-95.
48. Valayannopoulos, V., et al., Mucopolysaccharidosis VI. *Orphanet J Rare Dis*, 2010. **5**: p. 5.
49. Tomanin, R., et al., Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. *Hum Mutat*, 2018.

- 39**(12): p. 1788-1802.
50. Brands, M.M., et al., Up to five years experience with 11 mucopolysaccharidosis type VI patients. *Mol Genet Metab*, 2013. **109**(1): p. 70-6.
 51. Oussoren, E., et al., A long term follow-up study of the development of hip disease in Mucopolysaccharidosis type VI. *Mol Genet Metab*, 2017. **121**(3): p. 241-251.
 52. Giugliani, R., P. Harmatz, and J.E. Wraith, Management guidelines for mucopolysaccharidosis VI. *Pediatrics*, 2007. **120**(2): p. 405-18.
 53. Oussoren, E., et al., Bone, joint and tooth development in mucopolysaccharidoses: relevance to therapeutic options. *Biochim Biophys Acta*, 2011. **1812**(11): p. 1542-56.
 54. Prydz, K., Determinants of Glycosaminoglycan (GAG) Structure. *Biomolecules*, 2015. **5**(3): p. 2003-22.
 55. Iozzo, R.V. and L. Schaefer, Proteoglycan form and function: A comprehensive nomenclature of proteoglycans. *Matrix Biol*, 2015. **42**: p. 11-55.
 56. Gallagher, J.T., M. Lyon, and W.P. Steward, Structure and function of heparan sulphate proteoglycans. *Biochem J*, 1986. **236**(2): p. 313-25.
 57. Turnbull, J.E., et al., Glycomics profiling of heparan sulfate structure and activity. *Methods Enzymol*, 2010. **480**: p. 65-85.
 58. Nugent, M.A., J. Zaia, and J.L. Spencer, Heparan sulfate-protein binding specificity. *Biochemistry (Mosc)*, 2013. **78**(7): p. 726-35.
 59. Bissell, M.J., H.G. Hall, and G. Parry, How does the extracellular matrix direct gene expression? *J Theor Biol*, 1982. **99**(1): p. 31-68.
 60. Lander, A.D. and S.B. Selleck, The elusive functions of proteoglycans: in vivo veritas. *J Cell Biol*, 2000. **148**(2): p. 227-32.
 61. Ori, A., M.C. Wilkinson, and D.G. Fernig, The heparanome and regulation of cell function: structures, functions and challenges. *Front Biosci*, 2008. **13**: p. 4309-38.
 62. Lindahl, U., et al., Proteoglycans and Sulfated Glycosaminoglycans. *Essentials of Glycobiology*. 3rd edition, ed. A. Varki, et al. 2017, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017.
 63. Freeze, H.H., T. Kinoshita, and R.L. Schnaar, Genetic Disorders of Glycan Degradation. *Essentials of Glycobiology*. 3rd edition, ed. A. Varki, et al. 2017, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017.
 64. Mikami, T. and H. Kitagawa, Biosynthesis and function of chondroitin sulfate. *Biochim Biophys Acta*, 2013. **1830**(10): p. 4719-33.
 65. Iozzo, R.V., The biology of the small leucine-rich proteoglycans. *Functional network of interactive proteins*. *J Biol Chem*, 1999. **274**(27): p. 18843-6.
 66. Nareyeck, G., et al., Differential interactions of decorin and decorin mutants with type I and type VI collagens. *Eur J Biochem*, 2004. **271**(16): p. 3389-98.
 67. Mizumoto, S., S. Yamada, and K. Sugahara, Molecular interactions between chondroitin-dermatan sulfate and growth factors/receptors/matrix proteins. *Current*

- Opinion in Structural Biology, 2015. **34**: p. 35-42.
68. Cheng, F., et al., Patterns of uronosyl epimerization and 4-/6-O-sulphation in chondroitin/dermatan sulphate from decorin and biglycan of various bovine tissues. *Glycobiology*, 1994. **4**(5): p. 685-96.
 69. Choi, H.U., et al., Characterization of the dermatan sulfate proteoglycans, DS-PGI and DS-PGII, from bovine articular cartilage and skin isolated by octyl-sepharose chromatography. *J Biol Chem*, 1989. **264**(5): p. 2876-84.
 70. Trowbridge, J.M. and R.L. Gallo, Dermatan sulfate: new functions from an old glycosaminoglycan. *Glycobiology*, 2002. **12**(9): p. 117R-25R.
 71. Lujan, T.J., et al., Effect of dermatan sulfate glycosaminoglycans on the quasi-static material properties of the human medial collateral ligament. *J Orthop Res*, 2007. **25**(7): p. 894-903.
 72. Taylor, K.R., J.A. Rudisill, and R.L. Gallo, Structural and sequence motifs in dermatan sulfate for promoting fibroblast growth factor-2 (FGF-2) and FGF-7 activity. *J Biol Chem*, 2005. **280**(7): p. 5300-6.
 73. Osborne, S.A., et al., Antithrombin activity and disaccharide composition of dermatan sulfate from different bovine tissues. *Glycobiology*, 2008. **18**(3): p. 225-34.
 74. Funderburgh, J.L., Keratan sulfate: structure, biosynthesis, and function. *Glycobiology*, 2000. **10**(10): p. 951-8.
 75. Funderburgh, J.L., Keratan sulfate biosynthesis. *IUBMB Life*, 2002. **54**(4): p. 187-94.
 76. Sommarin, Y., et al., Osteoadherin, a cell-binding keratan sulfate proteoglycan in bone, belongs to the family of leucine-rich repeat proteins of the extracellular matrix. *J Biol Chem*, 1998. **273**(27): p. 16723-9.
 77. Nieduszynski, I.A., et al., Structural aspects of skeletal keratan sulphates. *Biochem Soc Trans*, 1990. **18**(5): p. 792-3.
 78. Nieduszynski, I.A., et al., There are two major types of skeletal keratan sulphates. *Biochem J*, 1990. **271**(1): p. 243-5.
 79. Krusius, T., et al., Identification of an O-glycosidic mannose-linked sialylated tetrasaccharide and keratan sulfate oligosaccharides in the chondroitin sulfate proteoglycan of brain. *J Biol Chem*, 1986. **261**(18): p. 8237-42.
 80. Chai, W., et al., High prevalence of 2-mono- and 2,6-di-substituted manol-terminating sequences among O-glycans released from brain glycopeptides by reductive alkaline hydrolysis. *Eur J Biochem*, 1999. **263**(3): p. 879-88.
 81. Hascall, V. and J.D. Esko, *Hyaluronan. Essentials of Glycobiology*. 3rd edition, ed. A. Varki, et al. 2017, Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017.
 82. Cowman, M.K. and S. Matsuoka, Experimental approaches to hyaluronan structure. *Carbohydr Res*, 2005. **340**(5): p. 791-809.
 83. Lee, J.Y. and A.P. Spicer, *Hyaluronan: a multifunctional, megaDalton, stealth*

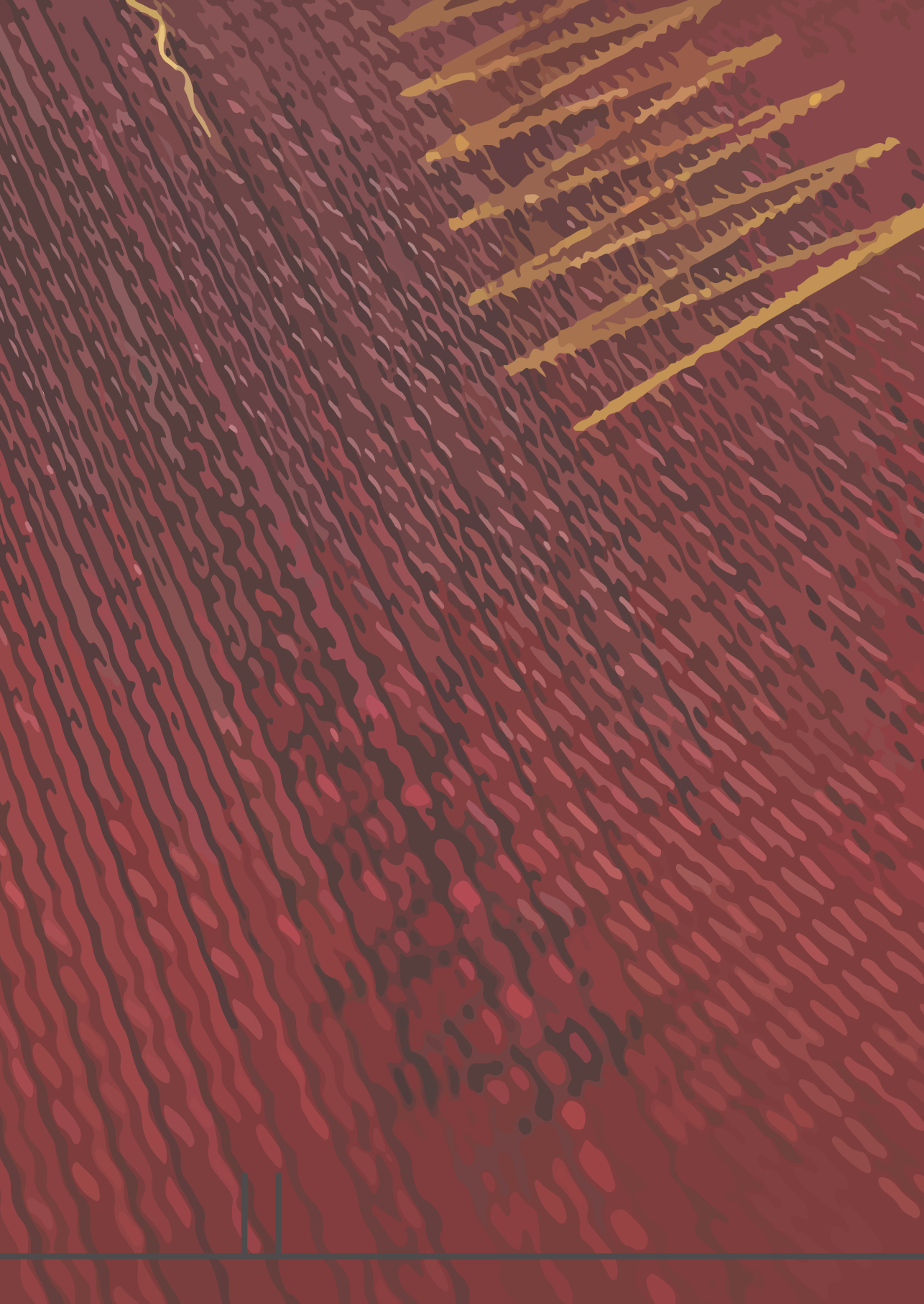
- molecule. *Curr Opin Cell Biol*, 2000. **12**(5): p. 581-6.
84. Couchman, J.R. and C.A. Pataki, An introduction to proteoglycans and their localization. *J Histochem Cytochem*, 2012. **60**(12): p. 885-97.
 85. Aya, K.L. and R. Stern, Hyaluronan in wound healing: rediscovering a major player. *Wound Repair Regen*, 2014. **22**(5): p. 579-93.
 86. Slevin, M., S. Kumar, and J. Gaffney, Angiogenic oligosaccharides of hyaluronan induce multiple signaling pathways affecting vascular endothelial cell mitogenic and wound healing responses. *J Biol Chem*, 2002. **277**(43): p. 41046-59.
 87. Weigel, P.H., G.M. Fuller, and R.D. LeBoeuf, A model for the role of hyaluronic acid and fibrin in the early events during the inflammatory response and wound healing. *J Theor Biol*, 1986. **119**(2): p. 219-34.
 88. Poole, C.A., Articular cartilage chondrons: form, function and failure. *J Anat*, 1997. **191 (Pt 1)**: p. 1-13.
 89. Gomez-Picos, P. and B.F. Eames, On the evolutionary relationship between chondrocytes and osteoblasts. *Front Genet*, 2015. **6**: p. 297.
 90. Zhang, L., Glycosaminoglycan (GAG) biosynthesis and GAG-binding proteins. *Prog Mol Biol Transl Sci*, 2010. **93**: p. 1-17.
 91. Simonaro, C.M., et al., Mechanism of glycosaminoglycan-mediated bone and joint disease: implications for the mucopolysaccharidoses and other connective tissue diseases. *Am J Pathol*, 2008. **172**(1): p. 112-22.
 92. Ausseil, J., et al., Early neurodegeneration progresses independently of microglial activation by heparan sulfate in the brain of mucopolysaccharidosis IIIB mice. *PLoS One*, 2008. **3**(5): p. e2296.
 93. Simonaro, C.M., et al., Involvement of the Toll-like receptor 4 pathway and use of TNF-alpha antagonists for treatment of the mucopolysaccharidoses. *Proc Natl Acad Sci U S A*, 2010. **107**(1): p. 222-7.
 94. Bertola, F., et al., IDUA mutational profiling of a cohort of 102 European patients with mucopolysaccharidosis type I: identification and characterization of 35 novel alpha-L-iduronidase (IDUA) alleles. *Hum Mutat*, 2011. **32**(6): p. E2189-210.
 95. Lee, O.J., et al., A study of the relationship between clinical phenotypes and plasma iduronate-2-sulfatase enzyme activities in Hunter syndrome patients. *Korean J Pediatr*, 2012. **55**(3): p. 88-92.
 96. Oussoren, E., et al., Residual alpha-L-iduronidase activity in fibroblasts of mild to severe Mucopolysaccharidosis type I patients. *Mol Genet Metab*, 2013. **109**(4): p. 377-81.
 97. Bergsma, A.J., et al., A genetic modifier of symptom onset in Pompe disease. *EBioMedicine*, 2019. **43**: p. 553-561.
 98. Weiss, G., Genetic mechanisms and modifying factors in hereditary hemochromatosis. *Nat Rev Gastroenterol Hepatol*, 2010. **7**(1): p. 50-8.

99. Cao, A. and P. Moi, Genetic modifying factors in beta-thalassemia. *Clin Chem Lab Med*, 2000. **38**(2): p. 123-32.
100. Liao, H., et al., Filaggrin mutations are genetic modifying factors exacerbating X-linked ichthyosis. *J Invest Dermatol*, 2007. **127**(12): p. 2795-8.
101. Fratantoni, J.C., C.W. Hall, and E.F. Neufeld, Hurler and Hunter syndromes: mutual correction of the defect in cultured fibroblasts. *Science*, 1968. **162**(3853): p. 570-2.
102. Harmatz, P., et al., Enzyme replacement therapy for mucopolysaccharidosis VI: evaluation of long-term pulmonary function in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J Inherit Metab Dis*, 2010. **33**(1): p. 51-60.
103. Gomes, D.F., et al., Clinical effectiveness of enzyme replacement therapy with galsulfase in mucopolysaccharidosis type VI treatment: Systematic review. *J Inherit Metab Dis*, 2019. **42**(1): p. 66-76.
104. Chen, H.H., et al., Enzyme replacement therapy for mucopolysaccharidoses; past, present, and future. *J Hum Genet*, 2019. **64**(11): p. 1153-1171.
105. Simonaro, C.M., Cartilage and chondrocyte pathology in the mucopolysaccharidoses: The role of glycosaminoglycan-mediated inflammation. *J Pediatr Rehabil Med*, 2010. **3**(2): p. 85-8.
106. Brands, M.M., et al., Mucopolysaccharidosis type VI phenotypes-genotypes and antibody response to galsulfase. *Orphanet J Rare Dis*, 2013. **8**: p. 51.
107. Ponder, K.P., Immune response hinders therapy for lysosomal storage diseases. *J Clin Invest*, 2008. **118**(8): p. 2686-9.
108. Wynn, R., Stem cell transplantation in inherited metabolic disorders. *Hematology Am Soc Hematol Educ Program*, 2011. **2011**: p. 285-91.
109. Turbeville, S., et al., Clinical outcomes following hematopoietic stem cell transplantation for the treatment of mucopolysaccharidosis VI. *Mol Genet Metab*, 2011. **102**(2): p. 111-5.
110. Auclair, D., et al., Intra-articular enzyme administration for joint disease in feline mucopolysaccharidosis VI: enzyme dose and interval. *Pediatr Res*, 2006. **59**(4 Pt 1): p. 538-43.
111. Auclair, D., et al., Long-term intra-articular administration of recombinant human N-acetylgalactosamine-4-sulfatase in feline mucopolysaccharidosis VI. *Mol Genet Metab*, 2007. **91**(4): p. 352-61.
112. Malinowska, M., et al., Genistein improves neuropathology and corrects behaviour in a mouse model of neurodegenerative metabolic disease. *PLoS One*, 2010. **5**(12): p. e14192.
113. Piotrowska, E., et al., Genistein-mediated inhibition of glycosaminoglycan synthesis as a basis for gene expression-targeted isoflavone therapy for mucopolysaccharidoses. *Eur J Hum Genet*, 2006. **14**(7): p. 846-52.
114. Friso, A., et al., Genistein reduces glycosaminoglycan levels in a mouse model of

- mucopolysaccharidosis type II. *Br J Pharmacol*, 2010. **159**(5): p. 1082-91.
115. Eliyahu, E., et al., Anti-TNF-alpha therapy enhances the effects of enzyme replacement therapy in rats with mucopolysaccharidosis type VI. *PLoS One*, 2011. **6**(8): p. e22447.
 116. Frohbergh, M., et al., Dose responsive effects of subcutaneous pentosan polysulfate injection in mucopolysaccharidosis type VI rats and comparison to oral treatment. *PLoS One*, 2014. **9**(6): p. e100882.
 117. Schuchman, E.H., et al., Pentosan polysulfate: a novel therapy for the mucopolysaccharidoses. *PLoS One*, 2013. **8**(1): p. e54459.
 118. Hein, L.K., et al., alpha-L-iduronidase premature stop codons and potential read-through in mucopolysaccharidosis I patients. *J Mol Biol*, 2004. **338**(3): p. 453-62.
 119. Brooks, D.A., V.J. Muller, and J.J. Hopwood, Stop-codon read-through for patients affected by a lysosomal storage disorder. *Trends Mol Med*, 2006. **12**(8): p. 367-73.
 120. Hinderer, C., et al., Intrathecal gene therapy corrects CNS pathology in a feline model of mucopolysaccharidosis I. *Mol Ther*, 2014. **22**(12): p. 2018-27.
 121. Hinderer, C., et al., Neonatal tolerance induction enables accurate evaluation of gene therapy for MPS I in a canine model. *Mol Genet Metab*, 2016. **119**(1-2): p. 124-30.
 122. Regenxbio, I., RGX-111 Gene Therapy in Patients With MPS I. 2021.
 123. Motas, S., et al., CNS-directed gene therapy for the treatment of neurologic and somatic mucopolysaccharidosis type II (Hunter syndrome). *JCI Insight*, 2016. **1**(9): p. e86696.
 124. Hinderer, C., et al., Delivery of an Adeno-Associated Virus Vector into Cerebrospinal Fluid Attenuates Central Nervous System Disease in Mucopolysaccharidosis Type II Mice. *Hum Gene Ther*, 2016. **27**(11): p. 906-915.
 125. Laoharawee, K., et al., Prevention of Neurocognitive Deficiency in Mucopolysaccharidosis Type II Mice by Central Nervous System-Directed, AAV9-Mediated Iduronate Sulfatase Gene Transfer. *Hum Gene Ther*, 2017. **28**(8): p. 626-638.
 126. Regenxbio, I., RGX-121 Gene Therapy in Patients With MPS II (Hunter Syndrome). 2021.
 127. Ruzo, A., et al., Liver production of sulfamidase reverses peripheral and ameliorates CNS pathology in mucopolysaccharidosis IIIA mice. *Mol Ther*, 2012. **20**(2): p. 254-66.
 128. Sorrentino, N.C., et al., A highly secreted sulphamidase engineered to cross the blood-brain barrier corrects brain lesions of mice with mucopolysaccharidoses type IIIA. *EMBO Mol Med*, 2013. **5**(5): p. 675-90.
 129. Tardieu, M., et al., Intracerebral administration of adeno-associated viral vector serotype rh.10 carrying human SGSH and SUMF1 cDNAs in children with mucopolysaccharidosis type IIIA disease: results of a phase I/II trial. *Hum Gene Ther*, 2014. **25**(6): p. 506-16.
 130. Lysogene, Intracerebral Gene Therapy for Sanfilippo Type A Syndrome. 2013.

131. Lysogene, Long-term Follow-up of Sanfilippo Type A Patients Treated by Intracerebral SAF-301 Gene Therapy. 2017.
132. Tardieu, M., et al., Intracerebral administration of rAAV2/5hNAGLU vector in children with MPS IIIB: results at 30 months of a phase I/II trial. *Molecular Genetics and Metabolism*, 2017. **120**(1): p. S130.
133. Lysogene, Study of AAVrh10-h.SGSH Gene Therapy in Patients With Mucopolysaccharidosis Type IIIA (MPS IIIA). 2022.
134. Fu, H., et al., Functional correction of neurological and somatic disorders at later stages of disease in MPS IIIA mice by systemic scAAV9-hSGSH gene delivery. *Mol Ther Methods Clin Dev*, 2016. **3**: p. 16036.
135. Abeona Therapeutics, I., Phase I/II Gene Transfer Clinical Trial of scAAV9.U1a.hSGSH. 2022.
136. Abeona Therapeutics, I., Gene Transfer Study of ABO-102 in Patients With Middle and Advanced Phases of MPS IIIA Disease. 2022.
137. Ellinwood, N.M., et al., Safe, efficient, and reproducible gene therapy of the brain in the dog models of Sanfilippo and Hurler syndromes. *Mol Ther*, 2011. **19**(2): p. 251-9.
138. UniQure Biopharma, B.V., S. Venn Life, and P. Institut, Intracerebral Gene Therapy in Children With Sanfilippo Type B Syndrome. 2019.
139. Tardieu, M., et al., Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: an uncontrolled phase 1/2 clinical trial. *Lancet Neurol*, 2017. **16**(9): p. 712-720.
140. Fu, H., et al., Correction of neurological disease of mucopolysaccharidosis IIIB in adult mice by rAAV9 trans-blood-brain barrier gene delivery. *Mol Ther*, 2011. **19**(6): p. 1025-33.
141. Fu, H., et al., Near-Complete Correction of Profound Metabolomic Impairments Corresponding to Functional Benefit in MPS IIIB Mice after IV rAAV9-hNAGLU Gene Delivery. *Mol Ther*, 2017. **25**(3): p. 792-802.
142. Abeona Therapeutics, I., Gene Transfer Clinical Trial for Mucopolysaccharidosis (MPS) IIIB. 2022.
143. Tessitore, A., et al., Biochemical, pathological, and skeletal improvement of mucopolysaccharidosis VI after gene transfer to liver but not to muscle. *Mol Ther*, 2008. **16**(1): p. 30-7.
144. Cotugno, G., et al., Long-term amelioration of feline Mucopolysaccharidosis VI after AAV-mediated liver gene transfer. *Mol Ther*, 2011. **19**(3): p. 461-9.
145. Fondazione, T., Gene Therapy in Patients With Mucopolysaccharidosis Disease. 2020.
146. Ponder, K.P., et al., Neonatal gene therapy with a gamma retroviral vector in mucopolysaccharidosis VI cats. *Mol Ther*, 2012. **20**(5): p. 898-907.
147. Chen, Y.H., et al., Sialic acid deposition impairs the utility of AAV9, but not peptide-modified AAVs for brain gene therapy in a mouse model of lysosomal storage

- disease. *Mol Ther*, 2012. **20**(7): p. 1393-9.
148. Wang, B., et al., Expression in blood cells may contribute to biochemical and pathological improvements after neonatal intravenous gene therapy for mucopolysaccharidosis VII in dogs. *Molecular Genetics and Metabolism*, 2006. **87**(1): p. 8-21.
 149. Smith, L.J., et al., Effect of neonatal gene therapy on lumbar spine disease in mucopolysaccharidosis VII dogs. *Mol Genet Metab*, 2012. **107**(1-2): p. 145-52.
 150. Visigalli, I., et al., Gene therapy augments the efficacy of hematopoietic cell transplantation and fully corrects mucopolysaccharidosis type I phenotype in the mouse model. *Blood*, 2010. **116**(24): p. 5130-9.
 151. Raffaele, I.S. and T. Fondazione, Gene Therapy With Modified Autologous Hematopoietic Stem Cells for the Treatment of Patients With Mucopolysaccharidosis Type I, Hurler Variant. 2021.
 152. Gleitz, H.F., et al., Brain-targeted stem cell gene therapy corrects mucopolysaccharidosis type II via multiple mechanisms. *EMBO Mol Med*, 2018. **10**(7).
 153. Langford-Smith, A., et al., Hematopoietic stem cell and gene therapy corrects primary neuropathology and behavior in mucopolysaccharidosis IIIA mice. *Mol Ther*, 2012. **20**(8): p. 1610-21.
 154. Sergijenko, A., et al., Myeloid/Microglial driven autologous hematopoietic stem cell gene therapy corrects a neuronopathic lysosomal disease. *Mol Ther*, 2013. **21**(10): p. 1938-49.
 155. Ellison, S.M., et al., Pre-clinical Safety and Efficacy of Lentiviral Vector-Mediated Ex Vivo Stem Cell Gene Therapy for the Treatment of Mucopolysaccharidosis IIIA. *Mol Ther Methods Clin Dev*, 2019. **13**: p. 399-413.
 156. University of, M., et al., Gene Therapy With Modified Autologous Hematopoietic Stem Cells for Patients With Mucopolysaccharidosis Type IIIA. 2024.
 157. Holley, R.J., et al., Macrophage enzyme and reduced inflammation drive brain correction of mucopolysaccharidosis IIIB by stem cell gene therapy. *Brain*, 2018. **141**(1): p. 99-116.
 158. Macsai, C.E., et al., Skeletal response to lentiviral mediated gene therapy in a mouse model of MPS VII. *Mol Genet Metab*, 2012. **106**(2): p. 202-13.
 159. Derrick-Roberts, A.L., et al., Reversal of established bone pathology in MPS VII mice following lentiviral-mediated gene therapy. *Mol Genet Metab*, 2016. **119**(3): p. 249-257.
 160. Broeders, M., et al., Sharpening the Molecular Scissors: Advances in Gene-Editing Technology. *iScience*, 2019. **23**(1): p. 100789.
 161. Ernst, M.P.T., et al., Ready for Repair? Gene Editing Enters the Clinic for the Treatment of Human Disease. *Mol Ther Methods Clin Dev*, 2020. **18**: p. 532-557.

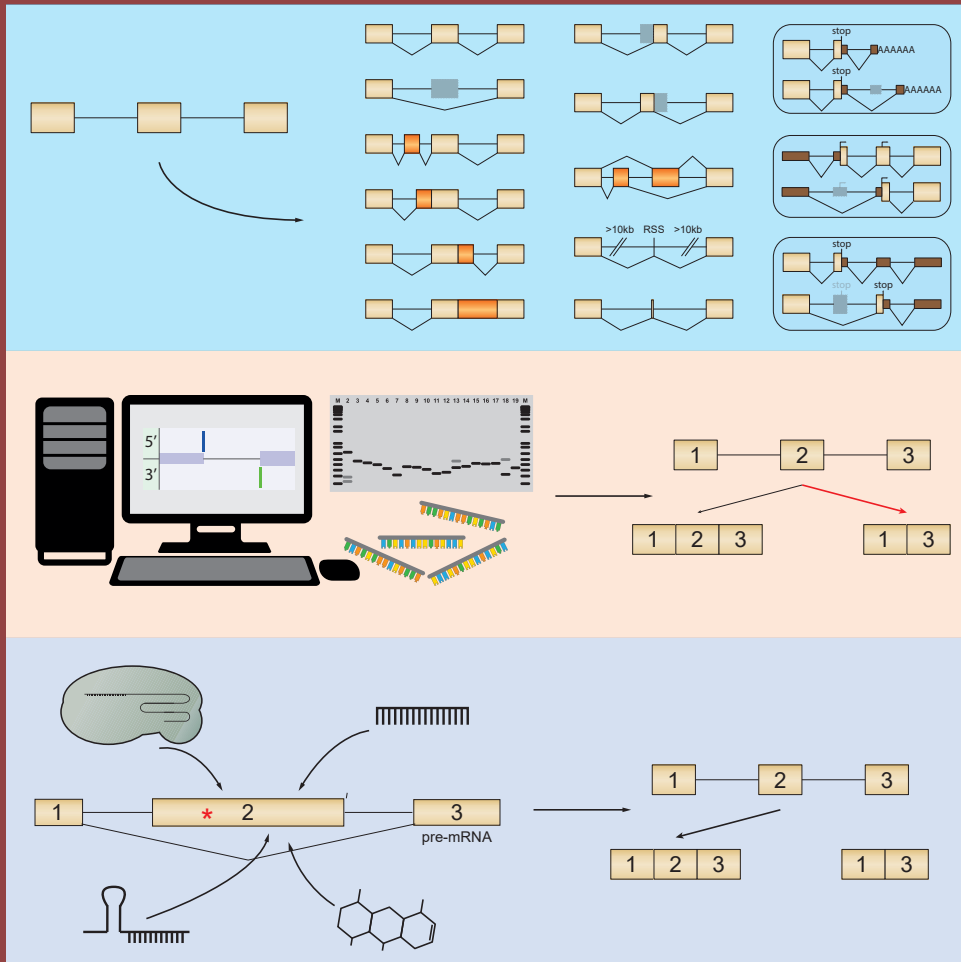


CHAPTER 2

ALTERNATIVE SPLICING IN GENETIC DISEASES: IMPROVED DIAGNOSIS AND NOVEL TREATMENT OPTIONS

Atze J. Bergsma, Erik van der Wal, **Mike Broeders**,
Ans T. van der Ploeg, W.W.M. Pim Pijnappel

Graphical abstract



Keywords: alternative splicing; genetic diagnostics; splicing therapy; human disease; next generation sequencing

Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options

**Atze J. Bergsma^{*,†,‡}, Erik van der Wal^{*,†,‡}, Mike Broeders^{*,†,‡},
Ans T. van der Ploeg^{†,‡}, W.W.M. Pim Pijnappel^{*,†,‡,1}**

^{*}Molecular Stem Cell Biology, Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands

[†]Department of Pediatrics, Erasmus Medical Center, Rotterdam, The Netherlands

[‡]Center for Lysosomal and Metabolic Diseases, Erasmus Medical Center, Rotterdam, The Netherlands

¹Corresponding author: e-mail address: w.pijnappel@erasmusmc.nl

Contents

| | |
|--|-----|
| 1. Introduction | 87 |
| 2. Different Outcomes of Alternative Splicing and Their Impact on Protein Expression | 88 |
| 3. Mechanisms of Splicing Regulation in Homeostasis and Disease | 91 |
| 3.1 Spliceosome | 91 |
| 3.2 The Spliceosome in Human Disease | 92 |
| 3.3 Splicing Enhancers and Repressors | 93 |
| 3.4 Tissue-Specific Master Splicing Factors | 96 |
| 3.5 RBM Family | 96 |
| 3.6 Rbfox Family | 97 |
| 3.7 Muscleblind Family | 99 |
| 3.8 Chromatin and Splicing Regulation | 100 |
| 3.9 Recursive Splicing | 101 |
| 4. Identification of Pathogenic Variants That Affect Splicing | 104 |
| 4.1 Molecular Diagnostics in the NGS Era | 104 |
| 4.2 In Silico Prediction of Functional Effects That Variants Have on Splicing | 105 |
| 4.3 Functional Assays | 109 |
| 4.4 Ultraviolet Crosslinking and Immunoprecipitation | 111 |
| 4.5 Are Splicing Variants Underestimated? | 113 |
| 5. Therapeutic Approaches to Modulate Splicing | 114 |
| 5.1 Small Molecule Modulators of Splicing | 114 |
| 5.2 Transsplicing | 116 |
| 5.3 Antisense Oligonucleotides | 117 |
| 5.4 Gene Therapy | 119 |
| 6. Future Perspectives | 122 |
| Acknowledgments | 123 |
| References | 124 |

Abstract

Alternative splicing is an important mechanism to regulate gene expression and to expand the repertoire of gene products in order to accommodate an increase in complexity of multicellular organisms. It needs to be precisely regulated, which is achieved via RNA structure, splicing factors, transcriptional regulation, and chromatin. Changes in any of these factors can lead to disease. These may include the core spliceosome, splicing enhancer/repressor sequences and their interacting proteins, the speed of transcription by RNA polymerase II, and histone modifications. While the basic principle of splicing is well understood, it is still very difficult to predict splicing outcome, due to the multiple levels of regulation. Current molecular diagnostics mainly uses Sanger sequencing of exons, or next-generation sequencing of gene panels or the whole exome. Functional analysis of potential splicing variants is scarce, and intronic variants are often not considered. This likely results in underestimation of the percentage of splicing variants. Understanding how sequence variants may affect splicing is not only crucial for confirmation of diagnosis and for genetic counseling, but also for the development of novel treatment options. These include small molecules, transsplicing, antisense oligonucleotides, and gene therapy. Here we review the current state of molecular mechanisms of splicing regulation and how deregulation can lead to human disease, diagnostics to detect splicing variants, and novel treatment options based on splicing correction.

LIST OF ABBREVIATIONS

| | |
|-----------------|--|
| 2'O-mePS | 2'O-methyl phosphorothioate |
| AON | antisense oligonucleotide |
| BMD | Becker muscular dystrophy |
| CADD | combined annotation-dependent depletion |
| cDNA | complementary DNA |
| CLIP | crosslinking and immunoprecipitation |
| DMD | Duchenne muscular dystrophy |
| ESC | embryonic stem cell |
| ESE | exonic splicing enhancer |
| ESS | exonic splicing silencer |
| GAA | acid α -glucosidase |
| HDR | homology directed repair |
| HGMD | human gene mutation database |
| hnRNPs | heterogeneous nuclear ribonucleoproteins |
| iCLIP | individual-nucleotide resolution UV CLIP |
| ISS | intronic splicing silencer |
| MOE | 2'O-methoxyethyl |
| mRNA | messenger RNA |
| NGD | No-go decay |
| NGS | next-generation sequencing |
| NHEJ | nonhomologous end joining |

NMD nonsense-mediated decay
NSD nonstop decay
PAR-CLIP photoactivatable-ribonucleoside enhanced CLIP
PMO phosphorodiamidate morpholino oligomer
rAAVs recombinant adeno-associated viruses
RRM RNA recognition motif
SF3B splicing factor 3B
SMA spinal muscular atrophy
SNP single-nucleotide polymorphism
snRNA small nuclear RNA
snRNP small nuclear ribonuclear protein
sQTL splicing quantitative trait locus
SR serine/arginine-rich
SREs splicing regulatory elements
SRSF SR splicing factor
SSOs splice-switching oligonucleotides
WES whole exome sequencing
WGS whole genome sequencing



1. INTRODUCTION

Determination of the pathogenicity of sequence variants is important for prediction of disease severity and genetic counseling. Furthermore, understanding the basic mechanism underlying disease caused by pathogenic variants is critical for development of new therapies. To date, approximately 192,000 disease-related lesions have been described in the human gene mutation database (HGMD). These occur in more than 30% of human protein-coding genes (<http://www.hgmd.cf.ac.uk>). Of all variants that are annotated in HGMD, 17,439 entries are believed to alter canonical splicing (HGMD, December 2016). These account for 9.1% of all disease-associated variants in the database, although this likely represents an underestimation as most efforts are directed toward missense and nonsense variants and because splicing is difficult to predict. With the advent of new next-generation sequencing (NGS) techniques (Goodwin et al., 2016), many more variants and genes that are linked to disease are likely to be added to the HGMD, broadening our insight of genotype/phenotype relations. However, there is an enormous pace at which new potentially pathogenic variants are discovered, with associated challenges to determine their pathogenicity. This review highlights advances in diagnostics, our understanding of the

different levels at which splicing is regulated in healthy individuals and in disease, and the development of potential novel treatment options based on splicing correction.



2. DIFFERENT OUTCOMES OF ALTERNATIVE SPLICING AND THEIR IMPACT ON PROTEIN EXPRESSION

Current estimates state that the human genome contains close to 21,000 protein-coding genes (Moraes and Goes, 2016). Alternative promoter usage, transcription termination, RNA processing, and posttranslational modifications further expand the repertoire for protein diversity. A striking example of diversity generated by alternative splicing is the *Dscam* gene in *Drosophila*, from which potentially 38,016 different transcripts can be expressed (Park and Graveley, 2007). Interestingly, almost all human gene transcripts undergo alternative splicing and are able to produce multiple types of transcripts (Barbosa-Morais et al., 2012; Pan et al., 2008; Wang et al., 2008). These potentially could expand the repertoire of human proteins by approximately 10-fold (Nilsen and Graveley, 2010). Indeed, the major fraction of alternatively spliced messenger RNA (mRNA) transcripts is engaged by ribosomes, suggesting these transcripts are translated into protein (Weatheritt et al., 2016). It will be important to confirm this hypothesis with proteomic approaches (Tress et al., 2017). Another important role for alternative splicing is the regulation of mRNA expression by generating out of frame products resulting in mRNA-decay, for example intron retention (Wong et al., 2016). One such case is the autoregulatory negative feedback loop that controls expression of Chtop (chromatin target of Prmt1). In this loop, Chtop binds its own mRNA and regulates intron 2 retention. Subsequently, a premature termination codon in the 5' end of intron 2 leads to nonsense-mediated decay (NMD) of the mRNA (Izumikawa et al., 2016). Furthermore, the extent of alternative splicing differs per type of tissue, with relatively high levels in brain compared to other tissues (Barbosa-Morais et al., 2012). Alternative splicing is thought to be important for normal development and homeostasis of various tissues including the nervous system (Raj and Blencowe, 2015), the immune system (Carpenter et al., 2014), and muscle (Jangi et al., 2014).

Multiple types of alternative splicing are known. Many genes constitutively express one major isoform which contains all canonical exons (Fig. 1). During alternative splicing, canonical splice sites can be skipped and/or noncanonical splice sites can be utilized to form different transcripts.

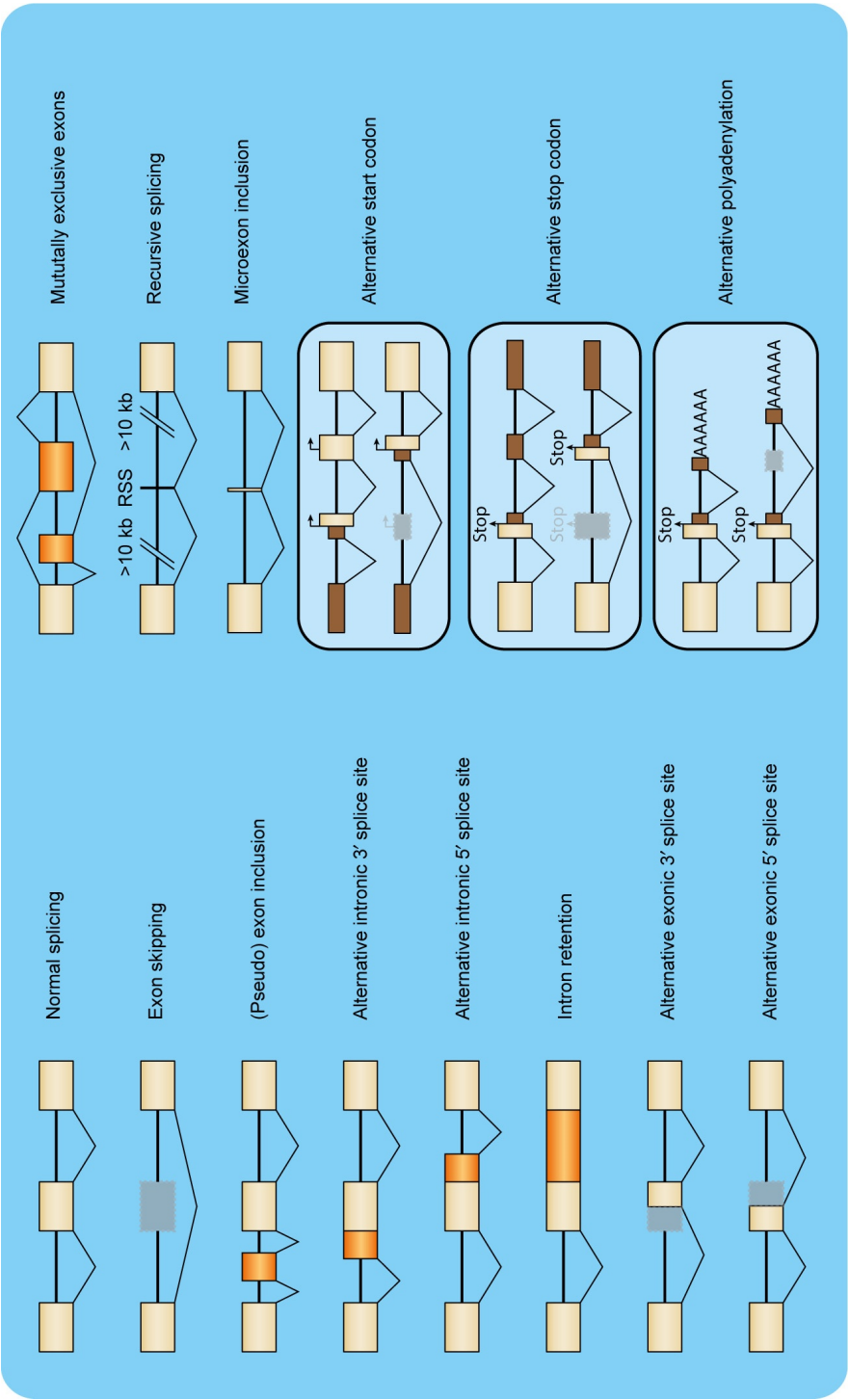


Fig. 1 Potential outcomes of alternative splicing. Constitutive exons are highlighted in beige, alternative exons and retained intronic sequences in orange, (partially) skipped exons in gray, 3' and 5' UTRs in brown, introns are depicted as black lines.

The consequences include exon skipping, exon inclusion, alternative 3' and 5' splice site usage, and intron retention (Fig. 1). Genetic variants can shift splicing patterns by altering the recognition of splice sites and/or splicing regulatory elements (SREs) by the spliceosome. This can have multiple outcomes.

(1) When the new length of the coding region in the mature mRNA cannot be divided by three, the result will be a reading frame shift and the generation of a premature stop-codon or the abolishment of a natural stop-codon. Normally, these products are degraded via the NMD pathway or the nonstop decay (NSD) pathway, respectively (reviewed by [Simms et al., 2017](#)). Aberrant proteins generated by NSD-mRNA are subsequently degraded by a ribosome-associated E3 ubiquitin ligase that is capable of detecting defective nascent proteins, targeting them for degradation. This acts as a protein quality control measure ([Bengtson and Joazeiro, 2010](#)). Activation of these decay pathways is often seen in human disease ([Miller and Pearce, 2014](#)). Interestingly, they also play important roles in normal physiology to provide an extra layer of regulation of gene expression. Examples include the determination of human stem cell fate ([Lou et al., 2016](#)), inflammation, and myeloid cell differentiation ([Saul et al., 2016](#)).

(2) Another outcome of alternative splicing can be that inclusion and/or exclusion of nucleotides leaves the reading frame intact, which generally does not result in mRNA decay via the NMD or NSD pathways. This includes the utilization of mutually exclusive exons, and it can potentially reshape the function of the encoded protein. In normal development and homeostasis, it provides an important mechanism to generate alternative proteins with tissue-specific expression patterns and distinct functions. For example, alternative splicing can alter protein localization via changes in signaling peptides ([Balthazar et al., 2017](#)), protein-protein interaction networks in neuronal tissue ([Ellis et al., 2012](#)), and organization of T-tubule structure during the maturation of muscle fibers via alternative splicing of four trafficking proteins ([Giudice et al., 2016](#)). Furthermore, inclusion of microexons, which are 3–50 bp in length, leaves the reading frame intact in 80%–90% of cases in neuronal tissue, suggesting that micro-exon inclusion is another important mechanism for protein diversification ([Irimia et al., 2014](#); [Li et al., 2015](#)) (Fig. 1). Variations in the human genome can lead to alternative splicing with pathogenic consequences. In human disease, aberrant splicing can generate protein products with loss or gain of function and/or toxicity, such as in several forms of Tauopathies (reviewed in [Park et al., 2016](#)). In-frame splice products can potentially activate another decay pathway, termed the no-go decay (NGD) pathway. This

pathway is initiated by barriers that block translation elongation in the ribosome, such as the presence of RNA structures like hairpins and pseudoknots (Doma and Parker, 2006). NGD is initiated in a similar manner compared to NSD, with endonucleolytic cleavage of stalled mRNAs close to the stall site, leaving an NSD-like substrate. This cleavage is stimulated by the DOM34:HBS1 complex both in NSD and in NGD (Tsuboi et al., 2012).

(3) Besides changes within the coding region of pre-mRNA, alternative splicing can also induce the usage of alternative translation start/stop sites and of different polyadenylation signals, which in turn causes changes in the 5' and 3' UTRs, respectively (Fig. 1). These variations can result in differential binding of RNA-binding proteins (RBPs), leading to the use of different polyadenylation signals or inclusion/exclusion of miRNA binding sites. One example is the upregulation of expression of functional *DUX4* mRNA. Upon contraction of the D4Z4 repeat at the telomeric end of chromosome 10, *DUX4* expression is induced. However, only in the context of a particular genetic background in which *DUX4* splicing leads to inclusion of a polyA signal in the mRNA, functional DUX4 protein is produced (Daxinger et al., 2015). DUX4 protein is toxic to skeletal muscle cells, and these events result in facioscapulohumeral muscular dystrophy.



3. MECHANISMS OF SPLICING REGULATION IN HOMEOSTASIS AND DISEASE

3.1 Spliceosome

Pre-mRNA splicing is regulated at several layers. The first is at the level of the core splicing machinery, which consists of a large RNA/protein complex termed the spliceosome. It exists in two forms: the major and the minor spliceosomes, of which the major spliceosome accounts for splicing of more than 99% of all introns (Turunen et al., 2013). The major spliceosome is composed of the small nuclear ribonuclear proteins (snRNPs) U1, U2, U4, U5, and U6, which assemble in a specific order (Fig. 2). The mechanism of action has been extensively reviewed by Lee and Rio (2015). The minor spliceosome has a similar mechanism as the major spliceosome. It recognizes slightly different consensus intronic sequences and utilizes the functional analogs U11, U12, U4atac, and U6atac of the U1, U2, U4, and U6 snRNPs from the major spliceosome, respectively. U5 remains shared between both complexes (Turunen et al., 2013). The spliceosome identifies target introns by recognition of conserved splice elements that are present in the native pre-mRNA sequence including the 5' and 3' splice sites, the branch point site

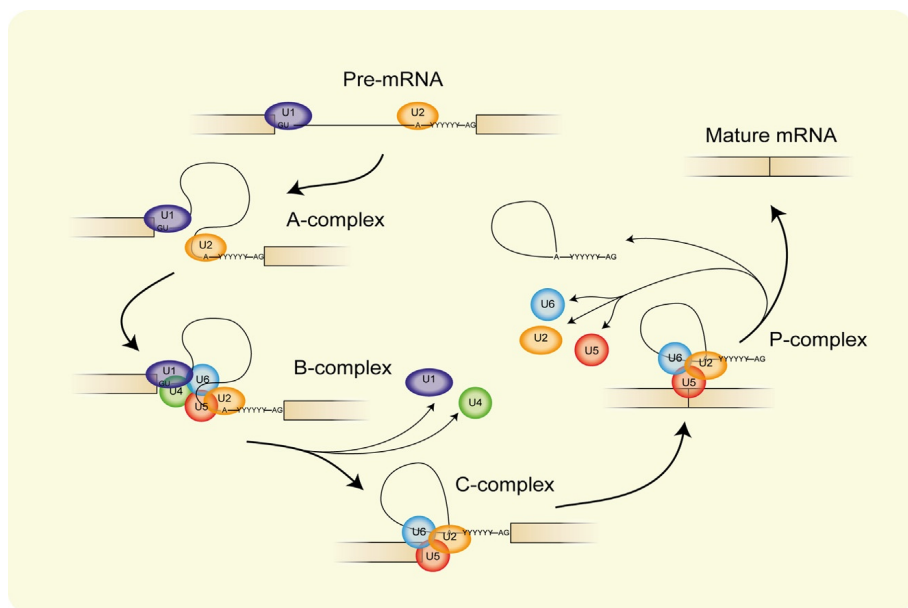


Fig. 2 Schematic representation of splicing performed by the major spliceosome. Pre-mRNA splicing starts with recognition of the 5' and 3' splice sites. Looping occurs to form the A-complex, generating a platform for the binding of the U4–U5–U6 tri-snRNP to form the B-complex. After activation of the B-complex, two transesterification steps follow to cut out the intron lariat via C- and P-complexes formation and ligation of the remaining exons, resulting in completion of the splicing process and formation of the mature mRNA.

and the polypyrimidine tract (Fig. 2). U1 recognizes the 5' splice site, while U2 recognizes the branch point and the 3' splice site to form the A-complex. In a next step, U4, U5, and U6 snRNPs join to form the B-complex. Two transesterification reactions follow in which the intron lariat is cut and exons are ligated during processing through the C- and P-complexes (reviewed in Moore and Sharp, 1993). It is well known that variations in any of the conserved splicing elements can result in aberrant splicing patterns. In fact, we calculated that 63% of splicing variants annotated in the HGMD are located within the two bases at the beginning and the end of introns (<http://www.hgmd.cf.ac.uk/ac/hoho2.php>). Pathogenic splicing variants that cause monogenic disorders have been identified in hundreds of disease-causing genes. These variants are often located at well-defined splice sites.

3.2 The Spliceosome in Human Disease

Pathogenic variants in genes encoding for subunits of the major spliceosome occur in various cancers and in retinitis pigmentosa (see review in Singh and

Cooper, 2012). Variants causing retinitis pigmentosa were identified in U4/U5/U6 tri-snRNP-specific proteins, but these only affected splicing of retina-related genes (Ruzickova and Stanek, 2016). The question is why only the retina is affected by defects in the major spliceosome. It has been hypothesized that splicing activity is only partially decreased in retinitis pigmentosa, and that only weakly spliced introns present in retina-specific genes are affected (Linder et al., 2011).

Additional evidence for a role of the major spliceosome in human development and disease was reported recently for germ cell development and variants that may cause male infertility (Wu et al., 2016). A genome-wide association study, in which single-nucleotide polymorphisms (SNPs) are compared, identified significant differences in the genomes of 981 patients with nonobstructive azoospermia (NOA) disorder, compared to the genomes of 1,657 healthy controls. This demonstrated linkage with the locus containing the *SNRPA1* gene, which encodes for the U2A subunit of the major spliceosome. Missense variants in the human *SNRPA1* expressed in *Drosophila* resulted in deregulation of the major spliceosome and disrupted spermatogenesis in *Drosophila*. In addition, knockdown of other core subunits of the major spliceosome (U1, U2, U4, and U5) in *Drosophila* resulted in male infertility. These results suggest that the major spliceosome is involved in spermatogenesis, and that its dysfunction could explain infertility in patients with NOA. Variants in U4atac, U11/U12–65k, and U12 of the minor spliceosome cause Taybi–Linder syndrome, isolated growth hormone deficiency, and early onset cerebellar ataxia, respectively. This is caused by intron retention of transcripts that are processed by the minor spliceosome (Argente et al., 2014; Ederly et al., 2011; Elsaid et al., 2017).

The examples above show that variants in components of the core spliceosomal machinery can cause various human disorders that affect very distinct cell types (Fig. 3). While this seems counterintuitive, it may indicate that splicing outcome is a trade-off between many regulatory pathways, and that the requirements for splicing regulatory factors to ensure correct splicing can be highly cell type specific.

3.3 Splicing Enhancers and Repressors

Besides the spliceosome, additional layers for the regulation of splicing exist. SREs are encoded in pre-mRNA transcripts. These are called exonic splicing enhancers (ESEs), intronic splicing enhancers, exonic splicing silencers (ESSs) or intronic splicing silencers (ISSs), and can be located at any position along the pre-mRNA (Fig. 3). Over 300 splicing regulatory proteins exist

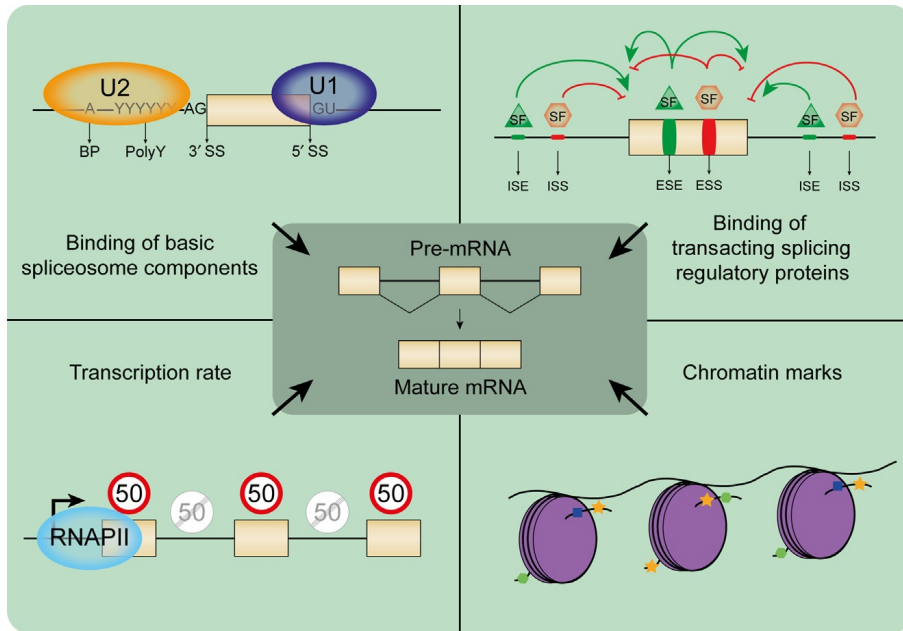


Fig. 3 Regulation of splicing. Splicing regulatory elements are utilized to guide binding of constitutive or tissue-specific splicing factors to modulate splicing. Transcription rate and changes in chromatin marks can also have a profound effect on splicing outcome.

that are thought to act via these elements ([Hoskins and Moore, 2012](#)). Following their binding to splice enhancer or silencer sequences, these proteins confer a positive or negative signal, influencing recognition of nearby splice sites by the spliceosome, thereby influencing splicing outcome ([Fig. 3](#)). These signals are important for the recognition of splice sites and have been shown to be able to function over distances >300 bp ([Schultz et al., 2016](#)). There are several different types of splicing regulatory protein groups, including heterogeneous nuclear ribonucleoproteins (hnRNPs), serine/ argenine-rich (SR) proteins, and tissue-specific RBPs. Each of these families can have both positive and negative effects on splicing.

The hnRNP family of splicing regulatory proteins is a group consisting of at least 20 proteins that exert different functions, dictated by their domain organization and cellular localization (reviewed in [Geuens et al., 2016](#)). hnRNPs are involved in multiple RNA-related processes during the regulation of transcription and translation ([Park et al., 2015](#)), such as RNA trafficking and localization ([Gao et al., 2008](#); [Shan et al., 2003](#)), loading of mRNA into exosomes ([Villarroya-Beltri et al., 2013](#)), telomere maintenance ([Pont et al., 2012](#)), and splicing. hnRNPs can influence exon

definition either positively or negatively to modulate splicing. For example, hnRNP A1 primarily binds to splicing silencers from which it negatively influences exon definition in various ways (Blanchette and Chabot, 1999; Krecic and Swanson, 1999; Okunola and Krainer, 2009). hnRNP A1 has also been implicated in stimulation of exon inclusion (Martinez-Contreras et al., 2006; Oh et al., 2013). Similarly, hnRNP L can bind to multiple SREs to influence exon definition either positively or negatively (Motta-Mena et al., 2010). Because of the heterogeneous functions of members of the hnRNP family, it remains difficult to predict splicing-outcome when variants disrupt binding of these proteins. However, it is thought that hnRNPs tend to have more negative than positive effects on exon definition (Geuens et al., 2016).

hnRNPs have also been implicated in disease. An example is hnRNP C. It is one of the first family members that have been shown to be essential for splicing (Choi et al., 1986). hnRNP C plays a role in packaging of native RNA transcripts (Choi et al., 1986; Dreyfuss et al., 1993) and suppression of inclusion of alternative exons (Konig et al., 2010), especially of cryptic exons generated by transposition of Alu-elements (Zarnack et al., 2013). Dysregulation of hnRNP C can lead to altered expression of cancer-related genes such as *BRCA* (Anantha et al., 2013), indicating its potential role in certain malignancies. Furthermore, hnRNPs have been shown to play a major role in the regulation of splicing in neuronal cell homeostasis and have been implicated in diseases including ALS/FTLD, fragile X syndrome, Alzheimer's disease, and spinal muscular atrophy (SMA) (reviewed in Geuens et al., 2016).

SR proteins contain an N-terminal RNA recognition motif (RRM) and a C-terminal SR domain (Long and Caceres, 2009). Like hnRNPs, SR proteins can influence splice site definition in both positive and negative ways (Pandit et al., 2013; Zhou and Fu, 2013). Among the first identified SR proteins are SR splicing factor 1 (SRSF1) (also known as SF2/ASF) and SRSF2 (also known as SC35) (Fu et al., 1992; Krainer et al., 1990; Manley and Krainer, 2010). So far, 12 SRSF factors have been described, all of which function in splicing regulation (reviewed in Howard and Sanford, 2015).

Pathogenic variants that disrupt SREs can result in altered splicing patterns and disease. An example can be found in patients with SMA. SMA is caused by pathogenic variants in the *SMN1* gene. Its sister gene *SMN2* is inactivated by a point variant in exon 7, which disrupts an ESE that is normally recognized by SRSF1 (Cartegni et al., 2006), and/or generates an ESS that is bound by hnRNP A1 and A2, resulting in exon skipping (Kashima

et al., 2007). Antisense therapies are now being evaluated that block an ISS located in *SMN2* intron 7 to promote canonical splicing (Singh et al., 2006; see Section 5.3). Another example is pyruvate dehydrogenase complex deficiency, which causes Leigh's encephalomyelopathies and lactic acidosis, and is predominantly caused by pathogenic variants in the *E1 α* gene. One variant has been identified that causes missplicing of *E1 α* exon 7 due to an increase in the strength of an SRSF2 binding site in intron 7, which results in activation of a cryptic splice donor site (Gabut et al., 2005).

3.4 Tissue-Specific Master Splicing Factors

Tissue-specific splicing proteins play important roles in cellular differentiation, cell-fate determination, and regulation of cellular homeostasis. A number of master splicing factors have been identified that regulate downstream splicing networks in a tissue-specific manner (Jangi and Sharp, 2014). Dysregulation or pathogenic variants in these factors can lead to several diseases. Examples are found within the RNA binding motif (RBM) family, the Rbfox family, and the muscleblind family.

3.5 RBM Family

RBM24 and its close paralog RBM38 have been found to regulate alternative splicing and to play an important role in myogenesis. Jin et al. showed that RBM24 is involved in the posttranscriptional regulation of myogenin and p21. RBM24 regulates the mRNA levels of myogenin and p21, in part, through a myogenin-dependent posttranscriptional regulatory pathway and promotes myogenic differentiation (Jiang et al., 2014; Jin et al., 2010). Knockdown of RBM24 resulted in a decreased half-life of myogenin, suppression of differentiation-associated cell cycle arrest, and a delayed myogenic differentiation (Jin et al., 2010; Miyamoto et al., 2009). Knockdown of RBM38 had similar effects on cell cycle arrest and myogenic differentiation (Miyamoto et al., 2009). Overexpression of either RBM24 or RBM38 had the opposite effect. It induced cell cycle arrest, and promoted myogenic differentiation (Miyamoto et al., 2009). Yang et al. identified RBM24 as a crucial muscle-specific splicing factor for skeletal muscle and cardiac development. In vitro splicing assays showed that muscle-specific exon inclusion in nuclear extracts of nonmuscle cells was promoted by recombinant RBM24, and that RBM24 is crucial for at least 68 splicing events, mainly exon inclusions (Yang et al., 2014).

RBM24 and RBM38 have been suggested to act as tumor suppressors. RBM38 is repressed in acute myeloid leukemia (AML) patients, and low levels of *RBM38* mRNA and downstream effector proteins may contribute to the differentiation block in AML (Wampfler et al., 2016). RBM24 is thought to act as a tumor suppressor through the upregulation of miR-25 in nasopharyngeal carcinoma (Hua et al., 2016). RBM24 has also been linked to Hirschsprung disease where it binds to *MIR143HG* and accelerates its transcript degradation (Du et al., 2016).

RBM20 directly regulates splicing of Titin through binding to certain regions of the pre-mRNA where it represses splicing (Li et al., 2013). In the remaining regions, where there is no binding of RBM20, Titin pre-mRNA is normally spliced, which generates a partially processed Titin pre-mRNA. Instead of being exported from the nucleus, these partially processed Titin pre-mRNAs are stored in the nucleus awaiting further processing. The 5' splice site and 3' splice site of the exons flanking the RBM20-repressed regions are then utilized to remove the repressed internal exons and introns (Li et al., 2013).

RBM20 has also been related to human heart disease. Pathogenic variants in *RBM20* lead to several cardiomyopathies like sudden cardiac death and dilated cardiomyopathy (Brauch et al., 2009). A deletion of the *RBM20*-RRM in a mouse model and knockout of *RBM20* in rat showed more expression of the cardiomyopathy-related compliant Titin isoforms and recapitulated the human *RBM20*-associated cardiomyopathy (Guo et al., 2012; Methawasin et al., 2014). Hinze et al. used a Titin N2B knockout mouse (*Titin*^{ΔN2B/ΔN2B}), which shows a complex cardiac phenotype including cardiac atrophy and dysfunction. Altering differential splicing of *RBM20* substrates showed positive effects on cardiac function in this model, which suggested *RBM20* as a therapeutic target (Hinze et al., 2016). In double-deficient mice (*Titin*^{ΔN2B/ΔN2B} *RBM20*^{ΔRRM/WT}), splicing of the compliant Titin isoforms was reduced and this was accompanied by positive effects on mRNA levels of genes related to the cAMP response, cardiac dimensions, and oxidative phosphorylation. In this splice-rescue mouse model, a reduction of 50% of *RBM20* function was sufficient to restore the diastolic dysfunction completely and reverted the cardiac atrophy, while no obvious extracardiac phenotypes with respect to weight, fertility, and grooming behavior were observed (Hinze et al., 2016).

3.6 Rbfox Family

The Rbfox family regulates alternative splicing in multiple cell types, including embryonic stem cells (ESCs), and cells from neuronal and myogenic

lineages (Gehman et al., 2012; Jin et al., 2003; Zhang et al., 2008). The Rbfox1 and Rbfox3 family members are expressed in muscle and neuronal tissues, whereas Rbfox2 regulates expression of more than 70 RBPs in mouse ESCs, thereby inducing alternative splicing-mediated NMD. This results in differentiation toward the mesodermal lineage (Jangi et al., 2014).

Rbfox1 has been linked to several disorders, one of which is autism spectrum disorder (ASD) (Martin et al., 2007; Sebat et al., 2007). In patients with ASD, downregulation of Rbfox1 in the brain was associated with altered splicing patterns of 48 ASD-susceptibility genes (Weyn-Vanhentenryck et al., 2014), and correlated with altered splicing of its predicted target exons (Bhalla et al., 2004; Voineagu et al., 2011). Hamada et al. showed that a specific isoform of Rbfox, Rbfox-isoform1 (iso1), plays an important role in synapse network formation and neuronal migration and that this may relate to ASD. Rbfox-iso1 knockdown in vitro resulted in a reduction of mature spine number and spine density, while in vivo knockdown showed defects in the radial migration and terminal translocation of cortical neurons. Involvement of Rbfox1-iso1 in the formation or maintenance of the synapse network was proposed (Hamada et al., 2016).

Gao et al. have shown that Rbfox1 is also an important regulator of alternative RNA splicing during heart failure, and proposed Rbfox1-mediated pre-mRNA splicing as a new potential therapeutic target (Gao et al., 2016). In a mouse model, Rbfox1 deficiency promoted pressure overload-induced heart failure and cardiac hypertrophy, whereas cardiac reexpression of Rbfox1 rescued this pathology. The same study showed that Rbfox1 expression is decreased in human failing hearts, and that Rbfox1 expression is essential for normal heart development in the developing zebrafish (Gao et al., 2016). The proposed mechanism involves the MEF2 transcription factor family. Rbfox regulates the switch from MEF2 isoform $\alpha 1$ to $\alpha 2$, which have differential effects on expression of cardiac hypertrophy genes. Similar Rbfox1 binding motifs are found near the targeted *MEF2* $\alpha 2$ exons across zebrafish, mice, rats, and humans, suggesting a conserved Rbfox1/*MEF2* mechanism.

In a similar fashion, dysregulation of Rbfox2 has been linked to other cardiac pathologies like hypoplastic left heart syndrome (HLHS) and cardiac pathology in diabetes. A study by Verma et al. showed that Rbfox2 is a major contributor to transcriptome changes in HLHS patients. In these patients, Rbfox2 carries pathogenic variants that alter its function in pre-mRNA splicing. In particular, Rbfox2 binding sites were found in 936 differentially

expressed transcripts in the right ventricle of infant HLHS patients (Verma et al., 2016). Nutter et al. showed that in diabetic hearts, 73% of the mis-spliced transcripts have an Rbfox2 binding site. A dominant-negative mechanism has been proposed for Rbfox2, in which an Rbfox2 isoform binds to wild-type Rbfox2 and subsequently inhibits Rbfox2-mediated splicing (Nutter et al., 2016). Rbfox3 has been less well studied in relation to disease, although a recent study linked variants in *RBFOX3* with sleep latency (Amin et al., 2016).

3.7 Muscleblind Family

Another well-studied splicing network is controlled by the muscleblind family. Members of this family play a role in cell differentiation, inferred from an improved efficiency of cellular reprogramming to induced pluripotent stem cells after knockdown of muscleblind-like splicing regulator 1 (MBNL1) and the MBNL-associated changes in splicing patterns upon mesodermal differentiation (Han et al., 2013; Venables et al., 2013). A variety of alternative splicing events involving cassette exons are directly regulated by MBNL1 and MBNL2 and these differ between ESCs and other cell types. In ESCs, overexpression of MBNL proteins shifted the splicing pattern toward a differentiation-like splicing pattern, whereas MBNL knockdown in differentiated cells promoted an ESC-like splicing pattern (Han et al., 2013). Another group showed that during cardiomyocyte maturation MBNL1-dependent splicing changes were enriched for spliceosomal genes as well as genes relevant to vesicular trafficking and membrane organization (Giudice et al., 2014). The function of MBNL and its effect on splicing regulation in myotonic dystrophy (DM) has been extensively reviewed in Cardani et al. (2014). A recent study showed that splicing events are responsive to the dose of MBNL1 protein. Splicing events were evaluated upon a range of MBNL1 protein levels, and each splicing event presented a unique dose-response curve. It has been suggested that the alternative splicing events can be used as a biomarker to accurately estimate the functional levels of MBNL1 protein in DM (Wagner et al., 2016).

The above tissue-specific regulators of splicing have many downstream targets and can be considered important master regulators. It will be interesting to further define the hierarchy of tissue-specific splicing regulation. This will be important to understand the consequences of misregulation of these factors for human development and disease.

3.8 Chromatin and Splicing Regulation

The cotranscriptional nature of pre-mRNA imposes transcription and chromatin structure as potential levels at which splicing may be regulated (Fig. 3). Chromatin structure differentiates intronic from exonic sequences (Schwartz et al., 2009b). One of the factors influencing this is GC content. In general, GC content tends to be higher in exonic regions. GC-rich regions are more likely to result in DNA methylation, as this occurs at CG dinucleotides (Gelfman and Ast, 2013). When DNA is methylated, it is more densely packed in chromatin, thereby slowing down passage of RNA-polymerase II (RNAPII) during transcription. This leads to a lower transcription rate, resulting in increased accumulation of RNAPII over intron–exon boundaries and exon bodies, and ultimately increases the chance of exon inclusion during cotranscriptional splicing as it allows more time for splicing (Adelman and Lis, 2012; Jonkers et al., 2014; Kwak and Lis, 2013). Furthermore, levels of specific histone modifications are different in exonic regions compared to intronic regions. One example is histone 3 lysine 36 trimethylation (H3K36me3), which is associated with transcriptional elongation and tends to coaccumulate with RNAPII at canonical splice sites and exonic regions (Ye et al., 2014). H3K36me3 also seems to be more enriched at constitutive exons compared to alternative exons (Kolasinska-Zwierz et al., 2009). The histone deacetylase complex RPD3S is recruited to the H3K36 di- and trimethylation marks and removes acetylation marks from nearby nucleosomes. Acetylation of histones generally leads to a more open chromatin structure, and removal of these marks leads to a more closed structure, resulting in a lowered transcription rate at these sites and consequently more time for correct splicing. Variants in the *SETD2* gene, which encodes for the enzyme responsible for H3K36 trimethylation, have been linked to human kidney cancer via alterations in RNA processing of highly transcribed genes including intron retention and aberrant splicing (Simon et al., 2014). Furthermore, alteration of H3K36me3 levels result in increased risk of renal cell carcinoma-specific death via changes in RNA processing (Ho et al., 2016).

Other histone marks that have been associated with splicing include H3K4me3, H4K20me1, and H3K79 di- and trimethylation (Andersson et al., 2009; Luco et al., 2010; Shindo et al., 2013; Ye et al., 2014). All these marks have established roles during transcriptional regulation, and their effects on splicing may either be explained by indirect effects on the speed of transcription or by providing binding platforms for splicing proteins that

modulate splicing cotranscriptionally. H3K4me3, which is associated with transcriptional activation, has been linked both with inclusion and exclusion of exons in several studies (Luco et al., 2010; Shindo et al., 2013). Enrichment in H3K79me1, which is involved in various roles of chromatin remodeling, correlates with exon inclusion in human embryonic and fetal lung fibroblast cell lines in a similar way as H3K36me3 (Shindo et al., 2013). H4K20me1 is enriched in exons and has been coupled to transcription elongation (Schwartz et al., 2009b). Global cassette exon inclusion levels positively correlate with global H4K20me1 levels of nucleosomes at the position of cassette exons in B-lymphocytes and chronic myelogenous leukemia cell lines (Liu et al., 2014).

Recruitment of splicing regulatory proteins can occur via chromatin-binding factors, as is the case with MRG15, which is part of several histone-modifying complexes. MRG15 is recruited to the H3K36me3 histone mark. By binding to the PTB splicing regulatory factor, MRG15 can form an active cotranscriptional complex that acts as a scaffold for splicing machinery components (Llorian et al., 2010; Luco et al., 2010). Another interaction between chromatin-binding proteins and splicing regulatory proteins can be found in the interaction between PSIP1, a H3K36me3 reader protein, with SRSF1 (Pradeepa et al., 2012). Furthermore CHD1, and ATP-dependent chromatin remodeling factor, is known to interact with the U2 snRNP (Sims et al., 2007).

Chromatin-binding proteins can act as scaffolds to recruit elements of the splicing machinery toward actively transcribed genes, specifically toward intronic regions. These and other findings strongly support the “exon definition” model, which states that exonic marks are recognized by the splicing machinery. This model is indeed favored in vertebrate species, where intronic lengths are considerably longer than exonic lengths (Hollander et al., 2016). Variants that lead to changes in the distribution of histone marks that are associated with exon definition could potentially lead to altered splicing.

3.9 Recursive Splicing

In very large introns, the three reactive sites in splicing, i.e., the 5' splice site, the branch point, and the 3' splice site, are separated by long stretches of RNA sequence (Fig. 4A and B). How these sites come together in a three-dimensional space is hard to explain with a conventional splicing mechanism. In humans, over 3400 introns are longer than 50 kb, and over

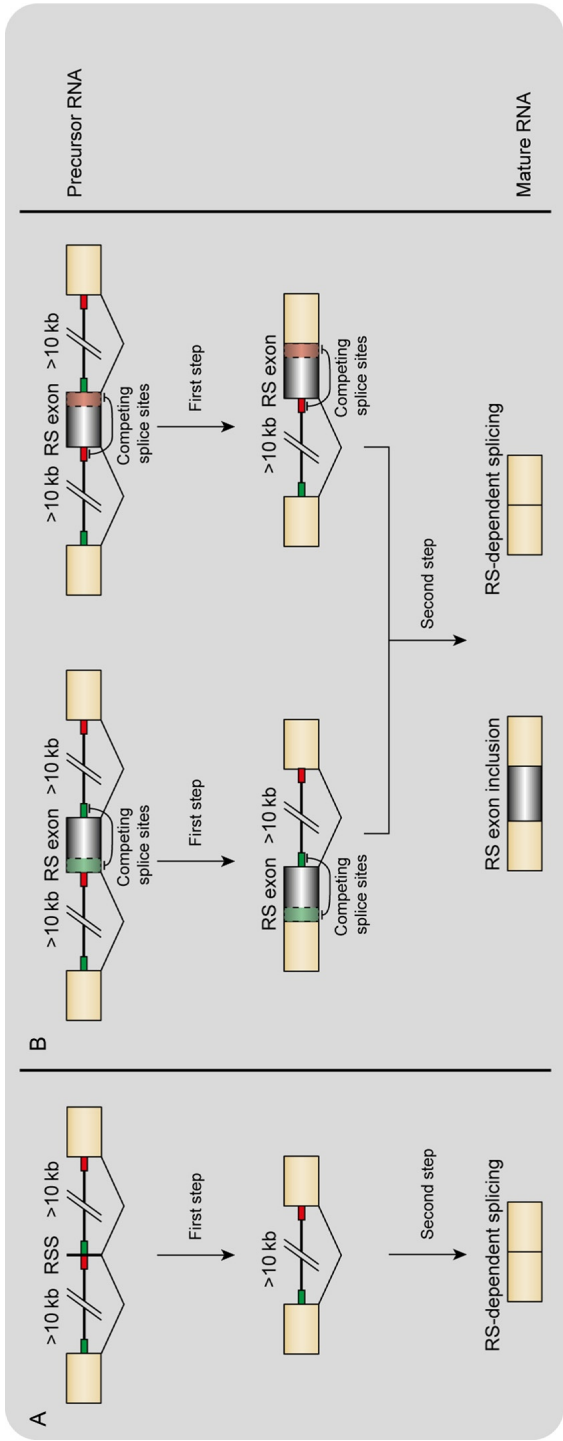


Fig. 4 Schematic representation of recursive splicing (RS). (A) Recursive splice sites deep in the intron are used to remove long intron sequences in a stepwise process. (B) RS utilizing a RS exon. Precursor RNAs are spliced using noncanonical splice sites to form a RS exon. In the first step, one part of the intron is spliced out creating the RS exon. This RS exon can be retained or removed in the second step, depending on which of two competing splice sites is used. 5'-Splice motif depicted in green, 3'-Splice motif in red.

1200 introns are longer than 100 kb (Bradnam and Korf, 2008). A refined mechanism for splicing of large introns was proposed and termed recursive splicing (RS). RS is the removal of large introns using noncanonical splice sites deep in the intron to splice out the intron in a stepwise manner (Fig. 4A and B). This phenomenon was already described in 1998 when Hatton et al. reported the stepwise removal of an intron of the Ultrabithorax gene in *Drosophila melanogaster* (Hatton et al., 1998). Only recently has it become clear that RS is widespread in *Drosophila* and has been conserved between species (Duff et al., 2015; Kelly et al., 2015; Sibley et al., 2015).

A recent study showed a sequential and principally obligatory process for genes that have recursive splice sites in *Drosophila* (Duff et al., 2015). Using deep-sequencing methods, 197 functional recursive splice sites were identified that were highly conserved between *Drosophila* strains. One-hundred and fifteen *Drosophila* genes were reported to undergo RS. Sibley et al. reported that nine genes in the human brain utilized RS to remove an intron (Sibley et al., 2015). A study using human primary endothelial cells detected intermediate splicing products from eight recursive splice sites in a 134 kb intron of the *SAMD4A* gene that appeared and disappeared during transcript production (Kelly et al., 2015). In contrast to RS seen in *Drosophila*, in which the recursive introns are completely spliced out (Fig. 4A), humans harbor a “recursive splicing exon” that seems crucial for removing the long intron and can be retained (Fig. 4B). Two functions of the RS exon in human have been proposed. First, to serve as quality control. RS exons are completely spliced out in normal mRNA, but can be retained when the exon is generated by a faulty splicing event or from an aberrant promotor sequence. A retained RS exon in an mRNA transcript usually leads to NMD due to premature stop codons. Second, the recognition of recursive splice sites is improved through the process of exon definition, a mechanism that defines splice sites on either side of the exon by recruiting splice promoting proteins (Sibley et al., 2015).

To the best of our knowledge, no disease has been linked to RS yet, however, mutating three recursive splice sites in *SAMD4A* using CRISPR/Cas9 resulted in a 35%–50% reduction of mRNA levels (Kelly et al., 2015). Furthermore, an antisense oligonucleotide (AON) that blocked the recursive splice site in the zebrafish *cadm2a* gene resulted in a twofold reduction of mRNA levels in vivo (Sibley et al., 2015). These results highlight the importance of RS for gene expression, and they open the possibility that human disorders may exist that are caused by disruption of this process, for example by deep intronic variants that are usually missed in standard diagnostics.

Interestingly, a recent study by Gazzoli et al. reported that RS occurs in many introns of the dystrophin gene, and that current therapies for Duchenne muscular dystrophy (DMD) based on skipping of exons 45–55 might be affected by RS (Gazzoli et al., 2016).



4. IDENTIFICATION OF PATHOGENIC VARIANTS THAT AFFECT SPLICING

4.1 Molecular Diagnostics in the NGS Era

Massively parallel sequencing, or NGS, has the major benefit that it can query multiple genomic regions in the same sample. Regions of interest can be specifically enriched using different techniques including solid and liquid hybridization capture techniques and multiplexed PCR-based methods. This technique can be beneficial for single gene screening of variants, particularly in large genes like dystrophin (Wang et al., 2014), and the breast cancer genes *BRCA1* and *BRCA2* (Neveling et al., 2017), for which robust screening methods have been developed.

Enrichment and subsequent sequencing of genes that are known to be linked to a specific phenotype is known as gene panel sequencing. This technique is currently the most common NGS technique used in a diagnostic setting. Examples of gene panels include those for inherited retinal disease (Carrigan et al., 2016), early onset epileptic encephalopathy (Gokben et al., 2016), specific types of cancer (Stanislaw et al., 2016), eye disorders (Consugar et al., 2015), hearing disorders (Tekin et al., 2016), neurodegenerative disorders (Kruger et al., 2016), cardiac disorders (Proost et al., 2017), and many others. Gene panel sequencing is a popular tool in clinical genetic laboratories because costs for sequencing and data analysis are significantly lower than for whole genome sequencing (WGS) techniques.

Whole exome sequencing (WES), in which the protein-coding part of the genome is analyzed by high-throughput sequencing, is becoming the next standard in DNA sequencing of patients that lack a diagnosis. WES enrichment is performed in the same way as gene panel sequencing, in fact many gene panels are generated by filtering WES data. In WES, several enrichment strategies can be used to enrich for all coding regions of the entire genome, including in-solution hybridization, selective circularization PCR, and high-multiplex PCR techniques (reviewed in Ballester et al., 2016). In this way, the analysis is performed in a less biased manner and novel potentially disease-causing genes may be identified. Many genes have been identified by WES that have a previously unrecognized link to disease

(Beaulieu et al., 2014; Chong et al., 2015; Wright et al., 2015). The wealth of variants detected with WES that could potentially be pathogenic necessitates extensive filtering. These filtering steps can include discarding variants that are above a certain threshold for minor allele frequency, as well as synonymous variants and intronic variants. However, this filtering increases the risk of discarding pathogenic variants for which the mechanism of pathogenicity is difficult to predict, for example, variants that affect splicing.

All NGS techniques mentioned earlier have one major limitation, which is that targeted enrichment is performed before sequence analysis. This generates an inherent bias of these approaches, and likely leaves causative variants undetected. Indeed, a recent survey has shown that NGS-based screening results in a genetic diagnosis in only 23%–26% of cases (Sawyer et al., 2016; Yang et al., 2013).

The optimal NGS technique for implementation in DNA diagnostics would be WGS, provided that costs and analysis time can be reduced in future improvements. This technique analyzes the complete genome, removing the bias of targeted enrichment seen in WES or gene panel sequencing. It should be noted however that technical bias can still occur, for instance GC-rich regions are difficult to sequence and repeat-containing sequences are excluded. WGS datasets take a relatively long time and large computer power to analyze because of their size. Furthermore, due to inclusion of all regions of the DNA, including noncoding parts, many more variants are detected, and the task to identify the truly pathogenic variant is even more daunting. Often datasets are extensively filtered for variants that are likely nonsignificant using algorithms that predict variant pathogenicity. The implementation of WGS in diagnostics has led to discovery of new disease genes. An example is the study from Nishiguchi et al., which resulted in the identification of causative variants in patients with autosomal recessive retinitis pigmentosa, including variants in the *NEK2* gene. These variants have not been linked previously to this disease (Nishiguchi et al., 2013).

4.2 In Silico Prediction of Functional Effects That Variants Have on Splicing

Various algorithms have been developed for splicing prediction. Currently, splicing prediction is routinely performed by diagnostics laboratories (Fig. 5). Many different online tools have been made available for prediction of splicing-related elements and/or to anticipate the effects that variants have on splicing (Table 1). Below, some of these are described in more detail.

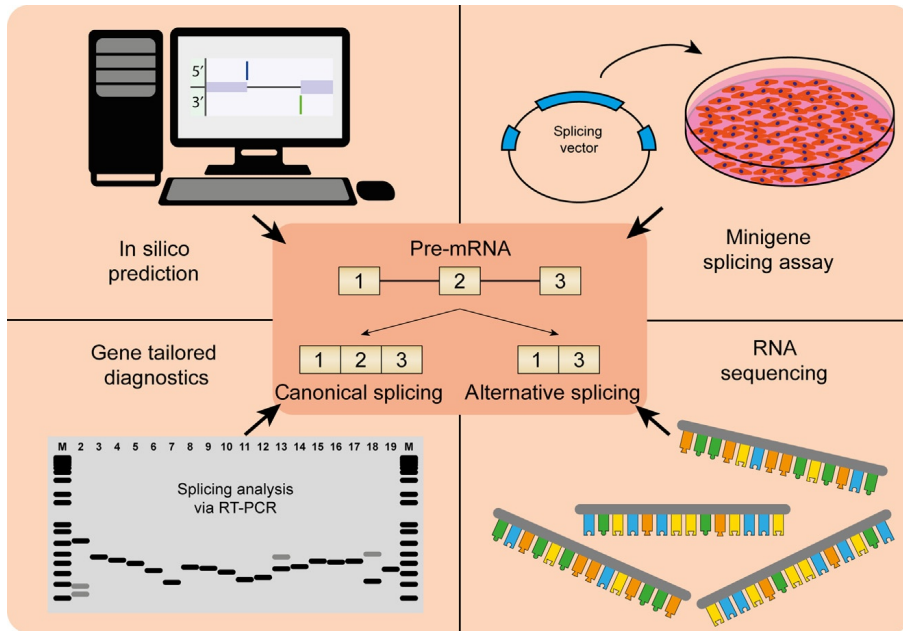


Fig. 5 Assays for testing alternative splicing.

A plethora of algorithms can be used to predict the presence of 3' and 5' splice sites (reviewed in [Jian et al., 2014](#)) as well as splicing enhancer and splicing silencer elements ([Sironi et al., 2004](#); [Soukarieh et al., 2016](#); [Yeo et al., 2004](#)). Human splicing finder (HSF) and SROOGLE are two widely used prediction tools that integrate many of different algorithms in one interface ([Desmet et al., 2009](#); [Schwartz et al., 2009a](#)). However, interpretation of the output in these programs can be challenging ([Van der Wal et al., 2017a](#)). Recently, the Automated Splice Site and Exon Definition Analyses (ASSEDA) tool has become available online. It predicts splice site utilization based on predicted splice site strengths ([Mucaki et al., 2013](#)).

Although these algorithms can be used to predict splicing outcome, caution should be taken when interpreting results from such predictions, as it does not always reflect the true molecular outcome. This was exemplified by Schleit et al. who functionally investigated splicing outcome of *COL1A1* mRNA in 40 individuals suspected of osteogenesis imperfecta ([Schleit et al., 2015](#)). Subsequently, variants present within the *COL1A1* gene of these individuals were analyzed with three different splicing prediction algorithms. The outcome of these predictions was correct in less than 75% of instances, compared to functional tests.

Table 1 Splice Prediction Tools Accessible Online

| Prediction Tool | Type of Motif Detected | Original Reference | Site |
|---------------------------|---|--|---|
| Splice site analyzer tool | Splice sites, branch site, SRE elements | Based on (Shapiro and Senapathy, 1987) | http://ibis.tau.ac.il/ssat/SpliceSiteFrame.htm |
| NetGene2 | Splice sites | Hebsgaard et al. (1996) | http://www.cbs.dtu.dk/services/NetGene2/ |
| NNSplice | Splice sites | Reese et al. (1997) | https://omictools.com/nnsplice-tool |
| GENSCAN | Splice sites | Burge and Karlin (1997) | http://genes.mit.edu/GENSCAN.html |
| SpliceView | Splice sites | Rogozin and Milanesi (1997) | http://bioinfo.itb.cnr.it/cgi-bin/oriel/wwwspliceview.pl |
| MaxEntScan | Splice sites | Yeo and Burge (2004) | http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html |
| SplicePredictor | Splice sites | Brendel and Kleffe (1998) | http://sites.usd.edu/bioinformatics/projects/splicepredictor |
| Spliceport | Splice sites | Dogan et al. (2007) | http://spliceport.cbcb.umd.edu/ |
| Human splicing finder | Splice sites, branch site, SRE elements | Desmet et al. (2009) | http://www.umd.be/HSF/ |
| CRYP-SKIP | Splice sites | Divina et al. (2009) | http://cryp-skip.img.cas.cz/ |
| EX-SKIP | Exonic splicing regulatory elements | Raponi et al. (2011) | http://ex-skip.img.cas.cz/ |
| SROOGLE | Splice sites, branch site, SRE elements | Schwartz et al. (2009a) | http://sroogle.tau.ac.il/ |
| GeneSplicer | Splice sites | Pertea et al. (2001) | https://ccb.jhu.edu/software/genesplicer/ |

Continued

Table 1 Splice Prediction Tools Accessible Online—cont'd

| Prediction Tool | Type of Motif Detected | Original Reference | Site |
|------------------|--|----------------------------|---|
| AVISPA | Splice sites | Barash et al. (2013) | http://avispa.biociphers.org |
| ESEFinder | Exonic splicing enhancers | Cartegni et al. (2003) | http://krainer01.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home |
| [RESCUE]-ESE | Exonic splicing regulatory elements | Fairbrother et al. (2002) | http://genes.mit.edu/burgelab/rescue-ese/ |
| ASPic | Splice sites | Martelli et al. (2011) | https://algotlab.eu/research/aspic/ |
| RegRNA2 | Big range of motifs, including user defined | Chang et al. (2013) | http://regrna2.mbc.nctu.edu.tw/ |
| Mutation Taster2 | Protein and RNA changes | Schwarz et al. (2014) | http://www.mutationtaster.org/ |
| Imhotep | Many different aspects (incl. DNA, RNA, and protein level) | Knecht et al. (2017) | http://www.uni-kiel.de/medinfo/cgi-bin/predictor/ |
| AASites | Classifies probable change in splicing | Faber et al. (2011) | http://genius.embnet.dkfz-heidelberg.de/menu/biounit/open-husar |
| SNPlice | Effect of variants on intron retention | Mudvari et al. (2015) | https://code.google.com/archive/p/snplice/ |
| Spliceman | Effect of variants on splicing | Lim and Fairbrother (2012) | http://fairbrother.biomed.brown.edu/spliceman/ |
| ASSEDA | Effect of variants on splicing | Mucaki et al. (2013) | http://splice.uwo.ca/ |

Most algorithms have been generated based on different datasets and use different calculations, which explains why splicing predictions can vary substantially between different algorithms. Alamut[®] is a commercial platform that uses five different algorithms to predict the strengths of both 3' and 5' splice sites, as well as algorithms that detect potential binding sites of

multiple splicing factors. The program helps the user by visualizing the potential presence of these sites, an example of which can be found in [Bergsma et al. \(2016\)](#). Certain programs focus not only on variants that affect splicing, but also on changes to protein sequence and stability. One example is Mutationtaster, which integrates information from different biomedical databases to test evolutionary conservation, splice-site changes, loss of protein features, and changes that might affect the amount of mRNA ([Schwarz et al., 2010](#)). In the era of NGS, high-throughput data interpretation is becoming more and more essential. A new version, Mutationtaster2, has been generated to analyze multiple variants in a short period of time from deep-sequencing data ([Schwarz et al., 2014](#)). Another program that is well suited for multiple predictions of results from NGS datasets is combined annotation-dependent depletion (CADD), which incorporates multiple factors and gives a single pathogenicity score to variants ([Kircher et al., 2014](#)). More examples of splicing prediction programs include IntSplice for prediction of the splicing consequences of intronic single-nucleotide variations, and IMHOTEP, a composite score integrating popular tools for predicting the functional consequences of nonsynonymous sequence variants ([Knecht et al., 2017](#); [Shibata et al., 2016](#)). Interestingly, a comparison of IMHOTEP against a variety of other prediction tools, including CADD, showed that it was superior in predicting variants that affect splicing ([Knecht et al., 2017](#)). Finally, to generate more accurate predictions, splice prediction programs have been generated based on functional data ([Di Giacomo et al., 2013](#); [Soukarieh et al., 2016](#)). Although algorithms based on functional data will likely give a better representation of splicing in vivo, these two algorithms have been based on a relatively small dataset. With current RNA sequencing techniques, algorithms based on larger datasets can be generated, which will likely lead to better splicing prediction algorithms in the near future.

The programs and algorithms described above each have their benefits and drawbacks. Algorithms and prediction tools for pathogenicity will become increasingly important with implementation of NGS techniques in diagnostics. However, functional data that confirm these predictions are often lacking. In practice, predictions always should be verified using a functional assay to confirm a particular diagnosis.

4.3 Functional Assays

A widely used approach to evaluate functional effects that variants may have on splicing is to introduce part of the genomic region of the gene of interest

in an artificial minigene system (Fig. 5). This technique can be used to study splicing of genes that have extremely low or no detectable expression in patient material or if the material is not available. The system is utilized to evaluate splicing of well-described genes associated with disease, such as dystrophin, *CFTR*, *BRCA1*, and *BRCA2* (Bonnet et al., 2008; Giorgi et al., 2015; Wang et al., 2016). These assays may utilize a dedicated exon trapping vector in which only the exon of interest and short (30–150bp) flanking intronic regions are included. This system represents a quick and easy way to analyze the effect of selected variants on pre-mRNA splicing. Several recent studies have utilized minigene assays to validate the effect of variants on alternative splicing, including variants affecting splicing of *SLC26A4* mRNA causing autoimmune thyroid diseases (Bouattour et al., 2017) and the *PKD2* gene causing polycystic kidney disease (Gonzalez-Paredes et al., 2016). A disadvantage is that the assay is particularly biased, because the exon is placed outside its in vivo context, with altered transcription, transcript stability, and flanking intronic sequences. An improvement is to insert the exon of interest including neighboring introns and exons in an expression vector, provided that the total insert is not too long. Some of the above-mentioned disadvantages also apply to this option. Minigenes are useful tools to test potential pathogenic effects of variants, but care should be taken with translating results to regulation of splicing in vivo, as alternative splicing can differ between cell/tissue types.

With current NGS techniques, it is now possible to perform genome-wide mRNA expression analyses for diagnostic purposes to directly detect aberrant splicing in mRNA isolated from patient-derived cells (Fig. 5). Short sequence reads can be generated, from which splicing products need to be reassembled in silico, which can be a challenge. Numerous mapping protocols, programs, and pipelines have been generated to deal with this problem, and each of these have their own benefits and drawbacks (Engstrom et al., 2013). Recent advances have enabled long-read RNA sequencing, a third-generation sequencing technique, enabling sequencing of large stretches of contiguous mRNA avoiding in silico reconstruction of isoforms (reviewed in Rhoads and Au, 2015). This technique has led to the detection of many previously unknown splice isoforms (Sharon et al., 2013). Furthermore, it allows for the generation of full-length personal transcriptomes, highlighting the potential for clinical implementation (Tilgner et al., 2014).

Several challenges remain to be overcome before RNA-based NGS methods can be implemented in routine diagnostics (Byron et al., 2016). One reason for this is that it requires patient material, which can be

impractical, or even impossible to obtain, depending on the tissue in which the gene of interest is expressed. In addition, expression levels of the gene of interest need to be high enough to detect aberrant splicing, which will affect the required depth of sequencing. However, for certain genes with ubiquitous expression patterns, gene-tailored diagnostics based on mRNA analysis should be possible. For example, it is feasible to perform unbiased PCR-based analyses of all exons from fibroblast cultures to detect splicing variants of the acid α -glucosidase (*GAA*) gene causing Pompe disease, a metabolic myopathy (Fig. 5). This approach has already revealed novel splicing variants (Bergsma et al., 2015, 2016). In principle, it should be possible to develop diagnostic kits tailored for the detection of splicing variants in genes that are expressed in easy-to-obtain human material such as blood, urine, and skin. At present, PCR-based methods are the cheapest and easiest way for gene-tailored mRNA analysis, but in the future it is likely that these will be replaced by NGS-based techniques.

Recently, genome-editing techniques such as CRISPR/Cas9 have opened new possibilities for investigating functional effects of alternative splicing. CRISPR/Cas9 can be used to induce splicing variants to test for pathogenic effects, as exemplified in Zhang et al. (2016). Another application is the generation of transgenic mice to modulate expression of specific alternative transcripts, in which inhibition of expression of a *Titin* isoform was used to study the effects on M-line formation in striated muscle (Charton et al., 2016). The relative ease of generating transgenic animal models will likely lead to more functional studies of alternative splicing in relation to disease.

4.4 Ultraviolet Crosslinking and Immunoprecipitation

RBPs are key players in posttranslational gene regulation and RNA processing (Gerstberger et al., 2014). To understand the mechanisms involved, systematic identification of endogenous protein–RNA interactions is crucial. In 2003, Ule et al. developed a method for the genome-wide identification of RBP-binding sites called ultraviolet (UV) crosslinking and immunoprecipitation (CLIP) (Ule et al., 2003). This method relies on crosslinking of the protein–RNA complexes, followed by purification and release of RNA fragments by protein digestion. The complementary DNA (cDNA) library generated from these RNA fragments can be subjected to deep sequencing. Bioinformatic analysis of these data can reveal the RBP–RNA interactions. A variety of bioinformatics pipelines for the

handling of genome-wide CLIP data has been developed over the years, which has been extensively reviewed by [Wang et al. \(2015\)](#).

Since its introduction, the CLIP method has evolved into three distinct versions: (1) high-throughput sequencing of RNA isolated by CLIP (HITS-CLIP), the first CLIP platform developed for the genome-wide-identification of RBP binding sites ([Licatalosi et al., 2008](#)). (2) Photoactivatable-ribonucleoside enhanced CLIP (PAR-CLIP), which improved the signal-to-noise ratio by incorporating photoreactive ribonucleoside analogs into living cells via the culture system prior to UV treatment ([Hafner et al., 2010](#)). (3) Individual-nucleotide resolution UV CLIP (iCLIP), which is based on truncation of cDNAs at the crosslink site and identifies the positions of crosslink sites at single-nucleotide resolution ([Konig et al., 2010](#)). New variants such as the enhanced CLIP (eCLIP), which decreases PCR amplification by a 1000-fold ([Van Nostrand et al., 2016](#)), infrared-CLIP (irCLIP), which uses an infrared dye-conjugated and biotinylated ligation adaptor for visualization of UV-crosslinked protein-RNA complexes ([Zarnegar et al., 2016](#)), and bromodeoxyuridine CLIP (BrdU-CLIP), which incorporates BrdUTP into the cDNA during reverse transcription ([Weyn-Vanhentenryck et al., 2014](#)), have been developed more recently with undoubtedly many more variants to follow. With the introduction of CLIP and all its variants, the RNA-regulation field has rapidly developed and this has greatly increased our understanding of RBP-RNA interactions over the past few years.

This is well exemplified by recent studies on RBM10, a splicing regulatory protein. Wang et al. used PAR-CLIP to identify binding sites for RBM10 in HEK 293 cells. A large number of binding sites were identified and these were located in close proximity to splice sites. A role for RBM10 in splicing regulation was proposed, especially in alternative splicing resulting in exon skipping ([Wang et al., 2013](#)). Pathogenic variants in *RBM10* result in TARP (talipes equinovarus, atrial septal defect, robin sequence, and persistent left superior vena cava) syndrome. iCLIP was used to understand how loss of RBM10 function might result in TARP ([Rodor et al., 2017](#)). RNA targets of RBM10 in a mouse mandibular embryonic cell line were identified, and this revealed a clear enrichment of RBM10 binding in the proximity of the branch point and the 3' splice site of protein-coding genes. In addition, RBM10 binding to small nuclear RNAs (snRNAs) of canonical and minor spliceosomes was shown. Together with full transcriptome analysis of gene expression and splicing changes in mouse ESCs with disrupted RBM10 expression and RBM10 knockout mouse

mandibular cells, this revealed an important role of RBM10 in the regulation of alternative splicing. This suggests that TARP syndrome might be induced by loss-of-function variants in *RBM10* that cause dysregulation of alternative splicing and/or altered expression of other genes. One gene in which splicing changes were observed was OPA1, which is involved in optic atrophy that is also recognized in TARP syndrome (Rodor et al., 2017). In a different study iCLIP demonstrated the role of RBM10 in NUMB splicing regulation, a gene that is involved in the Notch signaling pathway and that is frequently altered in lung cancer. RBM10 variants were shown to promote cell growth by alternative splicing of NUMB pre-mRNA leading to lung adenocarcinoma progression (Bechara et al., 2013).

4.5 Are Splicing Variants Underestimated?

Of all variants that are listed in HGMD, 16,356 entries are known to alter splicing (HGMD, December 2016). These represent 9.1% of all disease-associated variants present in the database. There is increasing evidence that this percentage may be an underestimation for a number of reasons. In the early days of sequence analysis, functional analysis of reverse-transcribed mRNAs using PCR techniques was often employed for finding variants that alter the length of mRNA transcripts. The first dataset specific for pathogenic splice variants presented 101 variants, and these were thought to account for at least 15% of pathogenic variants (Krawczak et al., 1992). Sanger sequencing of genomic DNA is to date the golden standard in genetic diagnostics. Owing to the high number of common SNPs (minor allele frequency of >1%) that are present in noncoding intronic regions, exonic sequences are prioritized for detection of pathogenic variants, leading to an overrepresentation of variants other than splicing variants in current databases. Indeed, a recent survey shows that approximately 25% of missense and nonsense variants described in the HGM database may have an effect on splicing (Sterne-Weiler et al., 2011), suggesting that one-third of all pathogenic single-nucleotide variants in the database potentially affect splicing. Synonymous variants, that are usually disregarded as they do not change the amino acid sequence, can also have an effect on splicing. A study performed by Supek et al. indicates that synonymous variants occur more often in oncogenes where they might act as driver variants by changing SREs, ultimately resulting in altered splicing (Supek et al., 2014).

Functional investigation of all known variants present in a single exon in *BRCA2* and *MLH1* yielded a high percentage of variants that affect splicing

(42% and 77%, respectively) (Di Giacomo et al., 2013; Soukarieh et al., 2016). One mathematical model even suggests that up to 62% of all disease-causing variants affect splicing (Lopez-Bigas et al., 2005). These predictions should be further investigated at the functional level.

A more comprehensive understanding of the splicing code will be necessary to estimate the impact that variants have on splicing in general. A recent study showed that disease-causing variants located in deep intronic regions affect splicing nine times more often than SNPs (Xiong et al., 2015). Furthermore, if missense variants have a low predicted impact on protein function, the likelihood that these variants lead to aberrant splicing is five times higher.

As well as rare variants, polymorphisms can also impact splicing. It has been estimated that up to 45% of synonymous SNPs potentially alter pre-mRNA splicing patterns (Mueller et al., 2015). Furthermore, splicing quantitative trait loci (sQTLs), which are regions in the DNA that correlate with splicing patterns and are defined by the presence of SNPs, are thought to be important contributors to phenotypic variation (Li et al., 2016). In a recent study, Takata et al. analyzed tissue from the prefrontal cortex of 206 individuals and identified genome-wide alternative splicing events controlled by sQTLs (Takata et al., 2017). Investigation of sQTLs SNPs located in schizophrenia-associated loci linked to four potential candidate genes which underwent alternative splicing and could be implicated in schizophrenia.

The data highlighted earlier indicate that the effect of (pathogenic) variants on splicing of pre-mRNA may be considerably underestimated and warrants future investigation.



5. THERAPEUTIC APPROACHES TO MODULATE SPLICING

Human disease caused by aberrant splicing is amenable to modulation in several ways. In recent years, new developments have presented promising therapeutic strategies.

5.1 Small Molecule Modulators of Splicing

Small molecules that modulate alternative splicing or inhibit the major spliceosome have been reported for the treatment of cancer and monogenic disorders. The advantages of using small molecules include that these can be administered orally, and that these may cross the blood–brain barrier.

In cancer, splicing regulatory sequences in (proto-)oncogenes can be mutated, and/or splicing factors can have altered expression. This can cause

alternative splicing and induce resistance to cancer drugs, or it can lead to escape of cancers cells from the immune system. This has been extensively reviewed in [Lee and Abdel-Wahab \(2016\)](#) and [Salton and Misteli \(2016\)](#) ([Fig. 6](#)). Small molecules have been developed that block early spliceosome formation via binding to the splicing factor 3B (SF3B) core spliceosome subunit, or that alter phosphorylation of splicing regulatory SR proteins. One of the earliest drugs to modulate splicing in cancer is the compound spliceostatin A, which targets the SF3B subunit of U2 snRNP, leading to cell cycle arrest in G1 and G2/M ([Kaida et al., 2007](#)). Unfortunately, a clinical trial with the SF3B inhibitor E7107 was suspended due to toxicity ([Hong et al., 2014](#)). Currently, one trial is recruiting patients for treatment with H3B-8800 ([NCT02841540](#)), another SF3B inhibitor.

Many small molecules have been reported that modulate aberrant splicing in monogenic disorders in vitro, but preclinical safety and efficacy for most of these compounds remain to be established (reviewed in [Bates et al., 2017](#); [Ohe and Hagiwara, 2015](#)). It is worth noting that high-throughput screens with compounds that are already used in the clinic may identify other compounds that also modulate splicing. An example is

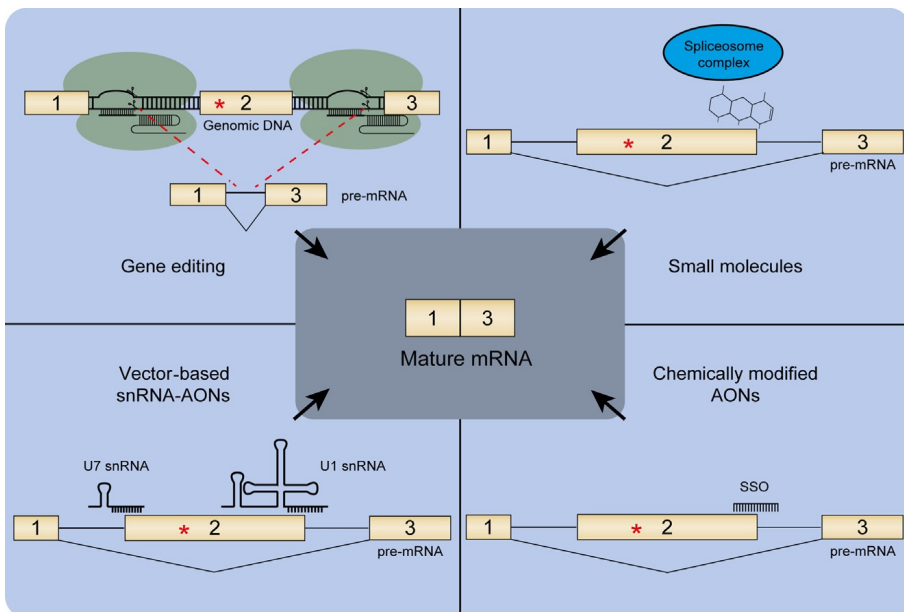


Fig. 6 Potential therapeutic strategies to interfere with splicing. *Red asterisks* indicate a hypothetical pathogenic variant. Potential therapeutic interventions include deletion of an exon from the DNA using gene-editing tools, small molecules that interfere with the activity of splicing proteins, vector-based snRNA-AONs, and chemically modified AONs.

provided by cardiotonic steroids. It was demonstrated that these regulate alternative splicing via depletion of SRSF3 and TRA2B, resulting in increased exon 10 inclusion in microtubule-associated protein tau (*MAPT*) in frontotemporal dementia in vitro (Anderson et al., 2012; Stoilov et al., 2008). Another example is DM type 1 (DM1), in which the master splicing factor MBNL1 is sequestered and depleted by CTG repeats in the myotonin protein kinase transcript, resulting in defects in alternative splicing. Metformin (Biguanides), which is commonly used for the treatment of type 2 diabetes, was shown to downregulate the master splicing factor RBM3 and alleviate some alternative splicing defects in DM1 (Laustriat et al., 2015). For SMA (discussed in more detail below), multiple splice-switching compounds were identified using large-scale drug screens. An example is Aclarubicin, which promotes inclusion of exon 7 of *SMN2* (Andreassi et al., 2001). More recently, additional molecules have been identified that can alter *SMN2* splicing to promote exon 7 inclusion using high-throughput screening. These include SMN-C1, SMN-C2, SMN-C3 (Naryshkin et al., 2014), and NVS-SM1 (Palacino et al., 2015). Both studies demonstrated enhanced survival of humanized *SMN7*delta mice after treatment with these compounds. Global gene expression analysis showed no major off-target effects, suggesting that these small molecules were selective. Lead optimization of SMN-C3 from Naryshkin et al. (2014) resulted in the compound RG7800, which was used in the clinical trial NCT02240355 (Ratni et al., 2016). This trial was unexpectedly terminated because of toxic effects found after 39 weeks in a parallel monkey study. To generate new compounds, chemical optimizations were performed on hits from Naryshkin et al. (2014), and a safety study was completed on a derivative of one of these hits (NCT02633709, Woll et al., 2016). NVS-SM1 described in Palacino et al. (2015) is currently used in a clinical trial to evaluate safety and efficacy (NCT02268552, Palacino et al., 2015). These findings highlight the potential of using small molecules for targeting splicing regulators for the treatment of monogenetic disorders and especially in the case of SMA. The challenge for the future will be to identify compounds with high selectivity and low toxicity, and to understand their mechanism of action.

5.2 Transsplicing

Another option for correcting aberrant splicing caused by the presence of splice altering variants is spliceosome-mediated RNA transsplicing (SMaRT) (reviewed in Berger et al., 2016). This technique is based on the property of the spliceosome to enable the combination of transcripts

from two independent RNA molecules to form one mRNA transcript. Several applications of SMaRT have been reported. Examples include cystic fibrosis (Mansfield et al., 2000), Huntington's disease (Rindt et al., 2012), inherited tauopathy (Avalé et al., 2013), and p53-defective cancers (He et al., 2015). The process of transsplicing rarely occurs in humans (Lei et al., 2016), and the challenge will be to increase the efficiency of this approach for clinical development.

5.3 Antisense Oligonucleotides

A specific way to influence gene expression can be accomplished with short (~15–30 nt) RNA-based molecules termed AONs that bind to RNA in a sequence-dependent manner. Several types of AONs can be distinguished. These include single-stranded and double-stranded AONs for mRNA degradation, and splice-switching oligonucleotides (SSOs). To induce mRNA degradation, either the RNA interference (RNAi) pathway (using dsRNA) or the RNase H pathway (using ssRNA) may be utilized. SSOs rely on the ability to bind to splicing regulator sequences at the pre-mRNA to modulate splicing outcome. AONs have chemically modified backbones to prevent degradation. The type of modification has consequences for the mechanism of action (e.g., (in)activity toward RNase H and RNAi pathways) and for biodistribution and pharmacokinetics. The difficulty to deliver RNA-based molecules to cells is an area of active research, including the development of conjugates and chemistries that improve cellular delivery.

In RNAi, double-stranded RNA is administered to cells, which is then processed and loaded on the RNA-induced silencing complex (RISC). The RNA binds to a target mRNA, which results in its degradation, or inhibition of protein translation (Elbashir et al., 2001). Although no RNAi drugs are on the market thus far, several are currently tested in clinical trials. The most clinically advanced example is RNAi-mediated treatment of diabetic macular edema, which is currently in Phase II trials (NCT01445899). RNAi has been used to modulate splicing of alternative exons by targeting sequences that are located in the vicinity of those exons (Allo et al., 2009). Although RNAi has been extensively investigated as a potential therapy, doubts have been raised on its clinical applicability (Krieg, 2011). Off-target effects are a concern (Jackson and Linsley, 2010). Moreover, the administration of double-stranded RNA can lead to an innate immune response via recognition of double-stranded RNA by Toll-like receptors (de Fougerolles et al., 2007). These drawbacks need to be overcome before RNAi can be applied to the clinic.

Single-stranded AONs can target RNA after which these are recognized by the endogenous RNA degradation pathways involving RNase H. The single-stranded AON is usually modified to prevent its degradation; however, these modifications often interfere with the RNase H pathway. To prevent this, modifications to the backbone of the RNA molecule are made at the nucleotides on the outside of the molecule, the result of which is termed a gapmer (Yu et al., 2013). In this way, exonucleases do not degrade the RNA target. RNA bases in the middle of the AON are not modified so that they are still recognized by RNase H. One of the antisense gapmers under development is Mipomersen for intervention in patients with heterozygous familial hypercholesterolemia. This second-generation AON gapmer targets the splice acceptor of exon 21 in the *APOB* gene, causing both exon skipping of exon 21 and cleavage of the mRNA by RNase H, both resulting in downregulation of expression of the apolipoprotein B enzyme (Kastelein et al., 2006). First indications of clinical trials using subcutaneous administration ($3 \times$ weekly) revealed that indeed ApoB production was lower in treated patients (Akdim et al., 2010; Akdim et al., 2011; Kastelein et al., 2006) (NCT00216463, NCT00231569). However, adverse events also occurred, including injection site reactions, increased liver fat percentage, and increased serum transaminase levels. Four Phase III clinical trials (NCT00607373, NCT00706849, NCT00794664, and NCT00694109) indicated that mipomersen treatment resulted in lower levels of LDL-C, non-HDL cholesterol and Lp(a). Furthermore, a dramatic reduction in cardiac events was seen, indicating that the drug has beneficial therapeutic effects (Duell et al., 2016). The drug was FDA approved in 2013 for homozygous hypercholesterolemia, however, it should be stated that it received a black box label for potential hepatotoxicity.

SSOs are designed to modulate pre-mRNA splicing (Fig. 6). These typically utilize chemically modified AONs that do not activate the RNase H or RNAi pathways. The first SSOs carried a 2'-O-methylphosphorotioate (2'-O-mePS) backbone and were targeted against cryptic splice sites in mRNAs derived from pathogenic variants in the β -globin gene (Dominski and Kole, 1993; Sierakowska et al., 1996). Currently, a host of different backbone structures are under investigation for application in the clinic. Frequently used backbone chemistries include 2'-O-mePS, 2'-O-methoxyethyl (MOE), phosphorodiamidate morpholino oligomer (PMO), and peptide nucleic acids (Geary et al., 2015). Recently, a new type of backbone chemistry, tricyclo-DNA, was suggested to enhance cellular uptake (Goyenvall et al., 2015). Methods to improve uptake include the addition

of cell-penetrating peptides (reviewed in [Boisguerin et al., 2015](#)) and delivery via nanoparticles (reviewed in [Petrilli et al., 2014](#)). Furthermore, coadministration of specific monosaccharides can also lead to improved uptake ([Cao et al., 2016](#); [Han et al., 2016](#)). For many disorders, preclinical development of SSO-mediated therapy has been performed, including Hutchinson–Gilford progeria syndrome ([Osorio et al., 2011](#)), type I Usher syndrome ([Lentz et al., 2013](#)), cystic fibrosis ([Igreja et al., 2016](#)), Leber congenital amaurosis ([Garanto et al., 2016](#)), certain types of cancer (reviewed in [Farooqi et al., 2014](#)), and Pompe disease. In the latter, a substrate reduction strategy used SSO-mediated reduction of glycogen synthase expression ([Clayton et al., 2014](#)), while direct approaches used promotion of canonical *GAA* splicing via inhibition of cryptic splicing ([Bergsma et al., 2016](#); [Van der Wal et al., 2017a, b](#)).

For two diseases, DMD and SMA, clinical development of AON therapy have also passed Phase III trials ([CliniclaTrials.gov](#) identifiers [NCT02255552](#) and [NCT02193074](#), respectively) after extensive preclinical work (reviewed in [Aartsma-Rus, 2010](#); [Singh et al., 2015](#)) and both drugs have been approved by the FDA. Eteplirsen, a PMO-based AON for DMD was approved by the FDA in September 2016 for IV administration (dose of 30 mg/kg weekly) ([Niks and Aartsma-Rus, 2017](#)). Nusinersen, an MOE phosphorothioate-modified AON for SMA, was approved by the FDA in December 2016 for intrathecal administration (three loading doses of 12 mg/kgL, fourth dose 30 days later, and every 3 months thereafter) ([Chiriboga et al., 2017](#)). New generation AONs are currently being developed to improve uptake and lower toxicity. For example, cellular uptake in heart, skeletal muscle, and the CNS is still inefficient following intravenous administration, and it can be anticipated that new generation SSOs with improved properties with respect to cellular uptake will be tested in future clinical trials. For SMA a new type of AON containing a morpholino backbone and a conjugated peptide showed improved uptake in vivo in a mouse model ([Shabanpoor et al., 2017](#)). Increasing amounts of reports now focus on improved AON therapeutic targeting strategies for previous and novel targets ([d'Ydewalle et al., 2017](#); [Osman et al., 2016](#)).

5.4 Gene Therapy

5.4.1 *U1/U7 snRNAs*

Delivery of AONs into target tissues is challenging, and recent advances in viral-based gene therapy can overcome this limitation ([Naldini, 2015](#)). Naturally occurring snRNAs, which are normally involved in processing of

pre-mRNAs, can be used for viral vector-based delivery of antisense sequences. U1 and U7 snRNAs have been used to modulate splicing (Imbert et al., 2017) (Fig. 6). The U7 snRNA is normally utilized in processing of histone pre-mRNAs. When the SM binding site is modified, it can function to target and block any desired sequence on a pre-mRNA of choice (Soldati and Schumperli, 1988 reviewed in Schumperli and Pillai, 2004). Natural U1 snRNA consists of four stem loops (sl) that bind the core spliceosomal components U1–70k (sl 1), U1-A (sl 2), U1-C (sl 3), and U2 snRNP (sl 4). It also contains an SM binding site for SM proteins. This assembly generates a functional U1 snRNP, which is involved in intron definition by binding to splice donor sites (Lee and Rio, 2015; Stark et al., 2001). Bases 3–10 of the U1 snRNA are involved in this binding and can be modified by insertion of up to 50 bases and then targeted to any desired sequence. For example, targeting to mutated (cryptic) splice donor sites can result in promotion of intron definition. It should be noted that intronic splice donor sites are often well conserved and therefore potential off-targets should be determined.

Evidence that the modified U1/U7 snRNA system is capable of targeting a heterologous pre-mRNA and to modulate its splicing was shown for beta thalassemia. The IVS2-705T>G variant in the pre-mRNA of *β-globin* generates a cryptic splice site. Targeting this cryptic splice site with U1 or U7 snRNA in the pre-mRNA restored canonical splicing to up 56% and 65% of healthy individuals, respectively (Gorman et al., 1998, 2000). Many studies used the U7 snRNA system to modulate splicing in vitro. The targeting strategy is similar to that used with chemical stable AONs and the efficacy can be comparable. The U1 snRNA targeting implements a different strategy as it is semirestricted to the splice donor site. Proof of principle for a U1 snRNA-mediated therapy for SMA was reported by Rogalska et al., who demonstrated increased survival in a mouse model. Untreated mice died within 6 days, while insertion of a single germline copy of exon-specific U1 snRNA resulted in increased survival up to at least 250 days. Off-target effects were reduced by targeting a nonconserved intronic sequence upstream of the splice donor site, resulting in 12 genes with changed expression (Rogalska et al., 2016).

For in vivo targeting, the modified U1/U7 snRNA system can be inserted into different viruses including retroviruses, lentiviruses, adenoviruses, and recombinant adeno-associated viruses (rAAVs) (extensively reviewed in Imbert et al., 2017). rAAV is of particular interest because it can be targeted to a wide range of tissue types using different tropisms.

Furthermore, it is relative safe as it has low likelihood to integrate into the genome, and the transgene can be expressed over prolonged periods up to years. Disadvantages include antibody formation against the capsid (Kotterman and Schaffer, 2014), and, because rAAV is a noninsertional vector, decreased transgene expression over time due to cell division and cell loss. A long-term follow-up study of golden retriever muscular dystrophy dogs using rAAV-U7 snRNA for the skipping of exons 6 and 8, which was applied to skeletal muscle, demonstrated loss of rAAV copy number, and dystrophin positive fibers during a period of 5 years (Vulin et al., 2012). This may have been caused by turnover of transduced fibers, as the exon skipping strategy converted the DMD phenotype to a milder Becker muscular dystrophy (BMD) phenotype, which still has ongoing muscle wasting and regeneration.

5.4.2 Gene Editing

Permanent modulation of splicing has recently become possible with the development of gene-editing techniques, which use designer nucleases such as ZFNs, TALENs, or CRISPR/Cas9 (Fig. 6). These can generate double-strand breaks at predetermined locations in genomic DNA (Maggio and Goncalves, 2015). Repair of these breaks is performed by the host, either by using the nonhomologous end joining (NHEJ) pathway or by homology directed repair (HDR). The NHEJ pathway can be exploited to knockout genes (by disrupting the reading frame) or to generate deletions, while the HDR pathway requires a donor template with homology arms that can be used to repair or modify target genes. The NHEJ pathway is also active in nondividing cells, which allows gene editing of terminally differentiated cells (Lieber, 2010). The field is currently addressing the low efficiencies of HDR in nondividing cells, and a recent study demonstrated that transgene integration via homology-independent targeted integration (HITI) was possible via the NHEJ pathway (Suzuki et al., 2016).

Strategies that apply to splicing may include restoration of the reading frame by generation of a small nucleotide deletion or insertion, deletion of an exon to bypass variation hotspots and to restore the reading frame, or disruption of cryptic splice sites. To apply this strategy, designer nucleases have to be delivered to the target cells *in vivo* or *ex vivo*. *In vitro* studies demonstrated the potential of this approach by reframing dystrophin transcripts via removal of various exons to produce a shorter but functional protein, which is associated with a milder BMD phenotype (Iyombe-Engembe et al., 2016; Maggio et al., 2016; Young et al., 2016).

The first proof of principle study for in vivo targeting used either electroporation or adenoviral infection to deliver CRISPR/Cas9 in *mdx* mice. This successfully removed exon 23 in the dystrophin gene (Xu et al., 2016). Using the smaller SaCas9 (Ran et al., 2015) three studies demonstrated in vivo delivery of CRISPR/cas9 to muscle and heart tissue using systemic administration of rAAV8/9 (Long et al., 2016; Nelson et al., 2016; Tabebordbar et al., 2016). Muscle strength and force measurements showed improvements as compared with untreated *mdx* mice. Tabebordbar et al. also demonstrated that muscle stem cells were targeted, which is important for regeneration of myofibers with a reframed dystrophin protein (Tabebordbar et al., 2016). These studies presented a novel approach to modulate splicing using designer nucleases. Challenges for the future include the efficiency of in vivo gene editing, and the safety profile of such methods. In particular, undesired double-stranded DNA breaks generated by off-target effects would pose a threat for the development of tumorigenic cells. Ex vivo application of gene editing followed by cell engraftment is currently under clinical development for the treatment of HIV (Tebas et al., 2014), and has so far not reported serious adverse events.



6. FUTURE PERSPECTIVES

This review has highlighted the complexity of splicing and its crucial role in gene regulation in healthy and diseased individuals. Yet, our understanding of how sequence variants may affect splicing outcome and how this relates to human disease is at its infancy. Our ability to detect sequence variants at the DNA level have reached the previously unthinkable milestone of WGS at an affordable price and high speed. This has also exponentially increased the number of questions regarding pathogenicity of sequence variants. Currently, DNA NGS is mostly confined to gene panels or whole exomes, but in the near future it can be expected that WGS analysis will be increasingly applied, provided that the costs will further decrease. This will put an even larger pressure on our ability to assess which variants may be pathogenic. Improved in silico prediction programs would be important. Next-generation programs would be needed that take into account the different layers at which splicing is regulated, including expression and activity of splicing factors, presence, and strengths of cis acting elements, transcription rate, chromatin environment, activation of decay pathways, and other splicing mechanisms such as RS.

Functional assays to detect aberrant splicing will continue to be essential to validate prediction programs, but also to improve these. Single gene-based assays using PCR analysis of all exons are fast, simple, and feasible for a subset of genes with ubiquitous expression. In the future, it can be expected that PCR analysis will be replaced by NGS-based analysis. Upcoming is genome-wide analysis of mRNA transcripts using long-read RNAseq, which allow the analysis of isoforms without *in silico* reconstruction from short reads, and can even be analyzed on the single cell level (Karlsson and Linnarsson, 2017). In addition, CLIP-based technology to detect RNA–protein interactions will remain a crucial approach to understand splicing mechanisms, and optimization and diversifications of CLIP will expand our options for in-depth mechanistic studies.

Another challenge for the future is to link splicing outcome to disease. Is the variant dominant negative? Is there leaky wild-type splicing? How much normal product is needed to correct the disease? Why are certain organs/tissues particularly sensitive to variants in general splicing factors? Can we link splicing outcome to disease severity? The answers to these questions will be helpful to develop novel treatment options. While a number of options are under development and some have entered the clinic, therapeutic intervention with splicing still faces some important challenges. For AONs, cellular uptake and tolerability are key issues. The development of next-generation AONs with improved properties is promising and may play an important role in upcoming trials for a variety of disorders. Small molecule drugs that interfere with expression or activity of splicing factors face the challenge of specificity. Gene therapy is a promising, rapidly developing field, and potentially offers a long- or life-term treatment option. The challenges depend on the type of gene therapy. Safety with respect to genome integrity and immune responses to viral vectors, and long-term efficacy are important aspects that need to be addressed. Current AAV and LV vectors appear effective and safe in a number of clinical trials. The excellent progress that has been made in the past few years in splicing research is expected to result in improved DNA and RNA-based diagnosis and methods to predict or determine splicing outcome, a better understanding of splicing mechanisms, and clinical development of more treatment options for human diseases based on splicing correction.

ACKNOWLEDGMENTS

We thank Dr. Lies Hoefsloot for discussion. We thank Katherine MacKenzie and Pablo Herrero Hernández for critical reading of the manuscript. The authors are

supported by grants from Sophia Children's Hospital Foundation (SSWO, grant S17-32), Stichting Metakids, The Prinses Beatrix Spierfonds/Stichting Spieren voor Spieren (grant W.OR13-21), Tex Net, the WE Foundation, ZonMW—the Netherlands Organization for Health Research and Development (project number 152001005), and by the Erasmus MC.

REFERENCES

- Aartsma-Rus, A., 2010. Antisense-mediated modulation of splicing: therapeutic implications for Duchenne muscular dystrophy. *RNA Biol.* 7 (4), 453–461.
- Adelman, K., Lis, J.T., 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* 13 (10), 720–731.
- Akdin, F., Stroes, E.S., Sijbrands, E.J., Tribble, D.L., Trip, M.D., Jukema, J.W., Flaim, J.D., Su, J., Yu, R., Baker, B.F., Wedel, M.K., Kastelein, J.J., 2010. Efficacy and safety of mipomersen, an antisense inhibitor of apolipoprotein B, in hypercholesterolemic subjects receiving stable statin therapy. *J. Am. Coll. Cardiol.* 55 (15), 1611–1618.
- Akdin, F., Tribble, D.L., Flaim, J.D., Yu, R., Su, J., Geary, R.S., Baker, B.F., Fuhr, R., Wedel, M.K., Kastelein, J.J., 2011. Efficacy of apolipoprotein B synthesis inhibition in subjects with mild-to-moderate hyperlipidaemia. *Eur. Heart J.* 32 (21), 2650–2659.
- Allo, M., Buggiano, V., Fededa, J.P., Petrillo, E., Schor, I., de la Mata, M., Agirre, E., Plass, M., Eyra, E., Elela, S.A., Klinck, R., Chabot, B., Kornblihtt, A.R., 2009. Control of alternative splicing through siRNA-mediated transcriptional gene silencing. *Nat. Struct. Mol. Biol.* 16 (7), 717–724.
- Amin, N., Allebrandt, K.V., van der Spek, A., Muller-Myhsok, B., Hek, K., Teder-Laving, M., Hayward, C., Esko, T., van Mill, J.G., Mbarek, H., Watson, N.F., Melville, S.A., Del Greco, F.M., Byrne, E.M., Oole, E., Kolcic, I., Chen, T.H., Evans, D.S., Coresh, J., Vogelzangs, N., Karjalainen, J., Willemsen, G., Gharib, S.A., Zgaga, L., Mihailov, E., Stone, K.L., Campbell, H., Brouwer, R.W., Demirkan, A., Isaacs, A., Dogas, Z., Marcante, K.D., Campbell, S., Borovecki, F., Luik, A.I., Li, M., Hottenga, J.J., Huffman, J.E., van den Hout, M.C., Cummings, S.R., Aulchenko, Y.S., Gehrman, P.R., Uitterlinden, A.G., Wichmann, H.E., Muller-Nurasyid, M., Fehrmann, R.S., Montgomery, G.W., Hofman, A., Kao, W.H., Oostra, B.A., Wright, A.F., Vink, J.M., Wilson, J.F., Pramstaller, P.P., Hicks, A.A., Polasek, O., Punjabi, N.M., Redline, S., Psaty, B.M., Heath, A.C., Merrow, M., Tranah, G.J., Gottlieb, D.J., Boomsma, D.I., Martin, N.G., Rudan, I., Tiemeier, H., van, I.W.F., Penninx, B.W., Metspalu, A., Meitinger, T., Franke, L., Roenneberg, T., van Duijn, C.M., 2016. Genetic variants in RBFOX3 are associated with sleep latency. *Eur. J. Hum. Genet.* 24 (10), 1488–1495.
- Anantha, R.W., Alcivar, A.L., Ma, J., Cai, H., Simhadri, S., Ule, J., Konig, J., Xia, B., 2013. Requirement of heterogeneous nuclear ribonucleoprotein C for BRCA gene expression and homologous recombination. *PLoS One* 8 (4), e61368.
- Anderson, E.S., Lin, C.H., Xiao, X., Stoilov, P., Burge, C.B., Black, D.L., 2012. The cardiotonic steroid digitoxin regulates alternative splicing through depletion of the splicing factors SRSF3 and TRA2B. *RNA* 18 (5), 1041–1049.
- Andersson, R., Enroth, S., Rada-Iglesias, A., Wadelius, C., Komorowski, J., 2009. Nucleosomes are well positioned in exons and carry characteristic histone modifications. *Genome Res.* 19 (10), 1732–1741.
- Andreassi, C., Jarecki, J., Zhou, J., Covert, D.D., Monani, U.R., Chen, X., Whitney, M., Pollok, B., Zhang, M., Androphy, E., Burghes, A.H., 2001. Aclurubicin treatment restores SMN levels to cells derived from type I spinal muscular atrophy patients. *Hum. Mol. Genet.* 10 (24), 2841–2849.
- Argente, J., Flores, R., Gutierrez-Arumi, A., Verma, B., Martos-Moreno, G.A., Cusco, I., Oghabian, A., Chowen, J.A., Frilander, M.J., Perez-Jurado, L.A., 2014. Defective minor

- spliceosome mRNA processing results in isolated familial growth hormone deficiency. *EMBO Mol. Med.* 6 (3), 299–306.
- Avale, M.E., Rodriguez-Martin, T., Gallo, J.M., 2013. Trans-splicing correction of tau isoform imbalance in a mouse model of tau mis-splicing. *Hum. Mol. Genet.* 22 (13), 2603–2611.
- Ballester, L.Y., Luthra, R., Kanagal-Shamanna, R., Singh, R.R., 2016. Advances in clinical next-generation sequencing: target enrichment and sequencing technologies. *Expert Rev. Mol. Diagn.* 16 (3), 357–372.
- Balthazar, P., Tucunduva, D.C., Luce, M.J., Scott, M.S., 2017. Widespread pre-translational regulation of the inclusion of signal peptides in human proteins. *Genomics* 109, 113–122.
- Barash, Y., Vaquero-Garcia, J., Gonzalez-Vallinas, J., Xiong, H.Y., Gao, W., Lee, L.J., Frey, B.J., 2013. AVISPA: a web tool for the prediction and analysis of alternative splicing. *Genome Biol.* 14 (10), R114.
- Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C.M., Wilson, M.D., Kim, P.M., Odom, D.T., Frey, B.J., Blencowe, B.J., 2012. The evolutionary landscape of alternative splicing in vertebrate species. *Science* 338 (6114), 1587–1593.
- Bates, D.O., Morris, J.C., Oltean, S., Donaldson, L.F., 2017. Pharmacology of modulators of alternative splicing. *Pharmacol. Rev.* 69 (1), 63–79.
- Beaulieu, C.L., Majewski, J., Schwartzentruber, J., Samuels, M.E., Fernandez, B.A., Bernier, F.P., Brudno, M., Knoppers, B., Marcadier, J., Dymment, D., Adam, S., Bulman, D.E., Jones, S.J., Avar, D., Nguyen, M.T., Rousseau, F., Marshall, C., Wintle, R.F., Shen, Y., Scherer, S.W., Consortium, F.C., Friedman, J.M., Michaud, J.L., Boycott, K.M., 2014. FORGE Canada consortium: outcomes of a 2-year national rare-disease gene-discovery project. *Am. J. Hum. Genet.* 94 (6), 809–817.
- Bechara, E.G., Sebestyen, E., Bernardis, I., Eyra, E., Valcarcel, J., 2013. RBM5, 6, and 10 differentially regulate NUMB alternative splicing to control cancer cell proliferation. *Mol. Cell* 52 (5), 720–733.
- Bengtson, M.H., Joazeiro, C.A., 2010. Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467 (7314), 470–473.
- Berger, A., Maire, S., Gaillard, M.C., Sahel, J.A., Hantraye, P., Bemelmans, A.P., 2016. mRNA trans-splicing in gene therapy for genetic diseases. *Wiley Interdiscip. Rev. RNA* 7 (4), 487–498.
- Bergsma, A.J., Kroos, M., Hoogeveen-Westerveld, M., Halley, D., van der Ploeg, A.T., Pijnappel, W.W., 2015. Identification and characterization of aberrant GAA pre-mRNA splicing in Pompe disease using a generic approach. *Hum. Mutat.* 36 (1), 57–68.
- Bergsma, A.J., In 't Groen, S.L., Verheijen, F.W., van der Ploeg, A.T., Pijnappel, W.P., 2016. From cryptic toward canonical pre-mRNA splicing in Pompe disease: a pipeline for the development of antisense oligonucleotides. *Mol. Ther. Nucleic Acids* 5, e361.
- Bhalla, K., Phillips, H.A., Crawford, J., McKenzie, O.L., Mulley, J.C., Eyre, H., Gardner, A.E., Kremmidiotis, G., Callen, D.F., 2004. The de novo chromosome 16 translocations of two patients with abnormal phenotypes (mental retardation and epilepsy) disrupt the A2BP1 gene. *J. Hum. Genet.* 49 (6), 308–311.
- Blanchette, M., Chabot, B., 1999. Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J.* 18 (7), 1939–1952.
- Boisguerin, P., O'Donovan, L., Gait, M.J., Lebleu, B., 2015. In vitro assays to assess exon skipping in Duchenne muscular dystrophy. *Methods Mol. Biol.* 1324, 317–329.
- Bonnet, C., Krieger, S., Vezain, M., Rousselin, A., Tournier, I., Martins, A., Berthet, P., Chevrier, A., Dugast, C., Layet, V., Rossi, A., Lidereau, R., Frebourg, T.,

- Hardouin, A., Tosi, M., 2008. Screening BRCA1 and BRCA2 unclassified variants for splicing mutations using reverse transcription PCR on patient RNA and an ex vivo assay based on a splicing reporter minigene. *J. Med. Genet.* 45 (7), 438–446.
- Bouattour, R.K., Maalej, S.B., Zouari-Bradai, E., Mnif, M., Abid, M., Kacem, H.H., 2017. Intronic variants of SLC26A4 gene enhance splicing efficiency in hybrid minigene assay. *Gene* 620, 10–14.
- Bradnam, K.R., Korf, I., 2008. Longer first introns are a general property of eukaryotic gene structure. *PLoS One* 3 (8), e3093.
- Brauch, K.M., Karst, M.L., Herron, K.J., de Andrade, M., Pellikka, P.A., Rodeheffer, R.J., Michels, V.V., Olson, T.M., 2009. Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *J. Am. Coll. Cardiol.* 54 (10), 930–941.
- Brendel, V., Kleffe, J., 1998. Prediction of locally optimal splice sites in plant pre-mRNA with applications to gene identification in *Arabidopsis thaliana* genomic DNA. *Nucleic Acids Res.* 26 (20), 4748–4757.
- Burge, C., Karlin, S., 1997. Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* 268 (1), 78–94.
- Byron, S.A., Van Keuren-Jensen, K.R., Engelthaler, D.M., Carpten, J.D., Craig, D.W., 2016. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. *Nat. Rev. Genet.* 17 (5), 257–271.
- Cao, L., Han, G., Lin, C., Gu, B., Gao, X., Moulton, H.M., Seow, Y., Yin, H., 2016. Fructose promotes uptake and activity of oligonucleotides with different chemistries in a context-dependent manner in mdx mice. *Mol. Ther. Nucleic Acids* 5, e329.
- Cardani, R., Giagnacovo, M., Rossi, G., Renna, L.V., Bugiardini, E., Pizzamiglio, C., Botta, A., Meola, G., 2014. Progression of muscle histopathology but not of spliceopathy in myotonic dystrophy type 2. *Neuromuscul. Disord.* 24 (12), 1042–1053.
- Carpenter, S., Ricci, E.P., Mercier, B.C., Moore, M.J., Fitzgerald, K.A., 2014. Post-transcriptional regulation of gene expression in innate immunity. *Nat. Rev. Immunol.* 14 (6), 361–376.
- Carrigan, M., Duignan, E., Malone, C.P., Stephenson, K., Saad, T., McDermott, C., Green, A., Keegan, D., Humphries, P., Kenna, P.F., Farrar, G.J., 2016. Panel-based population next-generation sequencing for inherited retinal degenerations. *Sci. Rep.* 6, 33248.
- Cartegni, L., Wang, J., Zhu, Z., Zhang, M.Q., Krainer, A.R., 2003. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Res.* 31 (13), 3568–3571.
- Cartegni, L., Hastings, M.L., Calarco, J.A., de Stanchina, E., Krainer, A.R., 2006. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. *Am. J. Hum. Genet.* 78 (1), 63–77.
- Chang, T.H., Huang, H.Y., Hsu, J.B., Weng, S.L., Horng, J.T., Huang, H.D., 2013. An enhanced computational platform for investigating the roles of regulatory RNA and for identifying functional RNA motifs. *BMC Bioinf.* 14 (Suppl. 2), S4.
- Charton, K., Suel, L., Henriques, S.F., Moussu, J.P., Bovolenta, M., Taillepierre, M., Becker, C., Lipson, K., Richard, I., 2016. Exploiting the CRISPR/Cas9 system to study alternative splicing in vivo: application to titin. *Hum. Mol. Genet.* 25 (20), 4518–4532.
- Chiriboga, A.S., Konno, T., van Gerpen, J.A., 2017. Personality changes, executive dysfunction, and motor and memory impairment. *JAMA Neurol.* 74 (2), 245–246.
- Choi, Y.D., Grabowski, P.J., Sharp, P.A., Dreyfuss, G., 1986. Heterogeneous nuclear ribonucleoproteins: role in RNA splicing. *Science* 231 (4745), 1534–1539.
- Chong, J.X., Buckingham, K.J., Jhangiani, S.N., Boehm, C., Sobreira, N., Smith, J.D., Harrell, T.M., McMillin, M.J., Wiszniewski, W., Gambin, T., Coban Akdemir, Z.H., Doheny, K., Scott, A.F., Avramopoulos, D., Chakravarti, A., Hoover-Fong, J., Mathews, D., Witmer, P.D., Ling, H., Hetrick, K., Watkins, L., Patterson, K.E.,

- Reinier, F., Blue, E., Muzny, D., Kircher, M., Bilguvar, K., Lopez-Giraldez, F., Sutton, V.R., Tabor, H.K., Leal, S.M., Gunel, M., Mane, S., Gibbs, R.A., Boerwinkle, E., Hamosh, A., Shendure, J., Lupski, J.R., Lifton, R.P., Valle, D., Nickerson, D.A., Centers for Mendelian, G., Bamshad, M.J., 2015. The genetic basis of Mendelian phenotypes: discoveries, challenges, and opportunities. *Am. J. Hum. Genet.* 97 (2), 199–215.
- Clayton, N.P., Nelson, C.A., Weeden, T., Taylor, K.M., Moreland, R.J., Scheule, R.K., Phillips, L., Leger, A.J., Cheng, S.H., Wentworth, B.M., 2014. Antisense oligonucleotide-mediated suppression of muscle glycogen synthase 1 synthesis as an approach for substrate reduction therapy of Pompe disease. *Mol. Ther. Nucleic Acids* 3, e206.
- Consugar, M.B., Navarro-Gomez, D., Place, E.M., Bujakowska, K.M., Sousa, M.E., Fonseca-Kelly, Z.D., Taub, D.G., Janessian, M., Wang, D.Y., Au, E.D., Sims, K.B., Sweetser, D.A., Fulton, A.B., Liu, Q., Wiggs, J.L., Gai, X., Pierce, E.A., 2015. Panel-based genetic diagnostic testing for inherited eye diseases is highly accurate and reproducible, and more sensitive for variant detection, than exome sequencing. *Genet. Med.* 17 (4), 253–261.
- Daxinger, L., Tapscott, S.J., van der Maarel, S.M., 2015. Genetic and epigenetic contributors to FSHD. *Curr. Opin. Genet. Dev.* 33, 56–61.
- de Fougères, A., Vornlocher, H.P., Maraganore, J., Lieberman, J., 2007. Interfering with disease: a progress report on siRNA-based therapeutics. *Nat. Rev. Drug Discov.* 6 (6), 443–453.
- Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M., Beroud, C., 2009. Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* 37 (9), e67.
- Di Giacomo, D., Gaildrat, P., Abuli, A., Abdat, J., Frebourg, T., Tosi, M., Martins, A., 2013. Functional analysis of a large set of BRCA2 exon 7 variants highlights the predictive value of hexamer scores in detecting alterations of exonic splicing regulatory elements. *Hum. Mutat.* 34 (11), 1547–1557.
- Divina, P., Kvitkovicova, A., Buratti, E., Vorechovsky, I., 2009. Ab initio prediction of mutation-induced cryptic splice-site activation and exon skipping. *Eur. J. Hum. Genet.* 17 (6), 759–765.
- Dogan, R.I., Getoor, L., Wilbur, W.J., Mount, S.M., 2007. SplicePort—an interactive splice-site analysis tool. *Nucleic Acids Res.* 35 (Web server issue), W285–91.
- Doma, M.K., Parker, R., 2006. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. *Nature* 440 (7083), 561–564.
- Dominski, Z., Kole, R., 1993. Restoration of correct splicing in thalassemic pre-mRNA by antisense oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 90 (18), 8673–8677.
- Dreyfuss, G., Matunis, M.J., Pinol-Roma, S., Burd, C.G., 1993. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* 62, 289–321.
- Du, C., Shen, Z., Zang, R., Xie, H., Li, H., Chen, P., Hang, B., Xu, X., Tang, W., Xia, Y., 2016. Negative feedback circuitry between MIR143HG and RBM24 in Hirschsprung disease. *Biochim. Biophys. Acta* 1862 (11), 2127–2136.
- Duell, P.B., Santos, R.D., Kirwan, B.A., Witztum, J.L., Tsimikas, S., Kastelein, J.J., 2016. Long-term mipomersen treatment is associated with a reduction in cardiovascular events in patients with familial hypercholesterolemia. *J. Clin. Lipidol.* 10 (4), 1011–1021.
- Duff, M.O., Olson, S., Wei, X., Garrett, S.C., Osman, A., Bolisetty, M., Plocik, A., Celniker, S.E., Graveley, B.R., 2015. Genome-wide identification of zero nucleotide recursive splicing in *Drosophila*. *Nature* 521 (7552), 376–379.
- d'Ydewalle, C., Ramos, D.M., Pyles, N.J., Ng, S.Y., Gorz, M., Pilato, C.M., Ling, K., Kong, L., Ward, A.J., Rubin, L.L., Rigo, F., Bennett, C.F., Sumner, C.J., 2017. The antisense transcript SMN-AS1 regulates SMN expression and is a novel therapeutic target for spinal muscular atrophy. *Neuron* 93 (1), 66–79.

- Ederly, P., Marcaillou, C., Sahbatou, M., Labalme, A., Chastang, J., Touraine, R., Tubacher, E., Senni, F., Bober, M.B., Nampoothiri, S., Jouk, P.S., Steichen, E., Berland, S., Toutain, A., Wise, C.A., Sanlaville, D., Rousseau, F., Clerget-Darpoux, F., Leutenegger, A.L., 2011. Association of TALS developmental disorder with defect in minor splicing component U4atac snRNA. *Science* 332 (6026), 240–243.
- Elbashir, S.M., Lendeckel, W., Tuschl, T., 2001. RNA interference is mediated by 21- and 22-nucleotide RNAs. *Genes Dev.* 15 (2), 188–200.
- Ellis, J.D., Barrios-Rodiles, M., Colak, R., Irimia, M., Kim, T., Calarco, J.A., Wang, X., Pan, Q., O'Hanlon, D., Kim, P.M., Wrana, J.L., Blencowe, B.J., 2012. Tissue-specific alternative splicing remodels protein–protein interaction networks. *Mol. Cell* 46 (6), 884–892.
- Elsaid, M.F., Chalhoub, N., Ben-Omran, T., Kumar, P., Kamel, H., Ibrahim, K., Mohamoud, Y., Al-Dous, E., Al-Azwani, I., Malek, J.A., Suhre, K., Ross, M.E., Aleem, A.A., 2017. Mutation in noncoding RNA RNU12 causes early onset cerebellar ataxia. *Ann. Neurol.* 81 (1), 68–78.
- Engstrom, P.G., Steijger, T., Sipos, B., Grant, G.R., Kahles, A., Ratsch, G., Goldman, N., Hubbard, T.J., Harrow, J., Guigo, R., Bertone, P., Consortium, R., 2013. Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat. Methods* 10 (12), 1185–1191.
- Faber, K., Glatting, K.H., Mueller, P.J., Risch, A., Hotz-Wagenblatt, A., 2011. Genome-wide prediction of splice-modifying SNPs in human genes using a new analysis pipeline called AASites. *BMC Bioinf.* 12 (Suppl. 4), S2.
- Fairbrother, W.G., Yeh, R.F., Sharp, P.A., Burge, C.B., 2002. Predictive identification of exonic splicing enhancers in human genes. *Science* 297 (5583), 1007–1013.
- Farooqi, A.A., Rehman, Z.U., Muntane, J., 2014. Antisense therapeutics in oncology: current status. *Onco Targets Ther.* 7, 2035–2042.
- Fu, X.D., Mayeda, A., Maniatis, T., Krainer, A.R., 1992. General splicing factors SF2 and SC35 have equivalent activities in vitro, and both affect alternative 5' and 3' splice site selection. *Proc. Natl. Acad. Sci. U.S.A.* 89 (23), 11224–11228.
- Gabut, M., Mine, M., Marsac, C., Brivet, M., Tazi, J., Soret, J., 2005. The SR protein SC35 is responsible for aberrant splicing of the E1alpha pyruvate dehydrogenase mRNA in a case of mental retardation with lactic acidosis. *Mol. Cell. Biol.* 25 (8), 3286–3294.
- Gao, Y., Tatavarty, V., Korza, G., Levin, M.K., Carson, J.H., 2008. Multiplexed dendritic targeting of alpha calcium calmodulin-dependent protein kinase II, neurogranin, and activity-regulated cytoskeleton-associated protein RNAs by the A2 pathway. *Mol. Biol. Cell* 19 (5), 2311–2327.
- Gao, C., Ren, S., Lee, J.H., Qiu, J., Chapski, D.J., Rau, C.D., Zhou, Y., Abdellatif, M., Nakano, A., Vondriska, T.M., Xiao, X., Fu, X.D., Chen, J.N., Wang, Y., 2016. RBFOX1-mediated RNA splicing regulates cardiac hypertrophy and heart failure. *J. Clin. Invest.* 126 (1), 195–206.
- Garanto, A., Chung, D.C., Duijkers, L., Corral-Serrano, J.C., Messchaert, M., Xiao, R., Bennett, J., Vandenberghe, L.H., Collin, R.W., 2016. In vitro and in vivo rescue of aberrant splicing in CEP290-associated LCA by antisense oligonucleotide delivery. *Hum. Mol. Genet.* 25 (12), 2552–2563.
- Gazzoli, I., Pulyakhina, I., Verwey, N.E., Ariyurek, Y., Laros, J.F., t Hoen, P.A., Aartsma-Rus, A., 2016. Non-sequential and multi-step splicing of the dystrophin transcript. *RNA Biol.* 13 (3), 290–305.
- Geary, R.S., Norris, D., Yu, R., Bennett, C.F., 2015. Pharmacokinetics, biodistribution and cell uptake of antisense oligonucleotides. *Adv. Drug Deliv. Rev.* 87, 46–51.
- Gehman, L.T., Meera, P., Stoilov, P., Shiue, L., O'Brien, J.E., Meisler, M.H., Ares Jr., M., Otis, T.S., Black, D.L., 2012. The splicing regulator Rbfox2 is required for both cerebellar development and mature motor function. *Genes Dev.* 26 (5), 445–460.

- Gelfman, S., Ast, G., 2013. When epigenetics meets alternative splicing: the roles of DNA methylation and GC architecture. *Epigenomics* 5 (4), 351–353.
- Gerstberger, S., Hafner, M., Tuschl, T., 2014. A census of human RNA-binding proteins. *Nat. Rev. Genet.* 15 (12), 829–845.
- Geuens, T., Bouhy, D., Timmerman, V., 2016. The hnRNP family: insights into their role in health and disease. *Hum. Genet.* 135 (8), 851–867.
- Giorgi, G., Casarin, A., Trevisson, E., Dona, M., Cassina, M., Graziano, C., Picci, L., Clementi, M., Salviati, L., 2015. Validation of CFTR intronic variants identified during cystic fibrosis population screening by a minigene splicing assay. *Clin. Chem. Lab. Med.* 53 (11), 1719–1723.
- Giudice, J., Xia, Z., Wang, E.T., Scavuzzo, M.A., Ward, A.J., Kalsotra, A., Wang, W., Wehrens, X.H., Burge, C.B., Li, W., Cooper, T.A., 2014. Alternative splicing regulates vesicular trafficking genes in cardiomyocytes during postnatal heart development. *Nat. Commun.* 5, 3603.
- Giudice, J., Loehr, J.A., Rodney, G.G., Cooper, T.A., 2016. Alternative splicing of four trafficking genes regulates myofiber structure and skeletal muscle physiology. *Cell Rep.* 17 (8), 1923–1933.
- Gokben, S., Onay, H., Yilmaz, S., Atik, T., Serdaroglu, G., Tekin, H., Ozkinay, F., 2016. Targeted next generation sequencing: the diagnostic value in early-onset epileptic encephalopathy. *Acta Neurol. Belg.* 117, 131–138.
- Gonzalez-Paredes, F.J., Ramos-Trujillo, E., Claverie-Martin, F., 2016. Three exonic mutations in polycystic kidney disease-2 gene (PKD2) alter splicing of its pre-mRNA in a minigene system. *Gene* 578 (1), 117–123.
- Goodwin, S., McPherson, J.D., McCombie, W.R., 2016. Coming of age: ten years of next-generation sequencing technologies. *Nat. Rev. Genet.* 17 (6), 333–351.
- Gorman, L., Suter, D., Emerick, V., Schumperli, D., Kole, R., 1998. Stable alteration of pre-mRNA splicing patterns by modified U7 small nuclear RNAs. *Proc. Natl. Acad. Sci. U.S.A.* 95 (9), 4929–4934.
- Gorman, L., Mercatante, D.R., Kole, R., 2000. Restoration of correct splicing of thalassemic beta-globin pre-mRNA by modified U1 snRNAs. *J. Biol. Chem.* 275 (46), 35914–35919.
- Goyenvall, A., Griffith, G., Babbs, A., El Andaloussi, S., Ezzat, K., Avril, A., Dugovic, B., Chaussonot, R., Ferry, A., Voit, T., Amthor, H., Buhr, C., Schurch, S., Wood, M.J., Davies, K.E., Vaillend, C., Leumann, C., Garcia, L., 2015. Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers. *Nat. Med.* 21 (3), 270–275.
- Guo, W., Schafer, S., Greaser, M.L., Radke, M.H., Liss, M., Govindarajan, T., Maatz, H., Schulz, H., Li, S., Parrish, A.M., Dauksaite, V., Vakeel, P., Klaassen, S., Gerull, B., Thierfelder, L., Regitz-Zagrosek, V., Hacker, T.A., Saupe, K.W., Dec, G.W., Ellinor, P.T., MacRae, C.A., Spallek, B., Fischer, R., Perrot, A., Ozcelik, C., Saar, K., Hubner, N., Gotthardt, M., 2012. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat. Med.* 18 (5), 766–773.
- Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano Jr., M., Jungkamp, A.C., Munschauer, M., Ulrich, A., Wardle, G.S., Dewell, S., Zavolan, M., Tuschl, T., 2010. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141 (1), 129–141.
- Hamada, N., Ito, H., Nishijo, T., Iwamoto, I., Morishita, R., Tabata, H., Momiyama, T., Nagata, K., 2016. Essential role of the nuclear isoform of RBFOX1, a candidate gene for autism spectrum disorders, in the brain development. *Sci. Rep.* 6, 30805.
- Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., Wang, E., Trcka, D., Thompson, T., O'Hanlon, D., Slobodeniuc, V., Barbosa-Morais, N.L., Burge, C.B., Moffat, J.,

- Frey, B.J., Nagy, A., Ellis, J., Wrana, J.L., Blencowe, B.J., 2013. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. *Nature* 498 (7453), 241–245.
- Han, G., Gu, B., Cao, L., Gao, X., Wang, Q., Seow, Y., Zhang, N., Wood, M.J., Yin, H., 2016. Hexose enhances oligonucleotide delivery and exon skipping in dystrophin-deficient mdx mice. *Nat. Commun.* 7, 10981.
- Hatton, A.R., Subramaniam, V., Lopez, A.J., 1998. Generation of alternative Ultrabithorax isoforms and stepwise removal of a large intron by resplicing at exon–exon junctions. *Mol. Cell* 2 (6), 787–796.
- He, X., Liu, F., Yan, J., Zhang, Y., Yan, J., Shang, H., Dou, Q., Zhao, Q., Song, Y., 2015. Trans-splicing repair of mutant p53 suppresses the growth of hepatocellular carcinoma cells in vitro and in vivo. *Sci. Rep.* 5, 8705.
- Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P., Brunak, S., 1996. Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.* 24 (17), 3439–3452.
- Hinze, F., Dieterich, C., Radke, M.H., Granzier, H., Gotthardt, M., 2016. Reducing RBM20 activity improves diastolic dysfunction and cardiac atrophy. *J. Mol. Med. (Berl.)* 94 (12), 1349–1358.
- Ho, T.H., Kapur, P., Joseph, R.W., Serie, D.J., Eckel-Passow, J.E., Tong, P., Wang, J., Castle, E.P., Stanton, M.L., Cheville, J.C., Jonasch, E., Brugarolas, J., Parker, A.S., 2016. Loss of histone H3 lysine 36 trimethylation is associated with an increased risk of renal cell carcinoma-specific death. *Mod. Pathol.* 29 (1), 34–42.
- Hollander, D., Naftelberg, S., Lev-Maor, G., Kornblihtt, A.R., Ast, G., 2016. How are short exons flanked by long introns defined and committed to splicing? *Trends Genet.* 32 (10), 596–606.
- Hong, D.S., Kurzrock, R., Naing, A., Wheeler, J.J., Falchook, G.S., Schiffman, J.S., Faulkner, N., Pilat, M.J., O'Brien, J., LoRusso, P., 2014. A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) spliceosome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest. New Drugs* 32 (3), 436–444.
- Hoskins, A.A., Moore, M.J., 2012. The spliceosome: a flexible, reversible macromolecular machine. *Trends Biochem. Sci.* 37 (5), 179–188.
- Howard, J.M., Sanford, J.R., 2015. The RNAissance family: SR proteins as multifaceted regulators of gene expression. *Wiley Interdiscip. Rev. RNA* 6 (1), 93–110.
- Hua, W.F., Zhong, Q., Xia, T.L., Chen, Q., Zhang, M.Y., Zhou, A.J., Tu, Z.W., Qu, C., Li, M.Z., Xia, Y.F., Wang, H.Y., Xie, D., Claret, F.X., Song, E.W., Zeng, M.S., 2016. RBM24 suppresses cancer progression by upregulating miR-25 to target MALAT1 in nasopharyngeal carcinoma. *Cell Death Dis.* 7 (9), e2352.
- Igreja, S., Clarke, L.A., Botelho, H.M., Marques, L., Amaral, M.D., 2016. Correction of a cystic fibrosis splicing mutation by antisense oligonucleotides. *Hum. Mutat.* 37 (2), 209–215.
- Imbert, M., Dias-Florencio, G., Goyenvall, A., 2017. Viral vector-mediated antisense therapy for genetic diseases. *Genes (Basel)* 8 (2), 51.
- Irimia, M., Weatheritt, R.J., Ellis, J.D., Parikshak, N.N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallieres, M., Tapial, J., Raj, B., O'Hanlon, D., Barrios-Rodiles, M., Sternberg, M.J., Cordes, S.P., Roth, F.P., Wrana, J.L., Geschwind, D.H., Blencowe, B.J., 2014. A highly conserved program of neuronal microexons is misregulated in autistic brains. *Cell* 159 (7), 1511–1523.
- Iyombe-Engembe, J.P., Ouellet, D.L., Barbeau, X., Rousseau, J., Chapdelaine, P., Lague, P., Tremblay, J.P., 2016. Efficient restoration of the dystrophin gene reading frame and protein structure in DMD myoblasts using the CinDel method. *Mol. Ther. Nucleic Acids* 5, e283.
- Izumikawa, K., Yoshikawa, H., Ishikawa, H., Nobe, Y., Yamauchi, Y., Philipsen, S., Simpson, R.J., Isobe, T., Takahashi, N., 2016. Chtop (Chromatin target of Prmt1)

- auto-regulates its expression level via intron retention and nonsense-mediated decay of its own mRNA. *Nucleic Acids Res.* 44 (20), 9847–9859.
- Jackson, A.L., Linsley, P.S., 2010. Recognizing and avoiding siRNA off-target effects for target identification and therapeutic application. *Nat. Rev. Drug Discov.* 9 (1), 57–67.
- Jangi, M., Sharp, P.A., 2014. Building robust transcriptomes with master splicing factors. *Cell* 159 (3), 487–498.
- Jangi, M., Boutz, P.L., Paul, P., Sharp, P.A., 2014. Rbfox2 controls autoregulation in RNA-binding protein networks. *Genes Dev.* 28 (6), 637–651.
- Jian, X., Boerwinkle, E., Liu, X., 2014. In silico tools for splicing defect prediction: a survey from the viewpoint of end users. *Genet. Med.* 16 (7), 497–503.
- Jiang, Y., Zhang, M., Qian, Y., Xu, E., Zhang, J., Chen, X., 2014. Rbm24, an RNA-binding protein and a target of p53, regulates p21 expression via mRNA stability. *J. Biol. Chem.* 289 (6), 3164–3175.
- Jin, Y., Suzuki, H., Maegawa, S., Endo, H., Sugano, S., Hashimoto, K., Yasuda, K., Inoue, K., 2003. A vertebrate RNA-binding protein fox-1 regulates tissue-specific splicing via the pentanucleotide GCAUG. *EMBO J.* 22 (4), 905–912.
- Jin, D., Hidaka, K., Shirai, M., Morisaki, T., 2010. RNA-binding motif protein 24 regulates myogenin expression and promotes myogenic differentiation. *Genes Cells* 15 (11), 1158–1167.
- Jonkers, I., Kwak, H., Lis, J.T., 2014. Genome-wide dynamics of pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *eLife* 3, e02407.
- Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., Tani, T., Horinouchi, S., Yoshida, M., 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* 3 (9), 576–583.
- Karlsson, K., Linnarsson, S., 2017. Single-cell mRNA isoform diversity in the mouse brain. *BMC Genomics* 18 (1), 126.
- Kashima, T., Rao, N., David, C.J., Manley, J.L., 2007. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. *Hum. Mol. Genet.* 16 (24), 3149–3159.
- Kastelein, J.J., Wedel, M.K., Baker, B.F., Su, J., Bradley, J.D., Yu, R.Z., Chuang, E., Graham, M.J., Crooke, R.M., 2006. Potent reduction of apolipoprotein B and low-density lipoprotein cholesterol by short-term administration of an antisense inhibitor of apolipoprotein B. *Circulation* 114 (16), 1729–1735.
- Kelly, S., Georgomanolis, T., Zirkel, A., Diermeier, S., O'Reilly, D., Murphy, S., Langst, G., Cook, P.R., Papantonis, A., 2015. Splicing of many human genes involves sites embedded within introns. *Nucleic Acids Res.* 43 (9), 4721–4732.
- Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., Shendure, J., 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 46 (3), 310–315.
- Knecht, C., Mort, M., Junge, O., Cooper, D.N., Krawczak, M., Caliebe, A., 2017. IMHOTEP—a composite score integrating popular tools for predicting the functional consequences of non-synonymous sequence variants. *Nucleic Acids Res.* 45 (3), e13.
- Kolasinska-Zwierz, P., Down, T., Latorre, I., Liu, T., Liu, X.S., Ahringer, J., 2009. Differential chromatin marking of introns and expressed exons by H3K36me3. *Nat. Genet.* 41 (3), 376–381.
- Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., Ule, J., 2010. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* 17 (7), 909–915.
- Kotterman, M.A., Schaffer, D.V., 2014. Engineering adeno-associated viruses for clinical gene therapy. *Nat. Rev. Genet.* 15 (7), 445–451.
- Krainer, A.R., Conway, G.C., Kozak, D., 1990. The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* 62 (1), 35–42.

- Krawczak, M., Reiss, J., Cooper, D.N., 1992. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Hum. Genet.* 90 (1–2), 41–54.
- Krecic, A.M., Swanson, M.S., 1999. hnRNP complexes: composition, structure, and function. *Curr. Opin. Cell Biol.* 11 (3), 363–371.
- Krieg, A.M., 2011. Is RNAi dead? *Mol. Ther.* 19 (6), 1001–1002.
- Kruger, S., Battke, F., Sprecher, A., Munz, M., Synofzik, M., Schols, L., Gasser, T., Grehl, T., Prudlo, J., Biskup, S., 2016. Rare variants in neurodegeneration associated genes revealed by targeted panel sequencing in a German ALS cohort. *Front. Mol. Neurosci.* 9, 92.
- Kwak, H., Lis, J.T., 2013. Control of transcriptional elongation. *Annu. Rev. Genet.* 47, 483–508.
- Laustriat, D., Gide, J., Barrault, L., Chautard, E., Benoit, C., Auboeuf, D., Boland, A., Battail, C., Artiguenave, F., Deleuze, J.F., Benit, P., Rustin, P., Franc, S., Charpentier, G., Furling, D., Bassez, G., Nissan, X., Martinat, C., Peschanski, M., Baghdoyan, S., 2015. In vitro and in vivo modulation of alternative splicing by the biguanide metformin. *Mol. Ther. Nucleic Acids* 4, e262.
- Lee, S.C., Abdel-Wahab, O., 2016. Therapeutic targeting of splicing in cancer. *Nat. Med.* 22 (9), 976–986.
- Lee, Y., Rio, D.C., 2015. Mechanisms and regulation of alternative pre-mRNA splicing. *Annu. Rev. Biochem.* 84, 291–323.
- Lei, Q., Li, C., Zuo, Z., Huang, C., Cheng, H., Zhou, R., 2016. Evolutionary insights into RNA trans-splicing in vertebrates. *Genome Biol. Evol.* 8 (3), 562–577.
- Lentz, J.J., Jodelka, F.M., Hinrich, A.J., McCaffrey, K.E., Farris, H.E., Spalitta, M.J., Bazan, N.G., Duelli, D.M., Rigo, F., Hastings, M.L., 2013. Rescue of hearing and vestibular function by antisense oligonucleotides in a mouse model of human deafness. *Nat. Med.* 19 (3), 345–350.
- Li, S., Guo, W., Dewey, C.N., Greaser, M.L., 2013. Rbm20 regulates titin alternative splicing as a splicing repressor. *Nucleic Acids Res.* 41 (4), 2659–2672.
- Li, Y.I., Sanchez-Pulido, L., Haerty, W., Ponting, C.P., 2015. RBFOX and PTBP1 proteins regulate the alternative splicing of micro-exons in human brain transcripts. *Genome Res.* 25 (1), 1–13.
- Li, Y.I., van de Geijn, B., Raj, A., Knowles, D.A., Petti, A.A., Golan, D., Gilad, Y., Pritchard, J.K., 2016. RNA splicing is a primary link between genetic variation and disease. *Science* 352 (6285), 600–604.
- Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., Darnell, J.C., Darnell, R.B., 2008. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456 (7221), 464–469.
- Lieber, M.R., 2010. The mechanism of double-strand DNA break repair by the non-homologous DNA end-joining pathway. *Annu. Rev. Biochem.* 79, 181–211.
- Lim, K.H., Fairbrother, W.G., 2012. Spliceman—a computational web server that predicts sequence variations in pre-mRNA splicing. *Bioinformatics* 28 (7), 1031–1032.
- Linder, B., Dill, H., Hirmer, A., Brocher, J., Lee, G.P., Mathavan, S., Bolz, H.J., Winkler, C., Lagerbauer, B., Fischer, U., 2011. Systemic splicing factor deficiency causes tissue-specific defects: a zebrafish model for retinitis pigmentosa. *Hum. Mol. Genet.* 20 (2), 368–377.
- Liu, H., Jin, T., Guan, J., Zhou, S., 2014. Histone modifications involved in cassette exon inclusions: a quantitative and interpretable analysis. *BMC Genomics* 15, 1148.
- Llorian, M., Schwartz, S., Clark, T.A., Hollander, D., Tan, L.Y., Spellman, R., Gordon, A., Schweitzer, A.C., de la Grange, P., Ast, G., Smith, C.W., 2010. Position-dependent

- alternative splicing activity revealed by global profiling of alternative splicing events regulated by PTB. *Nat. Struct. Mol. Biol.* 17 (9), 1114–1123.
- Long, J.C., Caceres, J.F., 2009. The SR protein family of splicing factors: master regulators of gene expression. *Biochem. J.* 417 (1), 15–27.
- Long, C., Amoasii, L., Mireault, A.A., McAnally, J.R., Li, H., Sanchez-Ortiz, E., Bhattacharyya, S., Shelton, J.M., Bassel-Duby, R., Olson, E.N., 2016. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* 351 (6271), 400–403.
- Lopez-Bigas, N., Audit, B., Ouzounis, C., Parra, G., Guigo, R., 2005. Are splicing mutations the most frequent cause of hereditary disease? *FEBS Lett.* 579 (9), 1900–1903.
- Lou, C.H., Dumdie, J., Goetz, A., Shum, E.Y., Brafman, D., Liao, X., Mora-Castilla, S., Ramaiah, M., Cook-Andersen, H., Laurent, L., Wilkinson, M.F., 2016. Nonsense-mediated RNA decay influences human embryonic stem cell fate. *Stem Cell Rep.* 6 (6), 844–857.
- Luco, R.F., Pan, Q., Tominaga, K., Blencowe, B.J., Pereira-Smith, O.M., Misteli, T., 2010. Regulation of alternative splicing by histone modifications. *Science* 327 (5968), 996–1000.
- Maggio, I., Goncalves, M.A., 2015. Genome editing at the crossroads of delivery, specificity, and fidelity. *Trends Biotechnol.* 33 (5), 280–291.
- Maggio, I., Stefanucci, L., Janssen, J.M., Liu, J., Chen, X., Mouly, V., Goncalves, M.A., 2016. Selection-free gene repair after adenoviral vector transduction of designer nucleases: rescue of dystrophin synthesis in DMD muscle cell populations. *Nucleic Acids Res.* 44 (3), 1449–1470.
- Manley, J.L., Krainer, A.R., 2010. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). *Genes Dev.* 24 (11), 1073–1074.
- Mansfield, S.G., Kole, J., Puttaraju, M., Yang, C.C., Garcia-Blanco, M.A., Cohn, J.A., Mitchell, L.G., 2000. Repair of CFTR mRNA by spliceosome-mediated RNA trans-splicing. *Gene Ther.* 7 (22), 1885–1895.
- Martelli, P.L., D'Antonio, M., Bonizzoni, P., Castrignano, T., D'Erchia, A.M., D'Onorio De Meo, P., Fariselli, P., Finelli, M., Licciulli, F., Mangiulli, M., Mignone, F., Pavesi, G., Picardi, E., Rizzi, R., Rossi, I., Valletti, A., Zauli, A., Zambelli, F., Casadio, R., Pesole, G., 2011. ASPicDB: a database of annotated transcript and protein variants generated by alternative splicing. *Nucleic Acids Res.* 39 (Database issue), D80–5.
- Martin, C.L., Duvall, J.A., Ilkin, Y., Simon, J.S., Arreaza, M.G., Wilkes, K., Alvarez-Retuerto, A., Whichello, A., Powell, C.M., Rao, K., Cook, E., Geschwind, D.H., 2007. Cytogenetic and molecular characterization of A2BP1/FOX1 as a candidate gene for autism. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 144B (7), 869–876.
- Martinez-Contreras, R., Fiset, J.F., Nasim, F.U., Madden, R., Cordeau, M., Chabot, B., 2006. Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol.* 4 (2), e21.
- Methawasin, M., Hutchinson, K.R., Lee, E.J., Smith 3rd, J.E., Saripalli, C., Hidalgo, C.G., Ottenheijm, C.A., Granzier, H., 2014. Experimentally increasing titin compliance in a novel mouse model attenuates the Frank-Starling mechanism but has a beneficial effect on diastole. *Circulation* 129 (19), 1924–1936.
- Miller, J.N., Pearce, D.A., 2014. Nonsense-mediated decay in genetic disease: friend or foe? *Mutat. Res. Rev. Mutat. Res.* 762, 52–64.
- Miyamoto, S., Hidaka, K., Jin, D., Morisaki, T., 2009. RNA-binding proteins Rbm38 and Rbm24 regulate myogenic differentiation via p21-dependent and -independent regulatory pathways. *Genes Cells* 14 (11), 1241–1252.
- Moore, M.J., Sharp, P.A., 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* 365 (6444), 364–368.

- Moraes, F., Goes, A., 2016. A decade of human genome project conclusion: scientific diffusion about our genome knowledge. *Biochem. Mol. Biol. Educ.* 44 (3), 215–223.
- Motta-Mena, L.B., Heyd, F., Lynch, K.W., 2010. Context-dependent regulatory mechanism of the splicing factor hnRNP L. *Mol. Cell* 37 (2), 223–234.
- Mucaki, E.J., Shirley, B.C., Rogan, P.K., 2013. Prediction of mutant mRNA splice isoforms by information theory-based exon definition. *Hum. Mutat.* 34 (4), 557–565.
- Mudvari, P., Movassagh, M., Kowsari, K., Seyfi, A., Kokkinaki, M., Edwards, N.J., Golestaneh, N., Horvath, A., 2015. SNPllice: variants that modulate intron retention from RNA-sequencing data. *Bioinformatics* 31 (8), 1191–1198.
- Mueller, W.F., Larsen, L.S., Garibaldi, A., Hatfield, G.W., Hertel, K.J., 2015. The silent sway of splicing by synonymous substitutions. *J. Biol. Chem.* 290 (46), 27700–27711.
- Naldini, L., 2015. Gene therapy returns to centre stage. *Nature* 526 (7573), 351–360.
- Naryshkin, N.A., Weetall, M., Dakka, A., Narasimhan, J., Zhao, X., Feng, Z., Ling, K.K., Karp, G.M., Qi, H., Woll, M.G., Chen, G., Zhang, N., Gabbeta, V., Vazirani, P., Bhattacharyya, A., Furia, B., Risher, N., Sheedy, J., Kong, R., Ma, J., Turpoff, A., Lee, C.S., Zhang, X., Moon, Y.C., Trifillis, P., Welch, E.M., Colacino, J.M., Babiak, J., Almstead, N.G., Peltz, S.W., Eng, L.A., Chen, K.S., Mull, J.L., Lynes, M.S., Rubin, L.L., Fontoura, P., Santarelli, L., Haehnke, D., McCarthy, K.D., Schmucki, R., Ebeling, M., Sivaramakrishnan, M., Ko, C.P., Paushkin, S.V., Ratni, H., Gerlach, I., Ghosh, A., Metzger, F., 2014. Motor neuron disease. SMN2 splicing modifiers improve motor function and longevity in mice with spinal muscular atrophy. *Science* 345 (6197), 688–693.
- Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., Yan, W.X., Asokan, A., Zhang, F., Duan, D., Gersbach, C.A., 2016. In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351 (6271), 403–407.
- Neveling, K., Mensenkamp, A.R., Derks, R., Kwint, M., Ouchene, H., Steehouwer, M., van Lier, B., Bosgoed, E., Rikken, A., Tychon, M., Zafeiropoulou, D., Castelein, S., Hehir-Kwa, J., Tjwan Thung, D., Hofste, T., Lelieveld, S.H., Bertens, S.M., Adan, I.B., Eijkelenboom, A., Tops, B.B., Yntema, H., Stokowy, T., Knappskog, P.M., Hoberg-Vetti, H., Steen, V.M., Boyle, E., Martin, B., Ligtenberg, M.J., Shendure, J., Nelen, M.R., Hoischen, A., 2017. BRCA testing by single-molecule molecular inversion probes. *Clin. Chem.* 63 (2), 503–512.
- Niks, E.H., Aartsma-Rus, A., 2017. Exon skipping: a first in class strategy for Duchenne muscular dystrophy. *Expert Opin. Biol. Ther.* 17 (2), 225–236.
- Nilsen, T.W., Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463 (7280), 457–463.
- Nishiguchi, K.M., Tearle, R.G., Liu, Y.P., Oh, E.C., Miyake, N., Benaglio, P., Harper, S., Koskiniemi-Kuendig, H., Venturini, G., Sharon, D., Koenekoop, R.K., Nakamura, M., Kondo, M., Ueno, S., Yasuma, T.R., Beckmann, J.S., Ikegawa, S., Matsumoto, N., Terasaki, H., Berson, E.L., Katsanis, N., Rivolta, C., 2013. Whole genome sequencing in patients with retinitis pigmentosa reveals pathogenic DNA structural changes and NEK2 as a new disease gene. *Proc. Natl. Acad. Sci. U.S.A.* 110 (40), 16139–16144.
- Nutter, C.A., Jaworski, E.A., Verma, S.K., Deshmukh, V., Wang, Q., Botvinnik, O.B., Lozano, M.J., Abass, I.J., Ijaz, T., Brasier, A.R., Garg, N.J., Wehrens, X.H., Yeo, G.W., Kuyumcu-Martinez, M.N., 2016. Dysregulation of RBFOX2 is an early event in cardiac pathogenesis of diabetes. *Cell Rep.* 15 (10), 2200–2213.
- Oh, H., Lee, E., Jang, H.N., Lee, J., Moon, H., Sheng, Z., Jun, Y., Loh, T.J., Cho, S., Zhou, J., Green, M.R., Zheng, X., Shen, H., 2013. hnRNP A1 contacts exon 5 to promote exon 6 inclusion of apoptotic Fas gene. *Apoptosis* 18 (7), 825–835.

- Ohe, K., Hagiwara, M., 2015. Modulation of alternative splicing with chemical compounds in new therapeutics for human diseases. *ACS Chem. Biol.* 10 (4), 914–924.
- Okunola, H.L., Krainer, A.R., 2009. Cooperative-binding and splicing-repressive properties of hnRNP A1. *Mol. Cell. Biol.* 29 (20), 5620–5631.
- Osman, E.Y., Washington 3rd, C.W., Kaifer, K.A., Mazzasette, C., Patitucci, T.N., Florea, K.M., Simon, M.E., Ko, C.P., Ebert, A.D., Lorson, C.L., 2016. Optimization of morpholino antisense oligonucleotides targeting the Intronic repressor element1 in spinal muscular atrophy. *Mol. Ther.* 24 (9), 1592–1601.
- Osorio, F.G., Navarro, C.L., Cadinanos, J., Lopez-Mejia, I.C., Quiros, P.M., Bartoli, C., Rivera, J., Tazi, J., Guzman, G., Varela, I., Depetris, D., de Carlos, F., Cobo, J., Andres, V., De Sandre-Giovannoli, A., Freije, J.M., Levy, N., Lopez-Otin, C., 2011. Splicing-directed therapy in a new mouse model of human accelerated aging. *Sci. Transl. Med.* 3 (106), 106ra107.
- Palacino, J., Swalley, S.E., Song, C., Cheung, A.K., Shu, L., Zhang, X., Van Hoosear, M., Shin, Y., Chin, D.N., Keller, C.G., Beibel, M., Renaud, N.A., Smith, T.M., Salcius, M., Shi, X., Hild, M., Servais, R., Jain, M., Deng, L., Bullock, C., McLellan, M., Schuierer, S., Murphy, L., Blommers, M.J., Blaustein, C., Berenshteyn, F., Lacoste, A., Thomas, J.R., Roma, G., Michaud, G.A., Tseng, B.S., Porter, J.A., Myer, V.E., Tallarico, J.A., Hamann, L.G., Curtis, D., Fishman, M.C., Dietrich, W.F., Dales, N.A., Sivasankaran, R., 2015. SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat. Chem. Biol.* 11 (7), 511–517.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., Blencowe, B.J., 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* 40 (12), 1413–1415.
- Pandit, S., Zhou, Y., Shiue, L., Coutinho-Mansfield, G., Li, H., Qiu, J., Huang, J., Yeo, G.W., Ares Jr., M., Fu, X.D., 2013. Genome-wide analysis reveals SR protein cooperation and competition in regulated splicing. *Mol. Cell* 50 (2), 223–235.
- Park, J.W., Graveley, B.R., 2007. Complex alternative splicing. *Adv. Exp. Med. Biol.* 623, 50–63.
- Park, S.J., Lee, H., Jo, D.S., Jo, Y.K., Shin, J.H., Kim, H.B., Seo, H.M., Rubinsztein, D.C., Koh, J.Y., Lee, E.K., Cho, D.H., 2015. Heterogeneous nuclear ribonucleoprotein A1 post-transcriptionally regulates Drp1 expression in neuroblastoma cells. *Biochim. Biophys. Acta* 1849 (12), 1423–1431.
- Park, S.A., Ahn, S.I., Gallo, J.M., 2016. Tau mis-splicing in the pathogenesis of neurodegenerative disorders. *BMB Rep.* 49 (8), 405–413.
- Pertea, M., Lin, X., Salzberg, S.L., 2001. GeneSplicer: a new computational method for splice site prediction. *Nucleic Acids Res.* 29 (5), 1185–1190.
- Petrilli, R., Eloy, J.O., Marchetti, J.M., Lopez, R.F., Lee, R.J., 2014. Targeted lipid nanoparticles for antisense oligonucleotide delivery. *Curr. Pharm. Biotechnol.* 15 (9), 847–855.
- Pont, A.R., Sadri, N., Hsiao, S.J., Smith, S., Schneider, R.J., 2012. mRNA decay factor AUF1 maintains normal aging, telomere maintenance, and suppression of senescence by activation of telomerase transcription. *Mol. Cell* 47 (1), 5–15.
- Pradeepa, M.M., Sutherland, H.G., Ule, J., Grimes, G.R., Bickmore, W.A., 2012. Psp1/Ledgf p52 binds methylated histone H3K36 and splicing factors and contributes to the regulation of alternative splicing. *PLoS Genet.* 8 (5), e1002717.
- Proost, D., Saenen, J., Vandeweyer, G., Rothier, A., Alaerts, M., Van Craenenbroeck, E.M., Van Crombruggen, J., Mortier, G., Wuyts, W., Vrints, C., Del Favero, J., Loeys, B., Van Laer, L., 2017. Targeted next-generation sequencing of 51 genes involved in primary electrical disease. *J. Mol. Diagn.* 19 (3), 445–459.
- Raj, B., Blencowe, B.J., 2015. Alternative splicing in the mammalian nervous system: recent insights into mechanisms and functional roles. *Neuron* 87 (1), 14–27.

- Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., Koonin, E.V., Sharp, P.A., Zhang, F., 2015. In vivo genome editing using *Staphylococcus aureus* Cas9. *Nature* 520 (7546), 186–191.
- Raponi, M., Kralovicova, J., Copson, E., Divina, P., Eccles, D., Johnson, P., Baralle, D., Vorechovsky, I., 2011. Prediction of single-nucleotide substitutions that result in exon skipping: identification of a splicing silencer in BRCA1 exon 6. *Hum. Mutat.* 32 (4), 436–444.
- Ratni, H., Karp, G.M., Weetall, M., Naryshkin, N.A., Paushkin, S.V., Chen, K.S., McCarthy, K.D., Qi, H., Turpoff, A., Woll, M.G., Zhang, X., Zhang, N., Yang, T., Dakka, A., Vazirani, P., Zhao, X., Pinard, E., Green, L., David-Pierson, P., Tuerck, D., Poirier, A., Muster, W., Kirchner, S., Mueller, L., Gerlach, I., Metzger, F., 2016. Specific correction of alternative survival motor neuron 2 splicing by small molecules: discovery of a potential novel medicine to treat spinal muscular atrophy. *J. Med. Chem.* 59 (13), 6086–6100.
- Reese, M.G., Eeckman, F.H., Kulp, D., Haussler, D., 1997. Improved splice site detection in *genie*. *J. Comput. Biol.* 4 (3), 311–323.
- Rhoads, A., Au, K.F., 2015. PacBio sequencing and its applications. *Genomics Proteomics Bioinformatics* 13 (5), 278–289.
- Rindt, H., Yen, P.F., Thebeau, C.N., Peterson, T.S., Weisman, G.A., Lorson, C.L., 2012. Replacement of huntingtin exon 1 by trans-splicing. *Cell. Mol. Life Sci.* 69 (24), 4191–4204.
- Rodor, J., FitzPatrick, D.R., Eyra, E., Caceres, J.F., 2017. The RNA-binding landscape of RBM10 and its role in alternative splicing regulation in models of mouse early development. *RNA Biol.* 14 (1), 45–57.
- Rogalska, M.E., Tajnik, M., Licastro, D., Bussani, E., Camparini, L., Mattioli, C., Pagani, F., 2016. Therapeutic activity of modified U1 core spliceosomal particles. *Nat. Commun.* 7, 11168.
- Rogozin, I.B., Milanesi, L., 1997. Analysis of donor splice sites in different eukaryotic organisms. *J. Mol. Evol.* 45 (1), 50–59.
- Ruzickova, S., Stanek, D., 2016. Mutations in spliceosomal proteins and retina degeneration. *RNA Biol.* 1–9.
- Salton, M., Misteli, T., 2016. Small molecule modulators of pre-mRNA splicing in cancer therapy. *Trends Mol. Med.* 22 (1), 28–37.
- Saul, M.J., Stein, S., Grez, M., Jakobsson, P.J., Steinhilber, D., Suess, B., 2016. UPF1 regulates myeloid cell functions and S100A9 expression by the hnRNP E2/miRNA-328 balance. *Sci. Rep.* 6, 31995.
- Sawyer, S.L., Hartley, T., Dymont, D.A., Beaulieu, C.L., Schwartzentruber, J., Smith, A., Bedford, H.M., Bernard, G., Bernier, F.P., Brais, B., Bulman, D.E., Warman Chardon, J., Chitayat, D., Deladoey, J., Fernandez, B.A., Frosk, P., Geraghty, M.T., Gerull, B., Gibson, W., Gow, R.M., Graham, G.E., Green, J.S., Heon, E., Horvath, G., Innes, A.M., Jabado, N., Kim, R.H., Koeneke, R.K., Khan, A., Lehmann, O.J., Mendoza-Londono, R., Michaud, J.L., Nikkel, S.M., Penney, L.S., Polychronakos, C., Richer, J., Rouleau, G.A., Samuels, M.E., Siu, V.M., Suchowersky, O., Tarnopolsky, M.A., Yoon, G., Zahir, F.R., Consortium, F.C., Care4Rare Canada, C., Majewski, J., Boycott, K.M., 2016. Utility of whole-exome sequencing for those near the end of the diagnostic odyssey: time to address gaps in care. *Clin. Genet.* 89 (3), 275–284.
- Schleit, J., Bailey, S.S., Tran, T., Chen, D., Stowers, S., Schwarze, U., Byers, P.H., 2015. Molecular outcome, prediction, and clinical consequences of splice variants in COL1A1, which encodes the pro α 1(I) chains of type I procollagen. *Hum. Mutat.* 36 (7), 728–739.

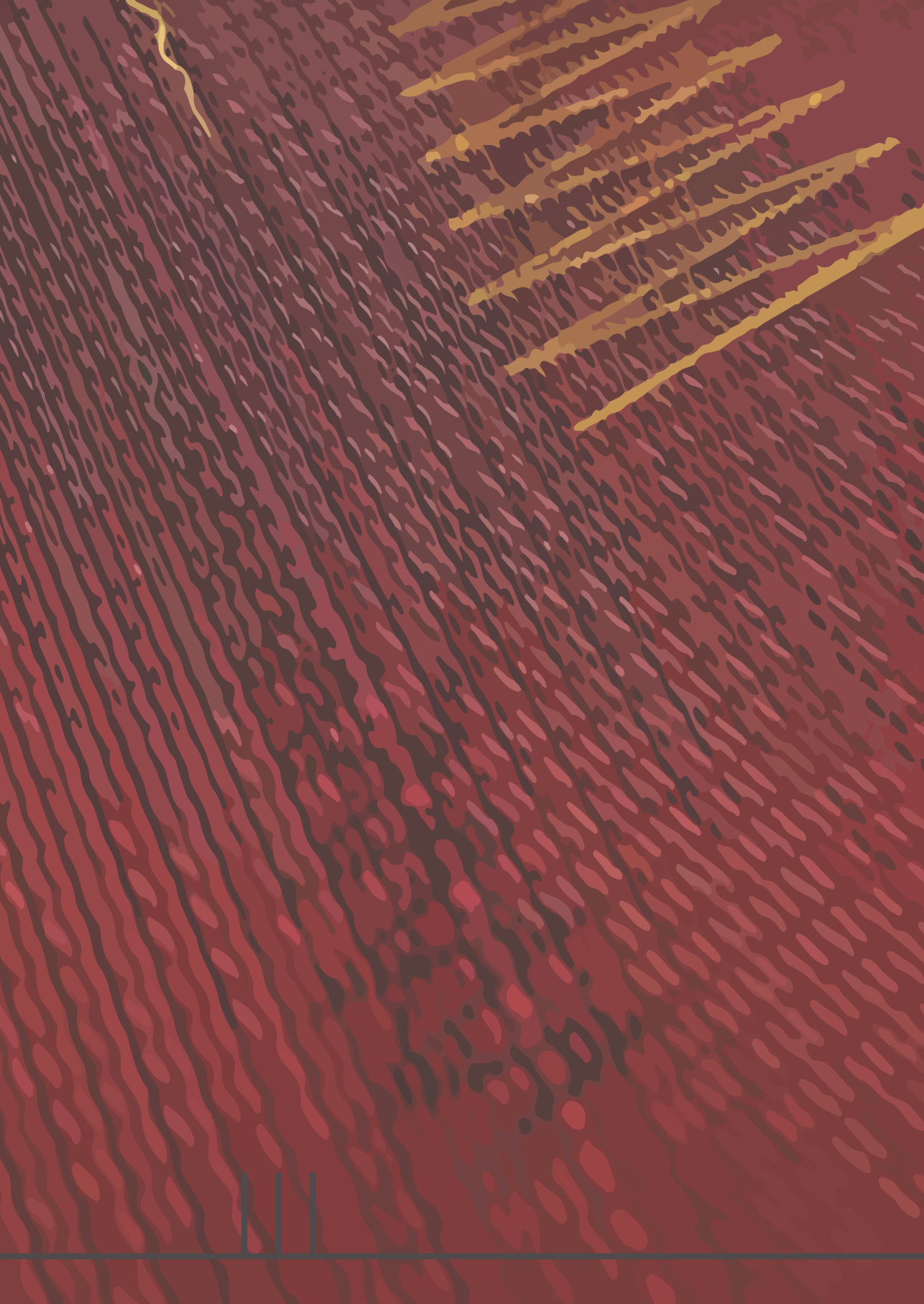
- Schultz, A.S., Preussner, M., Bunse, M., Karni, R., Heyd, F., 2016. Activation-dependent TRAF3 exon 8 alternative splicing is controlled by CELF2 and hnRNP C binding to an upstream intronic element. *Mol. Cell. Biol.* 37, e00488–16.
- Schumperli, D., Pillai, R.S., 2004. The special Sm core structure of the U7 snRNP: far-reaching significance of a small nuclear ribonucleoprotein. *Cell. Mol. Life Sci.* 61 (19–20), 2560–2570.
- Schwartz, S., Hall, E., Ast, G., 2009a. SROOGLE: webserver for integrative, user-friendly visualization of splicing signals. *Nucleic Acids Res.* 37 (Web server issue), W189–92.
- Schwartz, S., Meshorer, E., Ast, G., 2009b. Chromatin organization marks exon-intron structure. *Nat. Struct. Mol. Biol.* 16 (9), 990–995.
- Schwarz, J.M., Rodelsperger, C., Schuelke, M., Seelow, D., 2010. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* 7 (8), 575–576.
- Schwarz, J.M., Cooper, D.N., Schuelke, M., Seelow, D., 2014. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* 11 (4), 361–362.
- Sebat, J., Lakshmi, B., Malhotra, D., Troge, J., Lese-Martin, C., Walsh, T., Yamrom, B., Yoon, S., Krasnitz, A., Kendall, J., Leotta, A., Pai, D., Zhang, R., Lee, Y.H., Hicks, J., Spence, S.J., Lee, A.T., Puura, K., Lehtimäki, T., Ledbetter, D., Gregersen, P.K., Bregman, J., Sutcliffe, J.S., Jobanputra, V., Chung, W., Warburton, D., King, M.C., Skuse, D., Geschwind, D.H., Gilliam, T.C., Ye, K., Wigler, M., 2007. Strong association of de novo copy number mutations with autism. *Science* 316 (5823), 445–449.
- Shabanpoor, F., Hammond, S.M., Abendroth, F., Hazell, G., Wood, M.J., Gait, M.J., 2017. Identification of a peptide for systemic brain delivery of a morpholino oligonucleotide in mouse models of spinal muscular atrophy. *Nucleic Acid Ther.* 27, 130–143.
- Shan, J., Munro, T.P., Barbarese, E., Carson, J.H., Smith, R., 2003. A molecular mechanism for mRNA trafficking in neuronal dendrites. *J. Neurosci.* 23 (26), 8859–8866.
- Shapiro, M.B., Senapathy, P., 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res.* 15 (17), 7155–7174.
- Sharon, D., Tilgner, H., Grubert, F., Snyder, M., 2013. A single-molecule long-read survey of the human transcriptome. *Nat. Biotechnol.* 31 (11), 1009–1014.
- Shibata, A., Okuno, T., Rahman, M.A., Azuma, Y., Takeda, J., Masuda, A., Selcen, D., Engel, A.G., Ohno, K., 2016. IntSplice: prediction of the splicing consequences of intronic single-nucleotide variations in the human genome. *J. Hum. Genet.* 61 (7), 633–640.
- Shindo, Y., Nozaki, T., Saito, R., Tomita, M., 2013. Computational analysis of associations between alternative splicing and histone modifications. *FEBS Lett.* 587 (5), 516–521.
- Sibley, C.R., Emmett, W., Blazquez, L., Faro, A., Haberman, N., Briesse, M., Trabzuni, D., Ryten, M., Weale, M.E., Hardy, J., Modic, M., Curk, T., Wilson, S.W., Plagnol, V., Ule, J., 2015. Recursive splicing in long vertebrate genes. *Nature* 521 (7552), 371–375.
- Sierakowska, H., Sambade, M.J., Agrawal, S., Kole, R., 1996. Repair of thalassemic human beta-globin mRNA in mammalian cells by antisense oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 93 (23), 12840–12844.
- Simms, C.L., Thomas, E.N., Zaher, H.S., 2017. Ribosome-based quality control of mRNA and nascent peptides. *Wiley Interdiscip. Rev. RNA* 8 (1), e1366.
- Simon, J.M., Hacker, K.E., Singh, D., Brannon, A.R., Parker, J.S., Weiser, M., Ho, T.H., Kuan, P.F., Jonasch, E., Furey, T.S., Prins, J.F., Lieb, J.D., Rathmell, W.K., Davis, I.J., 2014. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome Res.* 24 (2), 241–250.
- Sims 3rd, R.J., Millhouse, S., Chen, C.F., Lewis, B.A., Erdjument-Bromage, H., Tempst, P., Manley, J.L., Reinberg, D., 2007. Recognition of trimethylated histone H3 lysine 4

- facilitates the recruitment of transcription postinitiation factors and pre-mRNA splicing. *Mol. Cell* 28 (4), 665–676.
- Singh, R.K., Cooper, T.A., 2012. Pre-mRNA splicing in disease and therapeutics. *Trends Mol. Med.* 18 (8), 472–482.
- Singh, N.K., Singh, N.N., Androphy, E.J., Singh, R.N., 2006. Splicing of a critical exon of human survival motor neuron is regulated by a unique silencer element located in the last intron. *Mol. Cell. Biol.* 26 (4), 1333–1346.
- Singh, N.N., Lee, B.M., DiDonato, C.J., Singh, R.N., 2015. Mechanistic principles of anti-sense targets for the treatment of spinal muscular atrophy. *Future Med. Chem.* 7 (13), 1793–1808.
- Sironi, M., Menozzi, G., Riva, L., Cagliani, R., Comi, G.P., Bresolin, N., Giorda, R., Pozzoli, U., 2004. Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic Acids Res.* 32 (5), 1783–1791.
- Soldati, D., Schumperli, D., 1988. Structural and functional characterization of mouse U7 small nuclear RNA active in 3' processing of histone pre-mRNA. *Mol. Cell. Biol.* 8 (4), 1518–1524.
- Soukari, O., Gaildrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frebourg, T., Tosi, M., Martins, A., 2016. Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. *PLoS Genet.* 12 (1), e1005756.
- Stanislaw, C., Xue, Y., Wilcox, W.R., 2016. Genetic evaluation and testing for hereditary forms of cancer in the era of next-generation sequencing. *Cancer Biol. Med.* 13 (1), 55–67.
- Stark, H., Dube, P., Luhrmann, R., Kastner, B., 2001. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature* 409 (6819), 539–542.
- Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D.N., Sanford, J.R., 2011. Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.* 21 (10), 1563–1571.
- Stoilov, P., Lin, C.H., Damoiseaux, R., Nikolic, J., Black, D.L., 2008. A high-throughput screening strategy identifies cardiotoxic steroids as alternative splicing modulators. *Proc. Natl. Acad. Sci. U.S.A.* 105 (32), 11218–11223.
- Supek, F., Minana, B., Valcarcel, J., Gabaldon, T., Lehner, B., 2014. Synonymous mutations frequently act as driver mutations in human cancers. *Cell* 156 (6), 1324–1335.
- Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., Kurita, M., Hishida, T., Li, M., Aizawa, E., Guo, S., Chen, S., Goebel, A., Soligalla, R.D., Qu, J., Jiang, T., Fu, X., Jafari, M., Esteban, C.R., Berggren, W.T., Lajara, J., Nunez-Delgado, E., Guillen, P., Campistol, J.M., Matsuzaki, F., Liu, G.H., Magistretti, P., Zhang, K., Callaway, E.M., Zhang, K., Belmonte, J.C., 2016. In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 540 (7631), 144–149.
- Tabebordbar, M., Zhu, K., Cheng, J.K., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., Cong, L., Zhang, F., Vandenbergh, L.H., Church, G.M., Wagers, A.J., 2016. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351 (6271), 407–411.
- Takata, A., Matsumoto, N., Kato, T., 2017. Genome-wide identification of splicing QTLs in the human brain and their enrichment among schizophrenia-associated loci. *Nat. Commun.* 8, 14519.
- Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G., Holmes, M.C., Gregory, P.D., Ando, D.G., Kalos, M., Collman, R.G., Binder-Scholl, G., Plesa, G., Hwang, W.T.,

- Levine, B.L., June, C.H., 2014. Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370 (10), 901–910.
- Tekin, D., Yan, D., Bademci, G., Feng, Y., Guo, S., Foster 2nd, J., Blanton, S., Tekin, M., Liu, X., 2016. A next-generation sequencing gene panel (MiamiOtoGenes) for comprehensive analysis of deafness genes. *Hear. Res.* 333, 179–184.
- Tilgner, H., Grubert, F., Sharon, D., Snyder, M.P., 2014. Defining a personal, allele-specific, and single-molecule long-read transcriptome. *Proc. Natl. Acad. Sci. U.S.A.* 111 (27), 9869–9874.
- Tress, M.L., Abascal, F., Valencia, A., 2017. Alternative splicing may not be the key to proteome complexity. *Trends Biochem. Sci.* 42 (2), 98–110.
- Tsuboi, T., Kuroha, K., Kudo, K., Makino, S., Inoue, E., Kashima, I., Inada, T., 2012. Dom34:hbs1 plays a general role in quality-control systems by dissociation of a stalled ribosome at the 3' end of aberrant mRNA. *Mol. Cell* 46 (4), 518–529.
- Turunen, J.J., Niemela, E.H., Verma, B., Frilander, M.J., 2013. The significant other: splicing by the minor spliceosome. *Wiley Interdiscip. Rev. RNA* 4 (1), 61–76.
- Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., Darnell, R.B., 2003. CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302 (5648), 1212–1215.
- Van der Wal, E., Bergsma, A.J., Pijnenburg, J.M., Van der Ploeg, A.T., Pijnappel, W.W.M., 2017a. Antisense oligonucleotides promote exon inclusion and correct the common c.-32-13T>G GAA splicing variant in Pompe disease. *Mol. Ther. Nucleic Acids* 7, 90–100.
- Van der Wal, E., Bergsma, A.J., Van Gestel, T.J.M., In 't Groen, S.L., Zaehres, H., Araúzo-Bravo, M.J., Schöler, H.R., Van der Ploeg, A.T., Pijnappel, W.W.M., 2017b. GAA deficiency in Pompe disease is alleviated by exon inclusion in iPSC-derived skeletal muscle cells. *Mol. Ther. Nucleic Acids* 7, 101–115.
- Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., Stanton, R., Rigo, F., Guttman, M., Yeo, G.W., 2016. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* 13 (6), 508–514.
- Venables, J.P., Lapasset, L., Gadea, G., Fort, P., Klinck, R., Irimia, M., Vignal, E., Thibault, P., Prinos, P., Chabot, B., Abou Elela, S., Roux, P., Lemaître, J.M., Tazi, J., 2013. MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. *Nat. Commun.* 4, 2480.
- Verma, S.K., Deshmukh, V., Nutter, C.A., Jaworski, E., Jin, W., Wadhwa, L., Abata, J., Ricci, M., Lincoln, J., Martin, J.F., Yeo, G.W., Kuyumcu-Martinez, M.N., 2016. Rbfox2 function in RNA metabolism is impaired in hypoplastic left heart syndrome patient hearts. *Sci. Rep.* 6, 30896.
- Villarroya-Beltri, C., Gutierrez-Vazquez, C., Sanchez-Cabo, F., Perez-Hernandez, D., Vazquez, J., Martin-Cofreces, N., Martinez-Herrera, D.J., Pascual-Montano, A., Mittelbrunn, M., Sanchez-Madrid, F., 2013. Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. *Nat. Commun.* 4, 2980.
- Voineagu, I., Wang, X., Johnston, P., Lowe, J.K., Tian, Y., Horvath, S., Mill, J., Cantor, R.M., Blencowe, B.J., Geschwind, D.H., 2011. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature* 474 (7351), 380–384.
- Vulin, A., Barthelemy, I., Goyenvallé, A., Thibaud, J.L., Beley, C., Griffith, G., Benchaouir, R., le Hir, M., Unterfinger, Y., Lorain, S., Dreyfus, P., Voit, T., Carlier, P., Blot, S., Garcia, L., 2012. Muscle function recovery in golden retriever muscular dystrophy after AAV1-U7 exon skipping. *Mol. Ther.* 20 (11), 2120–2133.

- Wagner, S.D., Struck, A.J., Gupta, R., Farnsworth, D.R., Mahady, A.E., Eichinger, K., Thornton, C.A., Wang, E.T., Berglund, J.A., 2016. Dose-dependent regulation of alternative splicing by MBNL proteins reveals biomarkers for myotonic dystrophy. *PLoS Genet.* 12 (9), e1006316.
- Wampfler, J., Federzoni, E.A., Torbett, B.E., Fey, M.F., Tschan, M.P., 2016. The RNA binding proteins RBM38 and DND1 are repressed in AML and have a novel function in APL differentiation. *Leuk. Res.* 41, 96–102.
- Wang, E.T., Sandberg, R., Luo, S., Khrebukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., Burge, C.B., 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* 456 (7221), 470–476.
- Wang, Y., Gogol-Doring, A., Hu, H., Frohler, S., Ma, Y., Jens, M., Maaskola, J., Murakawa, Y., Quedenau, C., Landthaler, M., Kalscheuer, V., Wieczorek, D., Wang, Y., Hu, Y., Chen, W., 2013. Integrative analysis revealed the molecular mechanism underlying RBM10-mediated splicing regulation. *EMBO Mol. Med.* 5 (9), 1431–1442.
- Wang, Y., Yang, Y., Liu, J., Chen, X.C., Liu, X., Wang, C.Z., He, X.Y., 2014. Whole dystrophin gene analysis by next-generation sequencing: a comprehensive genetic diagnosis of Duchenne and Becker muscular dystrophy. *Mol. Genet. Genomics* 289 (5), 1013–1021.
- Wang, T., Xiao, G., Chu, Y., Zhang, M.Q., Corey, D.R., Xie, Y., 2015. Design and bioinformatics analysis of genome-wide CLIP experiments. *Nucleic Acids Res.* 43 (11), 5263–5274.
- Wang, Z., Lin, Y., Qiu, L., Zheng, D., Yan, A., Zeng, J., Lan, F., 2016. Hybrid minigene splicing assay verified the pathogenicity of a novel splice site variant in the dystrophin gene of a Chinese patient with typical Duchenne muscular dystrophy phenotype. *Clin. Chem. Lab. Med.* 54 (9), 1435–1440.
- Weatheritt, R.J., Sterne-Weiler, T., Blencowe, B.J., 2016. The ribosome-engaged landscape of alternative splicing. *Nat. Struct. Mol. Biol.* 23 (12), 1117–1123.
- Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., Krainer, A.R., Darnell, R.B., Zhang, C., 2014. HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep.* 6 (6), 1139–1152.
- Woll, M.G., Qi, H., Turpoff, A., Zhang, N., Zhang, X., Chen, G., Li, C., Huang, S., Yang, T., Moon, Y.C., Lee, C.S., Choi, S., Almstead, N.G., Naryshkin, N.A., Dakka, A., Narasimhan, J., Gabbeta, V., Welch, E., Zhao, X., Risher, N., Sheedy, J., Weetall, M., Karp, G.M., 2016. Discovery and optimization of small molecule splicing modifiers of survival motor neuron 2 as a treatment for spinal muscular atrophy. *J. Med. Chem.* 59 (13), 6070–6085.
- Wong, J.J., Au, A.Y., Ritchie, W., Rasko, J.E., 2016. Intron retention in mRNA: no longer nonsense: known and putative roles of intron retention in normal and disease biology. *Bioessays* 38 (1), 41–49.
- Wright, C.F., Fitzgerald, T.W., Jones, W.D., Clayton, S., McRae, J.F., van Kogelenberg, M., King, D.A., Ambridge, K., Barrett, D.M., Bayzetenova, T., Bevan, A.P., Bragin, E., Chatzimichali, E.A., Gribble, S., Jones, P., Krishnappa, N., Mason, L.E., Miller, R., Morley, K.I., Parthiban, V., Prigmore, E., Rajan, D., Sifrim, A., Swaminathan, G.J., Tivey, A.R., Middleton, A., Parker, M., Carter, N.P., Barrett, J.C., Hurler, M.E., FitzPatrick, D.R., Firth, H.V., study, D.D.D., 2015. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 385 (9975), 1305–1314.
- Wu, H., Sun, L., Wen, Y., Liu, Y., Yu, J., Mao, F., Wang, Y., Tong, C., Guo, X., Hu, Z., Sha, J., Liu, M., Xia, L., 2016. Major spliceosome defects cause male infertility and are associated with nonobstructive azoospermia in humans. *Proc. Natl. Acad. Sci. U.S.A.* 113 (15), 4134–4139.

- Xiong, H.Y., Alipanahi, B., Lee, L.J., Bretschneider, H., Merico, D., Yuen, R.K., Hua, Y., Gueroussov, S., Najafabadi, H.S., Hughes, T.R., Morris, Q., Barash, Y., Krainer, A.R., Jovic, N., Scherer, S.W., Blencowe, B.J., Frey, B.J., 2015. RNA splicing. The human splicing code reveals new insights into the genetic determinants of disease. *Science* 347 (6218), 1254806.
- Xu, L., Park, K.H., Zhao, L., Xu, J., El Refaey, M., Gao, Y., Zhu, H., Ma, J., Han, R., 2016. CRISPR-mediated genome editing restores dystrophin expression and function in mdx mice. *Mol. Ther.* 24 (3), 564–569.
- Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., Hardison, M., Person, R., Bekheirnia, M.R., Leduc, M.S., Kirby, A., Pham, P., Scull, J., Wang, M., Ding, Y., Plon, S.E., Lupski, J.R., Beaudet, A.L., Gibbs, R.A., Eng, C.M., 2013. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* 369 (16), 1502–1511.
- Yang, J., Hung, L.H., Licht, T., Kostin, S., Looso, M., Khrameeva, E., Bindereif, A., Schneider, A., Braun, T., 2014. RBM24 is a major regulator of muscle-specific alternative splicing. *Dev. Cell* 31 (1), 87–99.
- Ye, Z., Chen, Z., Lan, X., Hara, S., Sunkel, B., Huang, T.H., Elnitski, L., Wang, Q., Jin, V.X., 2014. Computational analysis reveals a correlation of exon-skipping events with splicing, transcription and epigenetic factors. *Nucleic Acids Res.* 42 (5), 2856–2869.
- Yeo, G., Burge, C.B., 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.* 11 (2–3), 377–394.
- Yeo, G., Hoon, S., Venkatesh, B., Burge, C.B., 2004. Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc. Natl. Acad. Sci. U.S.A.* 101 (44), 15700–15705.
- Young, C.S., Hicks, M.R., Ermolova, N.V., Nakano, H., Jan, M., Younesi, S., Karumbayaram, S., Kumagai-Cresse, C., Wang, D., Zack, J.A., Kohn, D.B., Nakano, A., Nelson, S.F., Miceli, M.C., Spencer, M.J., Pyle, A.D., 2016. A single CRISPR-Cas9 deletion strategy that targets the majority of DMD patients restores dystrophin function in hiPSC-derived muscle cells. *Cell Stem Cell* 18 (4), 533–540.
- Yu, R.Z., Grundy, J.S., Geary, R.S., 2013. Clinical pharmacokinetics of second generation antisense oligonucleotides. *Expert Opin. Drug Metab. Toxicol.* 9 (2), 169–182.
- Zarnack, K., Konig, J., Tajnik, M., Martincorena, I., Eustermann, S., Stevant, I., Reyes, A., Anders, S., Luscombe, N.M., Ule, J., 2013. Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 152 (3), 453–466.
- Zarnegar, B.J., Flynn, R.A., Shen, Y., Do, B.T., Chang, H.Y., Khavari, P.A., 2016. irCLIP platform for efficient characterization of protein–RNA interactions. *Nat. Methods* 13 (6), 489–492.
- Zhang, C., Zhang, Z., Castle, J., Sun, S., Johnson, J., Krainer, A.R., Zhang, M.Q., 2008. Defining the regulatory network of the tissue-specific splicing factors fox-1 and fox-2. *Genes Dev.* 22 (18), 2550–2563.
- Zhang, Y., Schmid, B., Nielsen, T.T., Nielsen, J.E., Clausen, C., Hyttel, P., Holst, B., Freude, K.K., 2016. Generation of a human induced pluripotent stem cell line via CRISPR-Cas9 mediated integration of a site-specific homozygous mutation in CHMP2B. *Stem Cell Res.* 17 (1), 151–153.
- Zhou, Z., Fu, X.D., 2013. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* 122 (3), 191–207.

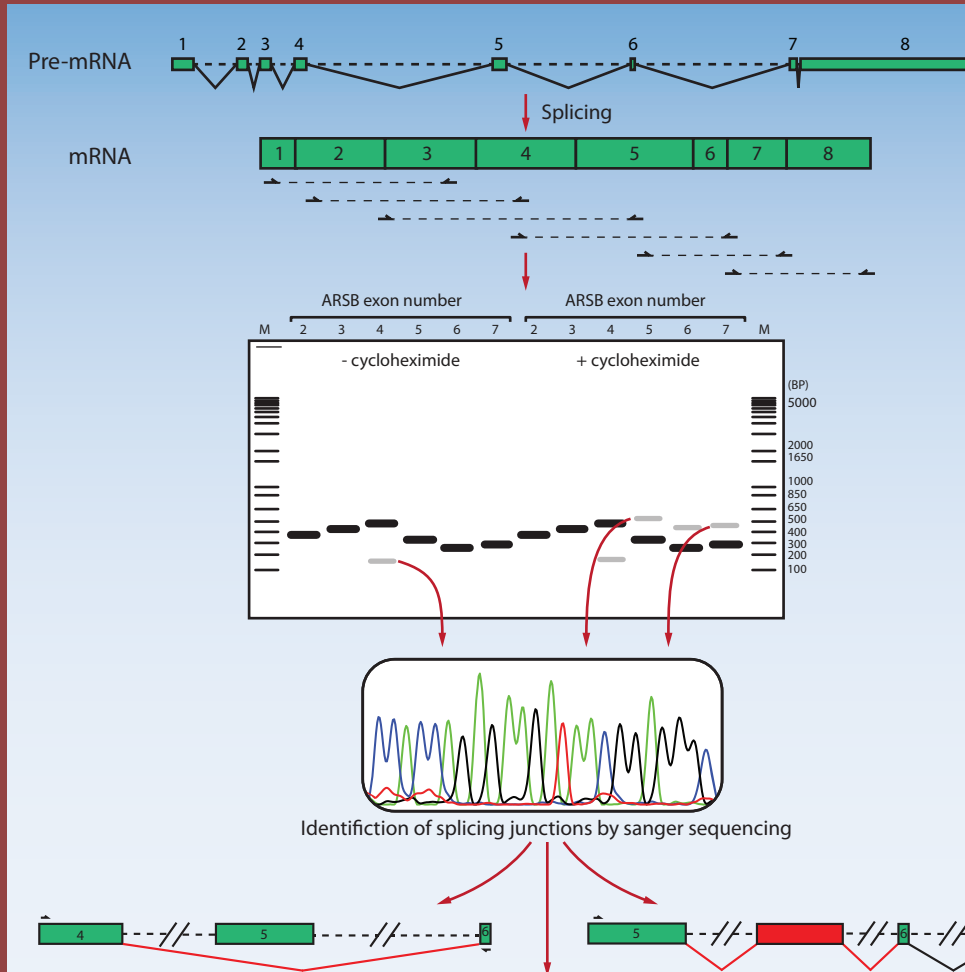


CHAPTER 3

A GENERIC ASSAY TO DETECT ABERRANT ARSB SPLICING AND MRNA DEGRADATION FOR THE MOLECULAR DIAGNOSIS OF MPS VI

Mike Broeders, Kasper Smits, Busra Goynuk, Esmee Oussoren, Hannerieke J.M.P.
van den Hout, Atze J. Bergsma, Ans T. van der Ploeg, W.W.M. Pim Pijnappel

Graphical abstract



Keywords: lysosomal disease; splicing; antisense oligonucleotides; pseudo exon; diagnosis; nonsense mediated decay

Molecular Therapy
Methods & Clinical Development

A Generic Assay to Detect Aberrant ARSB Splicing and mRNA Degradation for the Molecular Diagnosis of MPS VI

Mike Broeders,^{1,2,3} Kasper Smits,^{1,2,3} Busra Goynuk,^{1,2,3} Esme Oussoren,^{4,5} Hannerieke J.M.P. van den Hout,^{4,5} Atze J. Bergsma,^{1,2,3} Ans T. van der Ploeg,^{4,5} and W.W.M. Pim Pijnappel^{1,2,3}

¹Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; ²Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; ³Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, the Netherlands; ⁴Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, the Netherlands; ⁵Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, the Netherlands

Identification and characterization of disease-associated variants in monogenic disorders is an important aspect of diagnosis, genetic counseling, prediction of disease severity, and development of therapy. However, the effects of disease-associated variants on pre-mRNA splicing and mRNA degradation are difficult to predict and often missed. Here we present a generic assay for unbiased identification and quantification of arylsulfatase B (ARSB) mRNA for molecular diagnosis of patients with mucopolysaccharidosis VI (MPS VI). We found that healthy control individuals have inefficient ARSB splicing because of natural skipping of exon 5 and inclusion of two pseudoexons in introns 5 and 6. Analyses of 12 MPS VI patients with 10 different genotypes resulted in identification of a 151-bp intron inclusion caused by the c.1142+2T>C variant and detection of low ARSB expression from alleles with the c.629A>G variant. A special case showed skipping of exon 4 and low ARSB expression. Although no disease-associated DNA variant could be identified in this patient, the molecular diagnosis could be made based on RNA. These results highlight the relevance of RNA-based analyses to establish a molecular diagnosis of MPS VI. We speculate that inefficient natural splicing of ARSB may be a target for therapy based on promotion of canonical splicing.

INTRODUCTION

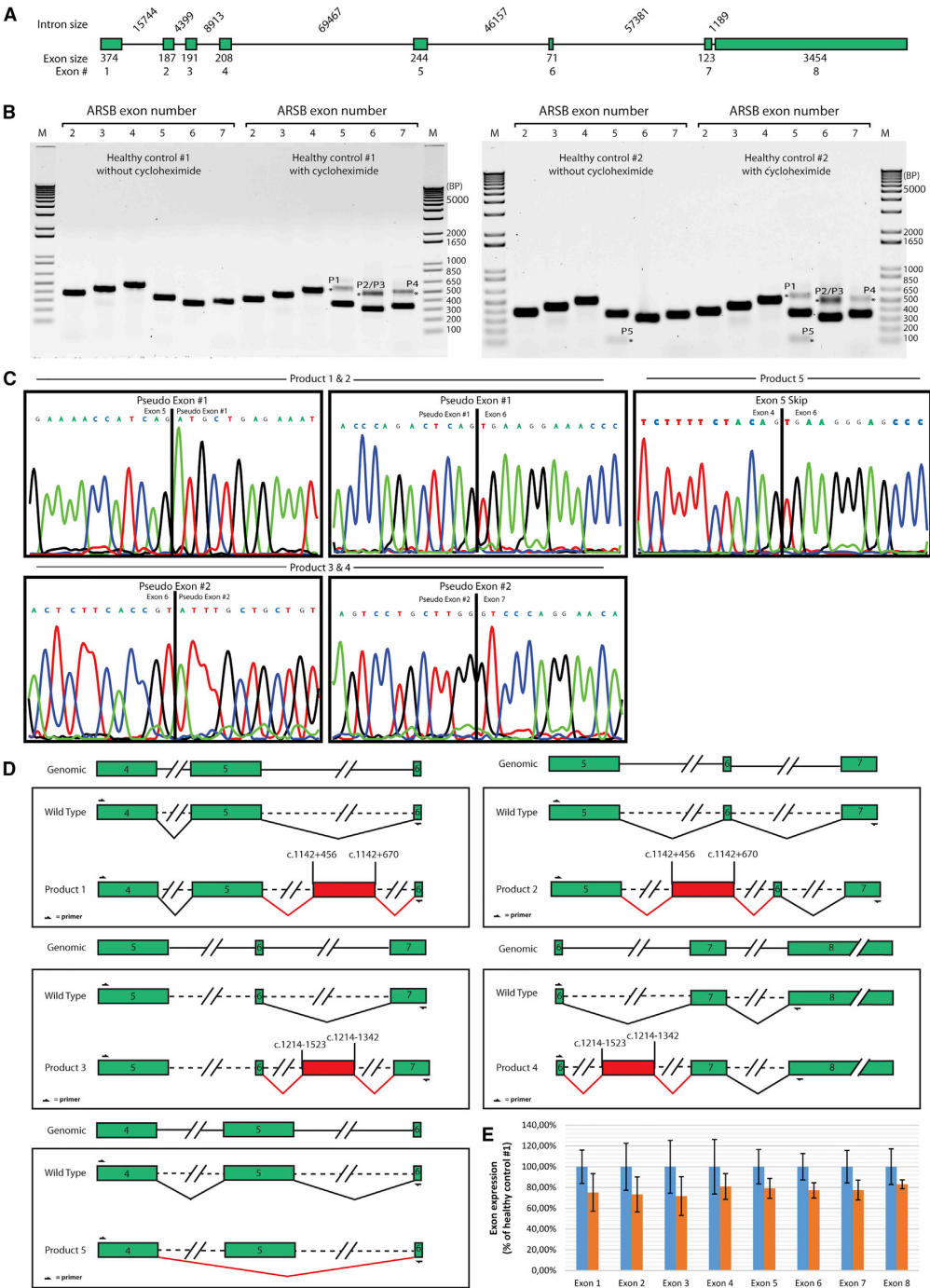
Mucopolysaccharidosis VI (MPS VI; MIM: 253200) is a lysosomal storage disease caused by disease-associated variants in the *arylsulfatase B* (ARSB) gene and has an autosomal recessive inheritance. ARSB enzyme deficiency results in failure to degrade glycosaminoglycans (GAGs) within lysosomes, resulting in their accumulation and lysosomal pathology in multiple tissues. The most prominently affected tissues include cartilage, bone, cornea, heart valves, and visceral organs.^{1–6} Patients with MPS VI can be diagnosed with a slowly or rapidly progressive form of the disease, although a broad clinical spectrum exists.^{5–8} Enzyme replacement therapy (ERT) for MPS VI, where the human recombinant ARSB enzyme is administered weekly

via intravenous infusion, has been available since 2005. ERT improves body growth and symptoms in visceral organs and extends life expectancy but has limited efficacy in connective tissues.^{8–12} The molecular diagnosis of MPS VI is based on ARSB enzyme deficiency and the presence of disease-associated variants in the ARSB gene.⁶ Current diagnostic assays for identification of DNA variants are based on sequence analysis of exons and short stretches of adjacent introns. However, disease-associated variants outside of these regions are missed in such analyses. In addition, the effects of variants on pre-mRNA splicing and/or mRNA expression remain unknown.

To date, over 200 unique variants in the ARSB gene have been reported,¹³ of which 37% are classified as likely pathogenic and 16% as pathogenic, and insufficient data are available for classification of ~45% of variants. Currently ~5% of the reported ARSB variants have been reported to affect splicing, but it is likely that this is an underestimation based on the reported frequency of splicing variants in human disease.^{14,15} This amounts to up to ~20% of disease-associated variants in general, implying that up to an additional ~30% of the 200 known ARSB variants for which insufficient information is available or that are currently classified as missense might have an undocumented effect on splicing, which should be investigated further.^{14,16,17} The effect of variants on splicing are difficult to predict *in silico*.^{18,19} Even when a strong effect on splicing is predicted, the outcome of aberrant splicing is unclear.²⁰ For instance, loss of a splice site may result in exon skipping or inclusion, intron retention, and/or utilization of a cryptic splice site. Depending on the number of nucleotides in the skipped or included sequence, the reading frame may remain intact or become disrupted. The latter can result in mRNA degradation via the nonsense-mediated decay (NMD) pathway and can remain unnoticed unless this pathway is inhibited experimentally;

Received 7 July 2020; accepted 11 September 2020;
<https://doi.org/10.1016/j.jomtm.2020.09.004>

Correspondence: W.W.M. Pim Pijnappel, Erasmus MC University Medical Center, Wytemaweg 80, 3015 CN Rotterdam, the Netherlands.
E-mail: w.pijnappel@erasmusmc.nl



(legend on next page)

for example, using cycloheximide (CHX). Identification of disease-associated variants and understanding their mechanism of action is becoming increasingly important to confirm the diagnosis, for genetic counseling, and for development of novel therapies.^{20,21} Newborn screening has been shown to be feasible for MPS VI.^{22–24} It will be of great importance for newborn screening programs to have a full understanding of ARSB disease-associated variants, including their putative effects on RNA processing and stability.

Previously, we developed a generic approach for identification and characterization of variants that affect GAA splicing in Pompe disease and applied this assay to identify multiple novel splicing events caused by disease-associated variants that were known or that were missed by standard diagnostics.^{25–27} Here we tailored this assay for detection of ARSB splicing in MPS VI and extended it to enable detection of ARSB transcripts that undergo mRNA decay. The ARSB splicing assay yielded novel information at the RNA level for 4 patients with three different genotypes of 12 MPS VI patients in the Netherlands, including an RNA-based molecular diagnosis of MPS VI in a patient lacking ARSB disease-associated DNA variants. In addition, it revealed inefficient canonical splicing of exons 5 and 6 in all analyzed patients and control individuals, providing a putative target for a generic strategy for MPS VI based on improving canonical splicing.

RESULTS

Healthy Control Individuals

The generic splicing assay consists of flanking-exon RT-PCR (using primers in exons flanking the exon of interest) and exon-internal RT-qPCR (using primers within the exon of interest) of all exons, followed by Sanger sequence analysis of splicing products, as described.²⁵ We tailored this assay to ARSB and included treatment with CHX to inhibit NMD to detect aberrant splicing products that changed the reading frame and caused a premature termination codon. Primary fibroblasts of two healthy control individuals were used to validate the approach. Flanking-exon RT-PCR analysis of all but the first and last exons showed that all canonical splicing products were detected at the expected sizes in healthy control individuals 1 and 2 (Figures 1A and 1B). Inhibition of NMD by treatment with CHX resulted in identification of additional aberrant products in healthy control individual 1 (products 1–4). Healthy control individual 2 showed identical products, and, in addition, showed a lowly expressed product (product 5) that was insensitive to NMD inhibition. Repeated analyses showed that product 5 can be easily missed in RT-PCR analysis, likely because of its low abundance, and that it could also be detected in control individual 1 in some cases (data not shown). Sequence

analysis showed that products 1 and 2 contained an inclusion of a pseudoexon of 214 nt from c.1142+456 to c.1142+670 in intron 5. Analysis of products 3 and 4 showed inclusion of a pseudoexon of 181 nt from c.1214–1523 to c.1214–1342 in intron 6. Product 5 was the result of a perfect skip of exon 5, resulting in loss of 244 nt (Figures 1C and 1D). All of these products resulted in an out-of-frame product that was predicted to undergo NMD. In agreement, products 1–4 were only detected after inhibition of NMD; however, product 5 was detected irrespective of CHX treatment, suggesting escape from NMD. RT-qPCR analysis showed quantification of all ARSB exons in healthy control individuals 1 and 2. The expression levels of all exons were similar in both healthy control individuals, with slightly lower levels in control individual 2 compared control individual 1 (Figure 1E). In summary, these results established the generic splicing assay for ARSB and showed inefficient ARSB splicing in healthy control individuals.

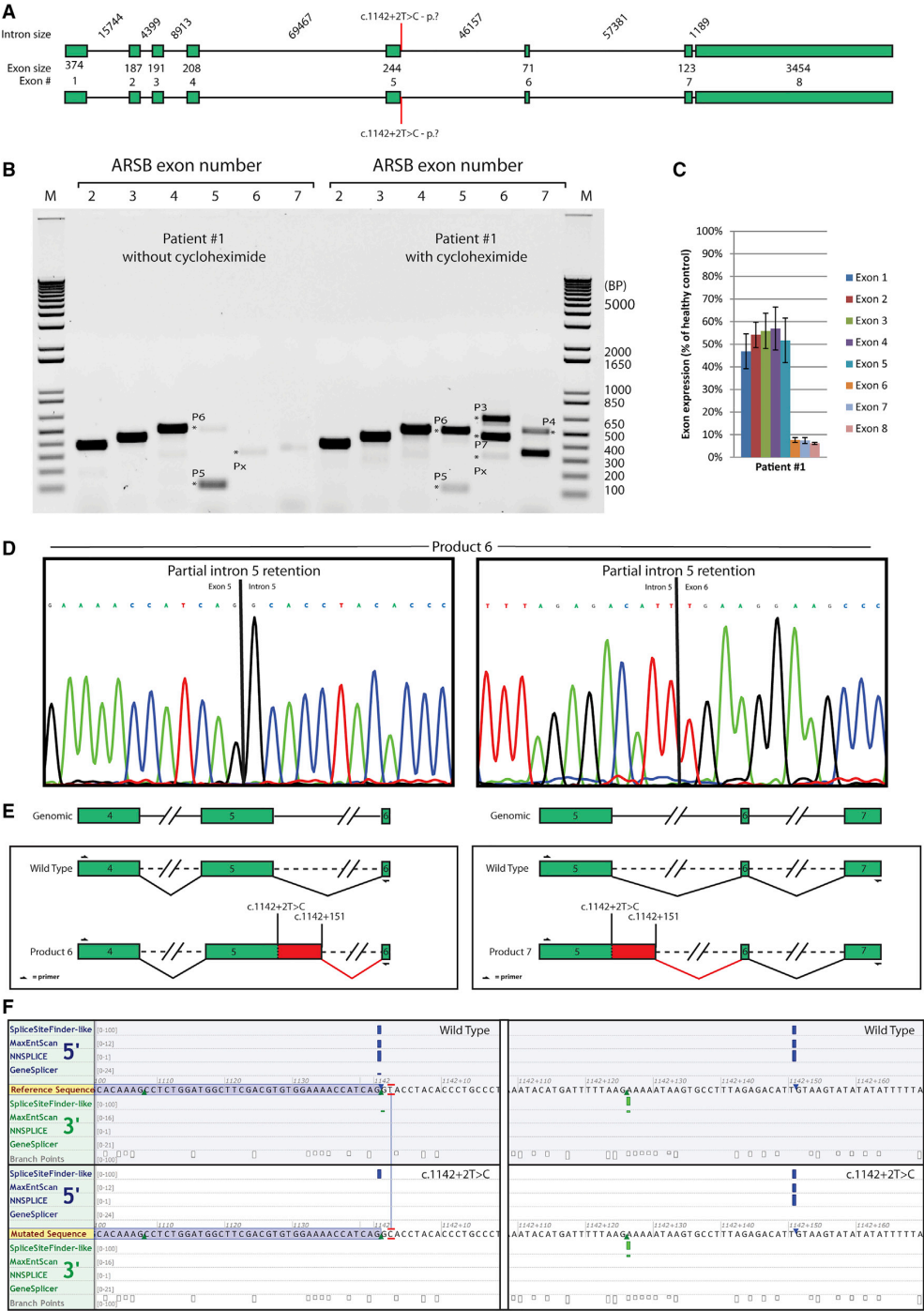
Patient 1

Patient #1 is homozygous for c.1142+2T>C, which is located near the splice donor of exon 5 (Figure 2A; Table 1). This variant has been published before by our center.⁷ No splicing analysis has been reported by us or others, but Garrido et al.^{28,29} reported that the similar variant c.1142+2T>A induces skipping of exon 5. *In silico* splicing prediction of both variants indicated loss of the canonical splice donor of exon 5 (Figure 2F, left panel; data not shown). Flanking-exon RT-PCR analysis of all exons showed the presence of canonical splicing products for exons 2–4 but very low expression of canonical splicing products for exons 5–7 (Figure 2B). After inhibition of NMD, several splicing products for exons 5–7 were detected, suggesting that the 3' part of the ARSB mRNA was mostly degraded under normal growth conditions (see below for identification of these products). Quantification of ARSB mRNA expression using exon-internal RT-qPCR analysis of cells grown in the absence of CHX showed that expression of exons 6–8 was below 10% of the levels in healthy control individuals. Exons 1–5 showed expression of ~50% of that of healthy control individuals, suggesting that this part of the mRNA partially escaped mRNA degradation. Flanking-exon RT-PCR analysis further showed two aberrant splice products for exon 5: one with a low molecular weight (MW) (product 5) and one with a higher MW than expected (product 6). Expression of exon 5 products in the absence of CHX was low in flanking-exon RT-PCR because of low expression of exon 6, in which the reverse primer was located. Sequence analysis indicated that product 5 was the result of complete skipping of exon 5. However, this exon 5 skip cannot be attributed to the c.1142+2T>C variant because it also occurs in healthy control individuals (Figure 1), and there was no indication that the level of exon 5 skipping was

Figure 1. Splicing Analysis of Two Healthy Control Individuals

(A) Schematic overview of the ARSB gene with intron size, exon size, and exon number. (B) Flanking-exon RT-PCR analysis of two healthy control individuals. Exon numbers are indicated above the lanes. PCR products were separated by electrophoresis on a 1% agarose gel. Asterisks indicate alternative splicing products detected in healthy individuals. Numbers besides the bands refer to the products analyzed in further detail. (C) Sequence analysis of the aberrant splicing products shown in (B). (D) Cartoons of splicing events detected in healthy control individuals. Exons are indicated as green boxes. Introns are depicted as lines. A broken line indicates that the intron is longer than suggested in this drawing. Canonical splicing is indicated in black and alternative splicing in red. (E) Exon-internal RT-qPCR analysis of healthy control individuals 1 and 2. Values are means relative to healthy control individual 1. GAPDH was used for normalization. Error bars indicate SD (n = 3).

Molecular Therapy
Methods & Clinical Development



(legend on next page)

increased compared with the healthy control individuals. Product 6 was identified after NMD inhibition. Sequence analysis of this product showed partial retention of intron 5 (Figures 2D and 2E, left panels). This was confirmed using flanking-exon RT-PCR analysis of exon 6, resulting in product 7, which was also larger than expected and showed the same partial inclusion of intron 5 (Figures 2D and 2E, right panels). Inclusion of the first 151 nt of intron 5 causes a shift in the reading frame, likely causing mRNA degradation starting at exon 6. Interestingly, partial retention of intron 5 seemed to prevent utilization of the cryptic splice acceptor seen in healthy control individuals in intron 5 at c.1142+456 (as judged from the absence of a band corresponding to product 1 (P1) and P2, as seen in Figure 1). Px likely represented the canonical splicing product for exon 6, based on size. Topo cloning followed by sequencing did not result in its identification, likely because of its very low abundance. Similar to healthy control individuals, inclusion of the pseudoexon in intron 6 was detected in this patient, as seen from the higher-MW P3 and P4 for flanking-exon RT-PCR analysis of exons 6 and 7, respectively (Figure 2B). Splicing prediction indicated the presence of a strong cryptic splice donor at position c.1142+151, which is the new donor that was indeed used in P6 and P7 and resulted in intron retention (Figure 2F, right panel). Taken together, these results indicate that c.1142+2T>C has the following effects: (1) it causes partial retention of 151 nt of intron 5, which leads to a frameshift and degradation of the downstream part of the mRNA; (2) it prevents utilization of the pseudoexon in intron 5; and (3) it has negligible effects on exon 5 skipping.

Patients 2–9

We applied the splicing assay to 8 additional patients with missense variants to assess whether these variants could also affect splicing. Patients 2 and 3 were homozygous for c.454C>T, located in exon 2. Patient 4 was heterozygous for c.629A>G and c.937C>G, which are located in exons 3 and 5, respectively. Patient 5–8 were homozygous for a disease-associated variant located in exon 5; patient 5 carried c.971G>T, patients 6 and 8 carried c.937C>T, and patient 7 carried c.995T>G. Patient 9 was homozygous for c.903C>G in exon 5 and c.1151G>A in exon 6. All of these missense variants have a predicted or demonstrated deleterious effect on ARSB enzyme activity.⁹ For all of these variants, *in silico* prediction indicated no effect on splicing (data not shown). Flanking-exon RT-PCR analysis of cells grown in the absence or presence of CHX indicated normal ARSB splicing in all cases (Figure S1). All analyzed patients showed inclusion of pseudoexons in introns 5 and 6, similar to healthy control individuals. We conclude that the missense variants that were present in patients 2–9 did not induce detectable aberrant ARSB splicing.

Patients 10 and 11

Patients 10 and 11 are siblings and compound heterozygous for the c.629A>G missense variant located in exon 3 and the c.979C>T nonsense variant located in exon 5 (Figure 3A). The c.979C>T variant results in a premature translation termination codon with predicted mRNA degradation via the NMD pathway. *In silico* splicing prediction of both variants indicated no effects on splicing (data not shown). Both variants have been classified previously as disease associated.^{9,30} Flanking-exon RT-PCR analysis showed all canonical splicing products in both patients, although the levels were severely reduced compared with healthy controls. The expression levels of all exons increased upon CHX treatment, indicating degradation by NMD. The same non-canonical products (i.e., products with the pseudoexons in introns 5 and 6) as in healthy controls were detected after NMD inhibition (Figures 3B and 3C). The exon-internal RT-qPCR assay confirmed the low levels of ARSB mRNA expression, which were approximately 20% of the values of healthy control individuals for all exons in both patients. These results indicate that these patients have low ARSB mRNA expression. Although a 50% reduction in mRNA expression can be explained by NMD of mRNA that is expressed from the c.979C>T allele, the cause of the additional 30% reduction in mRNA expression from the c.629A>G allele is unclear. Further research is needed to investigate whether this additional reduction in mRNA expression is caused by c.629A>G itself or by another, still unidentified ARSB sequence variant.

Patient 12

Patient 12 was diagnosed with MPS VI based on deficient ARSB enzymatic activity and clinical symptoms; however, no disease-associated variants in the ARSB gene were found at the DNA level by standard diagnostic analysis (Figure 4A). Flanking-exon RT-PCR showed that all canonical splicing products were present at very low levels (Figure 4B). This was likely due to degradation by NMD because expression of all exons was elevated following inhibition of NMD using CHX. A very low level of expression for the canonical exon 4 product was detected (top band in RT-PCR of exon 4; Figure 4). Exon-internal RT-qPCR analysis confirmed the low ARSB expression, with levels of all exons that were below 10% compared with the levels of healthy control individual 1 (Figure 4C). Besides all canonical splicing products, an aberrant product (P8) with a low MW was detected using flanking-exon RT-PCR analysis of exon 4. This product was the result of a perfect skip of exon 4, as shown by sequence analysis (Figure 4D). The size of exon 4 is 208 nt, and skipping of this exon results in a frameshift, likely causing NMD. This was indeed observed because the exon 4-skipped mRNA was more abundant following inhibition of NMD (Figure 4E). In addition, we observed that the pseudoexon in intron 5 (P1 and P2 in Figure 1)

Figure 2. Splicing Analysis of Patient 1

(A) A schematic overview of the ARSB gene, with the disease-associated variant indicated by a red solid line. (B) Flanking-exon RT-PCR analysis of patient 1. (C) Exon-internal RT-qPCR analysis of patient 1. Values are means relative to healthy control individual 1. GAPDH was used for normalization. Error bars indicate SD ($n = 3$). (D) Sequence analysis of the aberrant splicing products shown in (B). (E) Cartoons of aberrant splicing events detected in patient 1. (F) Alamut Visual splicing prediction for five algorithms of variant c.1142+2T>C compared with the wild type. The left panel shows the canonical splice donor, and the right panel shows the splice donor used in P6 and P7. Larger bars indicate higher scores in the algorithms.

Table 1. Characteristics of Included Patients

| Patient | ADNA (cDNA HGVS Nomenclature) | Location | ΔProtein (HGVS Nomenclature) | ADNA (cDNA HGVS Nomenclature) | Location | ΔProtein (HGVS Nomenclature) | Type of Disease Progression | ARSB Activity in Fibroblast (nmol/h*mg) | Age at Diagnosis (Years) | Ethnicity |
|---------|-------------------------------|-------------|------------------------------|-------------------------------|-------------|------------------------------|-----------------------------|---|--------------------------|----------------|
| 1 | c.1142+2T>C | intron 5 | p.? | c.1142+2T>C | intron 5 | p.? | rapidly | 84.8 | 2.9 | Pakistani |
| 2 | c.454C>T | exon 2 | p.(R152W) | c.454 C>T | exon 2 | p.(R152W) | slowly | 61.9 | 0.7 | Turkish |
| 3 | c.454C>T | exon 2 | p.(R152W) | c.454C>T | exon 2 | p.(R152W) | slowly | 79.9 | 7.5 | Turkish |
| 4 | c.629A>G | exon 3 | p.(Y210C) | c.937C>G | exon 5 | p.(P313A) | slowly | 50.2 | 10.3 | Dutch |
| 5 | c.971G>T | exon 5 | p.(G324V) | c.971 G>T | exon 5 | p.(G324V) | rapidly | 57.6 | 1.4 | Guinea |
| 6 | c.937C>T | exon 5 | p.(P313S) | c.937C>T | exon 5 | p.(P313S) | rapidly | 41.8 | 6 | Caribbean |
| 7 | c.995T>G | exon 5 | p.(V332G) | c.995T>G | exon 5 | p.(V332G) | rapidly | 69.4 | 2.7 | Moroccan |
| 8 | c.937C>G | exon 5 | p.(P313A) | c.937C>G | exon 5 | p.(P313A) | rapidly | ? | 4.6 | Dutch |
| 9 | c.[903C>G;1151G>A] | exons 5 & 6 | p.([N301K];[S384N]) | c.[903C>G;1151G>A] | exons 5 & 6 | p.([N301K];[S384N]) | rapidly | 85.7 | 1.9 | Turkish |
| 10 | c.629A>G | exon 3 | p.(Y210C) | c.979C>T | exon 5 | p.(R327X) | slowly | 40.7 | 6.4 | Dutch |
| 11 | c.629A>G | exon 3 | p.(Y210C) | c.979C>T | exon 5 | p.(R327X) | slowly | 82.3 | 5.9 | Dutch |
| 12 | | unknown | | | unknown | | rapidly | 36.3 | 3.1 | Dutch/Zimbabwe |

ARSB activity range in fibroblasts: healthy individuals, 320–1,080 nmol/h*mg; patients, 19–105 nmol/h*mg

seen in all individuals was lacking, likely because of the exon 4 skip, but that the pseudoexon in intron 6 (P3 and P4 in Figure 1) remained present.

To identify the genomic DNA variant that might be involved, we performed sequence analysis of all exons and their boundaries (up to 500 nt into the introns), but this failed to result in identification of any disease-associated variant. Because ARSB has very large intron sizes (8.9 kb for intron 3 and 69.4 kb for intron 4), it could be a candidate for recursive splicing (RS). The predicted consensus sequence AGGTRAGW used by Sibley et al.³¹ was present eight times at several locations in intron 4, but targeted sequencing of these regions did not reveal the presence of any ARSB DNA sequence variant, and RNA sequencing (RNA-seq) analysis did not reveal RS in ARSB (data not shown). Heterozygous SNPs were detected at the DNA level in intron 2 and exon 5 and at the RNA level at a 50/50 ratio in exon 5, arguing against a large gene deletion or a promoter variant.

To examine the genomic organization, a SNP array of the patient as well as of the parents was performed. This showed two maternal gains of 484 kb and 225 kb on the X chromosome and two paternal gains on chromosome 5 and 17 of 38 kb and 300 kb, respectively (Figure 4F). The X chromosome gains and the gain on chromosome 17 did not contain any genes known to be involved in splicing that could explain the results of the splicing assay. The gain on chromosome 5 was located in 5q14.1, which is also the location of the ARSB gene, and consisted of exons 2–4 of the ARSB gene. Although the exact location of this gain could not be determined, it is possible that it would disrupt one allele of patient 12 and cause skipping of exon 4. Nevertheless, it is unclear how this putative mono-allelic disruption could explain the results of the splicing assay because the results of the flanking-exon RT-PCR and the exon-internal RT-qPCR suggested degradation of both alleles. To examine this further, we analyzed the parents. Both parents showed expression levels of ARSB that were ~50% of levels in healthy control individuals. The father contributed the allele with the gain in 5q14.1, which might have disrupted ARSB expression to a large extent and might have caused skipping of exon 4. It is unclear why exon 4 skipping was minimally present in the father and so abundant in the patient. The mother likely contributed an allele with low ARSB expression, but it remained unclear what the underlying DNA variant was. Irrespective of the lack of an ARSB genotype for this patient, the results emphasize that analysis at the RNA level using the generic splicing assay enabled molecular diagnosis of MPS VI, consisting of low ARSB mRNA expression and skipping of exon 4.

DISCUSSION

Aberrant ARSB Splicing in Healthy Control Individuals

ARSB is a very large gene (total size, 208 kb) with exceptionally large introns; intron 1 is large at 15.7 kb, but introns 4, 5, and 6 are huge at 69.4, 46.1, and 57.3 kb, respectively. In comparison, the average size of a human protein-coding gene is 67 kb, and the average intron size is

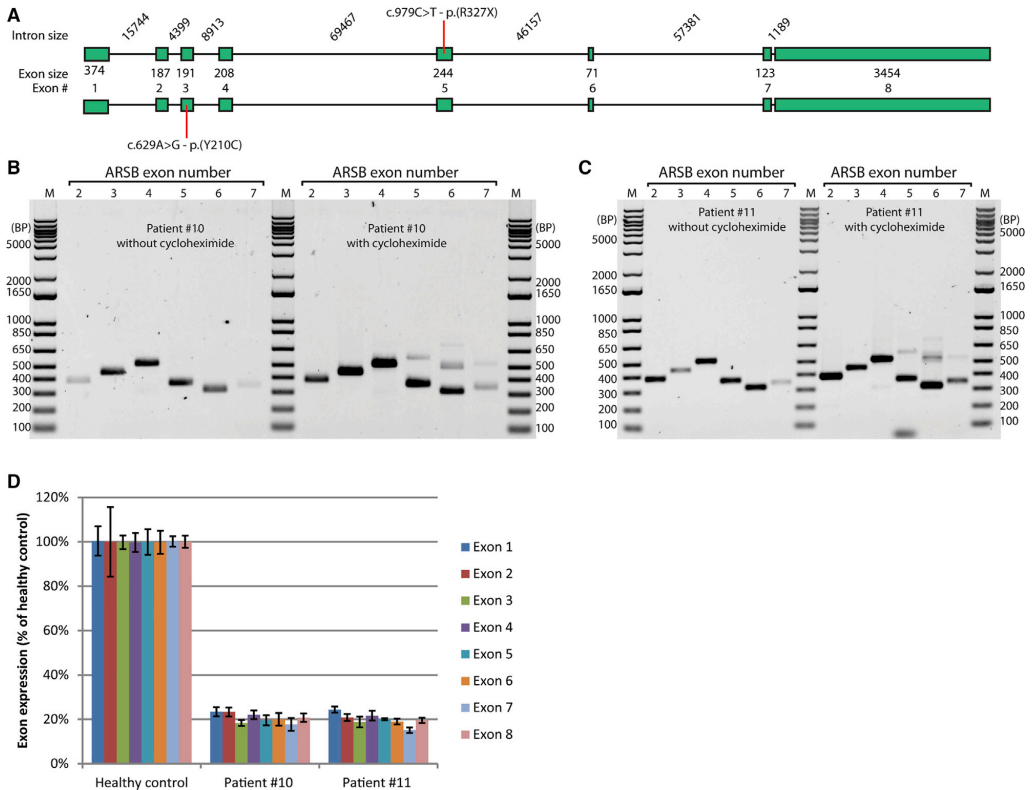


Figure 3. Splicing Analyses of Patients 10 and 11

(A) Schematic overview of the *ARSB* gene, with the disease-associated variant indicated by a red solid line. (B) Flanking-exon RT-PCR analysis of patient 10. (C) Flanking-exon RT-PCR analysis of patient 11. (D) Exon-internal RT-qPCR analysis of patients 10 and 11. Values are means relative to healthy control individual 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3).

~6 kb.³² This poses several challenges for production of ARSB protein. First, the pre-mRNA transcript is long, so takes more time than average for RNA polymerase II to generate a full-length transcript. Second, correct pre-mRNA splicing is a challenge because long introns contain, on average, more competitive cryptic splice sites, and long introns impose physical challenges for correct lariat formation during splicing.

Indeed, we found that splicing of *ARSB* is rather inefficient; exon 5 was skipped, and pseudoexons in introns 5 and 6 were utilized in a subset of transcripts. All of these aberrant splicing events resulted in an out-of-frame transcript. Skipping of exon 5 was present at low levels, produced an out-of-frame mRNA, but escaped NMD, as is sometimes the case.³³ Detection of exon 5 skipping in healthy control individuals should prevent misinterpretation of the effects of disease-associated *ARSB* variants in MPS VI.

Inclusion of pseudoexons in introns 5 and 6 did result in NMD and could only be detected after CHX treatment. Their inclusion was not obvious from *in silico* predictions (Figures S2 and S3). For the pseudoexon in intron 5, the splice acceptor of exon 6 had a relatively weak predicted strength, but the predicted strength of the cryptic splice acceptor at c.1142+456 was much weaker. Interestingly, the splice donor of this pseudoexon at c.1142+670 could not be predicted by any of the programs used. Cryptic splice sites of the pseudoexon in intron 6 were predicted *in silico* and were moderately strong. This highlights the need for experimental testing of splicing outcomes.

Relevance of Aberrant Natural Splicing

Although, at first sight, pseudoexon inclusion or exon skipping in healthy control individuals seem to be undesired forms of aberrant splicing, these are actually known to be among the mechanisms to regulate gene expression in various organisms ranging from plants

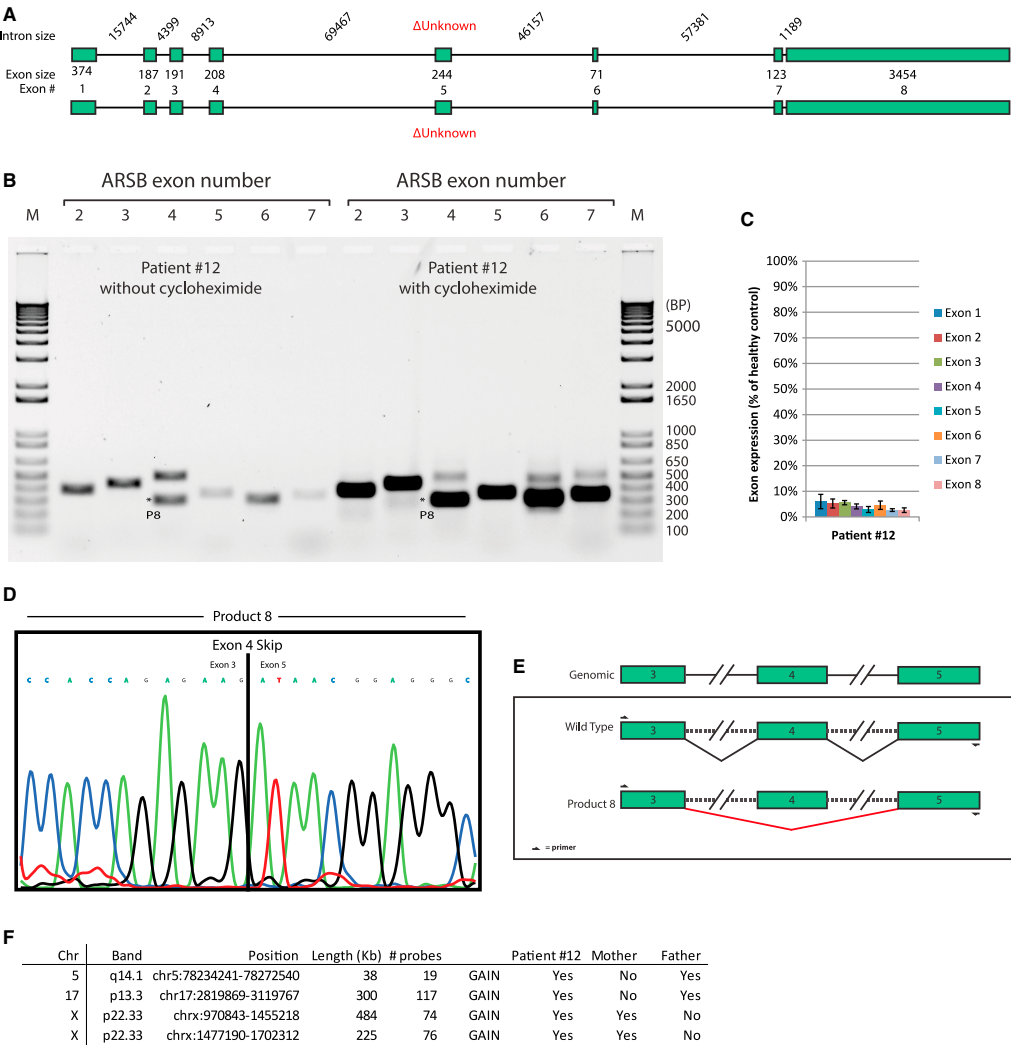


Figure 4. Splicing Analysis of Patient 12
(A) Schematic overview of the *ARSB* gene; no disease-associated variants were identified for this patient. (B) Flanking-exon RT-PCR analysis of patient 12. (C) Exon-internal RT-qPCR analysis of patient 12. Values are means relative to healthy control 1. *GAPDH* was used for normalization. Error bars indicate SD (n = 3). (D) Sequence analysis of the aberrant splicing products shown in (B). (E) Cartoon of the aberrant splicing event detected in patient 12. (F) SNP array results for patient 12 and parents.

to humans. For example, it has been shown that many genes involved in cell cycle regulation, splicing, development, tissue homeostasis, and immune regulation are subject to alternative splicing to produce aberrant transcripts in healthy individuals.^{20,34–36} This may reflect the need to rapidly regulate RNA expression levels during dynamic cellular processes. In *ARSB*, natural aberrant splicing might be useful to provide a level of regulation for turnover of GAGs in cartilage via *ARSB*-mediated degradation, which is known to be highly dynamic in response to mechanical stress.³⁷ Natural aberrant splicing can also play a role in human disease. Natural cryptic splice sites can overtake

canonical splicing when a canonical splice site is weakened; for example, as seen in β -thalassemia^{38–41} and in Pompe disease, caused by the common c.-32-13T>G (IVS1) variant.⁴² However, in the case of MPS VI, we did not observe simple competition between inclusion of a canonical versus a pseudoexon. For example, in patient 1, the c.1142+2T>C variant weakened the splice donor of exon 5, resulting in retention of the most 5' part of intron 5 and inhibition rather than promotion of utilization of the pseudoexon in intron 5. Cryptic splice sites and pseudoexons also offer the opportunity to design antisense oligonucleotides (AONs) that block utilization of cryptic splice sites and promote canonical splicing in human disease.^{20,21,43} We and others have shown recently that it is feasible for the lysosomal storage disease Pompe disease to restore GAA expression after aberrant splicing caused by several GAA variants, including c.1552-3C>G, c.1256A>T, c.2190-345A>G,²⁶ and the common IVS1 variant.^{42,44–46} However, AONs that blocked the cryptic splice sites of the pseudoexons in ARSB introns 5 and 6 failed to promote canonical ARSB splicing (Figure S4), suggesting a more complex underlying mechanism for ARSB splicing. We speculate that the large sizes of introns 5 and 6 offer too many alternative options for splicing to allow a simple competition model between utilization of canonical splice sites and the cryptic splice sites of the pseudoexons.

Unexpected Aberrant Splicing and mRNA Expression in MPS VI Patients

We found that the c.1142+2T>C variant in patient 2 caused intron retention of the most 5' part of intron 5 by utilizing a predicted cryptic splice acceptor site (patient 1; Figure 2). A similar variant, c.1142+2T>A, has been described in the literature. In that case, only cells from the father that carried c.1142+2T>A were analyzed in the absence of CHX. It was concluded that this variant causes skipping of exon 5. However, we found that exon 5 skipping also occurs in healthy control individuals, opening the possibility that c.1142+2T>A was not the cause of the exon 5 skip and that this variant may also cause intron 5 retention. This would require further analysis of cells grown in the presence of CHX because the intron retention product can be easily missed in normally growing cells because of NMD.

Patient 1 in the present study showed NMD of only the 3' part of the mRNA (Figure 2). At present we do not have an explanation for this observation. Aberrant splicing around exon 5 shows variable levels of NMD; exon 5 skipping occurs in healthy and diseased individuals, causes a frameshift, but does not induce NMD, whereas a premature termination codon in exon 5 does induce NMD (caused by c.979C>T in patients 10 and 11; Figure 3). Although exons 1–5 encode the ARSB signal peptide and the sulfatase domains, exons 6–8 encode the C terminus of the ARSB protein, for which no conserved domain is presently known.^{47,48} Future experiments are needed to elucidate the mechanisms underlying regulation of NMD around exon 5.

Patients 10 and 11 contained the c.629A>G variant on the second allele. ARSB mRNA expression was 20% rather than the expected 50% for all exons (the other allele underwent NMD), indicating that the c.629A>G allele was associated with low expression. We spec-

ulate that this missense was linked to another variant that was responsible for the low mRNA levels of this allele (for example, a promoter variant), which needs further analysis. Nevertheless, this result highlights the value of quantitative analysis of ARSB mRNA expression to detect effects on mRNA expression levels.

A Complex Case: Molecular Diagnosis at the mRNA Level

Patient 12 was deficient in ARSB enzyme activity, but no DNA variants could be identified by standard diagnostics. We found a number of aberrations, including low mRNA expression of all exons, skipping of exon 4, and gain of ARSB exons 2–4 in 5q14.1, which is the location of the ARSB gene. This helped to establish the molecular diagnosis at the mRNA level, but we could not identify the DNA variants that were responsible. Below we discuss possible scenarios for this patient.

The father contains gain of ARSB exons 2–4 in 5q14.1 on one allele. He has ~50% ARSB mRNA expression from both alleles (data not shown) and shows very low but detectable skipping of exon 4 (Figure S5). We hypothesize that the gain of ARSB 2–4 landed in the ARSB gene to induce exon 4 skipping and disrupt ARSB expression via NMD. However, the father had only minor levels of exon 4 skipping compared with patient 12. We hypothesize that patient 12 may contain a *de novo* variant in *cis* or in *trans* that results in worsening of exon 4 skipping and NMD.

The mother also has low ARSB mRNA expression of ~50% of the levels of healthy control individual 1 levels (data not shown). No disease-associated ARSB variant was detected in the mother. A heterozygous SNP was detected in exon 5 in mRNA from patient 12 at a 50% ratio in the absence and presence of CHX (data not shown), indicating that both alleles were expressed. This argued against a simple scenario in which the maternal allele underwent NMD. The underlying molecular mechanism remained obscure and may suggest a *de novo* variant in the maternal allele that caused low mRNA expression in patient 12.

Conclusions

In conclusion, of 12 MPS VI patients who were present in the Netherlands at the time of analysis, we identified novel aberrant splicing/mRNA expression events in 4 patients with three different genotypes. This highlights the need to implement systematic diagnostic analyses at the mRNA level to establish a molecular diagnosis and to understand the pathogenic mechanism of ARSB variants that can cause MPS VI. For future diagnostic implementation of the generic splicing assay, we anticipate that next-generation sequencing will be useful to make this assay suitable for diagnosis of a broad spectrum of human diseases within a single run.

MATERIALS AND METHODS

Patients and Healthy Control Individuals

Patients were diagnosed with MPS VI at the Center for Lysosomal and Metabolic Diseases of Erasmus MC, Rotterdam, the Netherlands. Diagnosis was based on clinical symptoms, ARSB enzyme deficiency in leukocytes and/or fibroblasts, and ARSB variant analysis. Analysis

was performed on patient material obtained with informed consent. For healthy control individuals, analysis was performed on material for an unrelated disease obtained with informed consent. The study protocol was approved by the Medical Ethical Committee at Erasmus Medical Center, Rotterdam.

Nomenclature

The nomenclature of variants and splice sites is according to Human Genome Variation Society (HGVS) standards (<http://www.hgvs.org/mutnomen/>). RefSeq NM_000046.5 was used as a *ARSB* reference transcript and NC_000005.10 as the genomic sequence.⁴⁹

Splicing Prediction

Splicing prediction was carried out as described before.²⁵ Alamut Visual version 2.6.1 was used to predict 5' and 3' splice junctions using five different algorithms (see the description of algorithms at <http://www.interactive-biosoftware.com/doc/alamutvisual/2.6/splicing.html>).

Cell Culture

Primary human fibroblasts were cultured in high-glucose DMEM (Lonza), 10% fetal bovine serum (Thermo Fisher Scientific), and 100 U/mL penicillin/streptomycin/glutamine (Lonza). Cells were grown in the presence of 5% CO₂. Cells were passaged at 80% to 90% confluency with TrypLE (Gibco). All cell lines were routinely tested for mycoplasma infection using the MycoAlert Mycoplasma Detection Kit (Lonza) and were negative. To inhibit NMD, 20 µg/ml CHX (Sigma) was added to the medium 48 h prior to RNA isolation. Phosphorodiamidate morpholino oligo (PMO) AON design and treatment at 20 µM final concentration was performed as described previously.^{26,42,44}

RNA Isolation and cDNA Preparation

RNA was isolated using the RNeasy Miniprep Kit with on-column DNase treatment (QIAGEN) using guidelines provided by the manufacturer. 400 ng of RNA was used for generation of cDNA using the iScript cDNA Synthesis Kit (Bio-Rad). The cDNA solution was diluted 10 times before use in RT-PCR or RT-qPCR.

Flanking-Exon RT-PCR Analysis and Exon-Internal RT-qPCR Analysis

Flanking-exon RT-PCR was performed with FastStart Taq Polymerase (Roche) on a Bio-Rad S1000 thermal cycler. Exon-internal RT-qPCR was carried out using iTaq SYBR Green Supermix (Bio-Rad) on a cfx96rts cycler (Bio-Rad). All samples were measured in triplicate and normalized against Glyceraldehyde-3-Phosphate Dehydrogenase (*GAPDH*). All primer sets used for RT-qPCR showed high efficiency and specificity based on melting-curve analysis and standard curve measurements. Primers are shown in Table S1.

Sequencing

Direct sequencing of products identified in flanking-exon RT-PCR was performed using the Big Dye Terminator Kit v.3.1 (Applied Biosystems). To obtain pure DNA samples, PCR products visible

on the gel were stabbed with a 20-µL pipette tip, and DNA on the tip was resuspended in 10 µL H₂O. A 1-µL aliquot was subsequently used in a new PCR (as described above). Excess primers and dinucleotide triphosphates (dNTPs) were removed using FastAP thermosensitive alkaline phosphatase (Thermo Fisher Scientific) according to the manufacturer's protocol. Alternatively, PCR products were cloned into a pCR2.1-TOPO (Thermo Fisher Scientific) cloning vector according to the manufacturer's protocol, and then M13 primers were used for sequencing (Table S1). Samples were purified with Sephadex G-50 (GE Healthcare), and the sequence was determined on an AB3130 genetic analyzer (Applied Biosystems).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.omtm.2020.09.004>.

AUTHOR CONTRIBUTIONS

M.B. and W.W.M.P.P. conceived and designed the study and drafted the manuscript. M.B., K.S., and B.G. performed the experiments. W.W.M.P.P. supervised the study. M.B., E.O., H.J.M.P.v.d.H., A.J.B., A.T.v.d.P., and W.W.M.P.P. were involved in data interpretation and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

This work has received funding from Zeldzame Ziekten Fonds/WE Foundation, Metakids (project number 2018-082), and Stofwisselkracht.

REFERENCES

1. Azevedo, A.C., Schwartz, I.V., Kalakun, L., Brustolin, S., Burin, M.G., Beheragaray, A.P., Leistner, S., Giugliani, C., Rosa, M., Barrios, P., et al. (2004). Clinical and biochemical study of 28 patients with mucopolysaccharidosis type VI. *Clin. Genet.* 66, 208–213.
2. Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D., Childs, B., Kinzler, K.W., and Vogelstein, B. (2001). *The Metabolic Bases of Inherited Disease*. (McGraw-Hill).
3. Spranger, J.W., Koch, F., McKusick, V.A., Natzschka, J., Wiedemann, H.R., and Zellweger, H. (1970). Mucopolysaccharidosis VI (Maroteaux-Lamy's disease). *Helv. Paediatr. Acta* 25, 337–362.
4. Stumpf, D.A., Austin, J.H., Crocker, A.C., and LaFrance, M. (1973). Mucopolysaccharidosis type VI (Maroteaux-Lamy syndrome). I. Sulfatase B deficiency in tissues. *Am. J. Dis. Child.* 126, 747–755.
5. Valayannopoulos, V., Nicely, H., Harmatz, P., and Turbeville, S. (2010). Mucopolysaccharidosis VI. *Orphanet J. Rare Dis.* 5, 5.
6. Giugliani, R., Harmatz, P., and Wraith, J.E. (2007). Management guidelines for mucopolysaccharidosis VI. *Pediatrics* 120, 405–418.
7. Swiedler, S.J., Beck, M., Bajbouj, M., Giugliani, R., Schwartz, I., Harmatz, P., Wraith, J.E., Roberts, J., Ketteridge, D., Hopwood, J.J., et al. (2005). Threshold effect of urinary glycosaminoglycans and the walk test as indicators of disease progression in a survey of subjects with Mucopolysaccharidosis VI (Maroteaux-Lamy syndrome). *Am. J. Med. Genet. A* 134A, 144–150.
8. Brands, M.M., Oussoren, E., Ruijter, G.J., Vollenbregt, A.A., van den Hout, H.M., Joosten, K.F., Hop, W.C., Plug, I., and van der Ploeg, A.T. (2013). Up to five years experience with 11 mucopolysaccharidosis type VI patients. *Mol. Genet. Metab.* 109, 70–76.

9. Brands, M.M., Hoogveen-Westerveld, M., Kroos, M.A., Nobel, W., Ruijter, G.J., Özkan, L., Plug, I., Grinberg, D., Vilageliu, L., Halley, D.J., et al. (2013). Mucopolysaccharidosis type VI phenotypes-genotypes and antibody response to gal-sulfase. *Orphanet J. Rare Dis.* 8, 51.
10. Decker, C., Yu, Z.F., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., Scarpa, M., et al. (2010). Enzyme replacement therapy for mucopolysaccharidosis VI: Growth and pubertal development in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J. Pediatr. Rehabil. Med.* 3, 89–100.
11. Harmatz, P., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., Scarpa, M., et al. (2008). Long-term follow-up of endurance and safety outcomes during enzyme replacement therapy for mucopolysaccharidosis VI: Final results of three clinical studies of recombinant human N-acetylgalactosamine 4-sulfatase. *Mol. Genet. Metab.* 94, 469–475.
12. Harmatz, P., Yu, Z.F., Giugliani, R., Schwartz, I.V., Guffon, N., Teles, E.L., Miranda, M.C., Wraith, J.E., Beck, M., Arash, L., et al. (2010). Enzyme replacement therapy for mucopolysaccharidosis VI: evaluation of long-term pulmonary function in patients treated with recombinant human N-acetylgalactosamine 4-sulfatase. *J. Inherit. Metab. Dis.* 33, 51–60.
13. Tomanin, R., Karageorgos, L., Zanetti, A., Al-Sayed, M., Bailey, M., Miller, N., Sakuraba, H., and Hopwood, J.J. (2018). Mucopolysaccharidosis type VI (MPS VI) and molecular analysis: Review and classification of published variants in the ARSB gene. *Hum. Mutat.* 39, 1788–1802.
14. Sterne-Weiler, T., Howard, J., Mort, M., Cooper, D.N., and Sanford, J.R. (2011). Loss of exon identity is a common mechanism of human inherited disease. *Genome Res.* 21, 1563–1571.
15. Stenson, P.D., Mort, M., Ball, E.V., Evans, K., Hayden, M., Heywood, S., Hussain, M., Phillips, A.D., and Cooper, D.N. (2017). The Human Gene Mutation Database: towards a comprehensive repository of inherited mutation data for medical research, genetic diagnosis and next-generation sequencing studies. *Hum. Genet.* 136, 665–677.
16. Lim, K.H., Ferraris, L., Filloux, M.E., Raphael, B.J., and Fairbrother, W.G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. *Proc. Natl. Acad. Sci. USA* 108, 11093–11098.
17. Soukariéh, O., Galdrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frébourg, T., Tosi, M., and Martins, A. (2016). Exonic Splicing Mutations Are More Prevalent than Currently Estimated and Can Be Predicted by Using In Silico Tools. *PLoS Genet.* 12, e1005756.
18. Moles-Fernández, A., Duran-Lozano, L., Montalban, G., Bonache, S., López-Perolio, I., Menéndez, M., Santamaría, M., Behar, R., Blanco, A., Carrasco, E., et al. (2018). Computational Tools for Splicing Defect Prediction in Breast/Ovarian Cancer Genes: How Efficient Are They at Predicting RNA Alterations? *Front. Genet.* 9, 366.
19. Schleif, J., Bailey, S.S., Tran, T., Chen, D., Stowers, S., Schwarze, U., and Byers, P.H. (2015). Molecular Outcome, Prediction, and Clinical Consequences of Splice Variants in COL1A1, Which Encodes the pro α 1(I) Chains of Type I Procollagen. *Hum. Mutat.* 36, 728–739.
20. Bergsma, A.J., van der Wal, E., Broeders, M., van der Ploeg, A.T., and Pim Pijnappel, W.W.M. (2018). Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options. *Int. Rev. Cell Mol. Biol.* 335, 85–141.
21. Kuijper, E.C., Bergsma, A.J., Pijnappel, W.W.M.P., and Aartsma-Rus, A. (2020). Opportunities and challenges for antisense oligonucleotide therapies. *J. Inherit. Metab. Dis.* Published online May 11, 2020. <https://doi.org/10.1002/jimd.12251>.
22. Arunkumar, N., Langan, T.J., Stapleton, M., Kubaski, F., Mason, R.W., Singh, R., Kobayashi, H., Yamaguchi, S., Suzuki, Y., Orii, K., et al. (2020). Newborn screening of mucopolysaccharidoses: past, present, and future. *J. Hum. Genet.* 65, 557–567.
23. Scott, C.R., Elliott, S., Hong, X., Huang, J.Y., Kumar, A.B., Yi, F., Pendem, N., Chennamaneni, N.K., and Gelb, M.H. (2020). Newborn Screening for Mucopolysaccharidoses: Results of a Pilot Study with 100 000 Dried Blood Spots. *J. Pediatr.* 216, 204–207.
24. Kubaski, F., de Oliveira Poswar, F., Michelin-Tirelli, K., Burin, M.G., Rojas-Málaga, D., Brusius-Facchin, A.C., Leistner-Segal, S., and Giugliani, R. (2020). Diagnosis of Mucopolysaccharidoses. *Diagnostics (Basel)* 10, E172.
25. Bergsma, A.J., Kroos, M., Hoogveen-Westerveld, M., Halley, D., van der Ploeg, A.T., and Pijnappel, W.W. (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. *Hum. Mutat.* 36, 57–68.
26. Bergsma, A.J., In 't Groen, S.L., Verheijen, F.W., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2016). From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides. *Mol. Ther. Nucleic Acids* 5, e361.
27. In 't Groen, S.L.M., de Faria, D.O.S., Iuliano, A., van den Hout, J.M.P., Douben, H., Dijkhuizen, T., Cassiman, D., Witters, P., Barba Romero, M.A., de Klein, A., et al. (2020). Novel GAA Variants and Mosaicism in Pompe Disease Identified by Extended Analyses of Patients with an Incomplete DNA Diagnosis. *Mol. Ther. Methods Clin. Dev.* 17, 337–348.
28. Garrido, E., Chabás, A., Coll, M.J., Blanco, M., Domínguez, C., Grinberg, D., Vilageliu, L., and Cormand, B. (2007). Identification of the molecular defects in Spanish and Argentinian mucopolysaccharidosis VI (Maroteaux-Lamy syndrome) patients, including 9 novel mutations. *Mol. Genet. Metab.* 92, 122–130.
29. Garrido, E., Cormand, B., Hopwood, J.J., Chabás, A., Grinberg, D., and Vilageliu, L. (2008). Maroteaux-Lamy syndrome: functional characterization of pathogenic mutations and polymorphisms in the arylsulfatase B gene. *Mol. Genet. Metab.* 94, 305–312.
30. Litjens, T., Brooks, D.A., Peters, C., Gibson, G.J., and Hopwood, J.J. (1996). Identification, expression, and biochemical characterization of N-acetylgalactosamine-4-sulfatase mutations and relationship with clinical phenotype in MPS-VI patients. *Am. J. Hum. Genet.* 58, 1127–1134.
31. Sibley, C.R., Emmett, W., Blazquez, L., Faro, A., Haberman, N., Briesse, M., Trabzuni, D., Ryten, M., Weale, M.E., Hardy, J., et al. (2015). Recursive splicing in long vertebrate genes. *Nature* 521, 371–375.
32. Piovesan, A., Caracausi, M., Antonaros, F., Pelleri, M.C., and Vitale, L. (2016). GeneBase 1.1: a tool to summarize data from NCBI gene datasets and its application to an update of human gene statistics. *Database (Oxford)* 2016, baw153.
33. Dyle, M.C., Kolakada, D., Cortazar, M.A., and Jagannathan, S. (2020). How to get away with nonsense: Mechanisms and consequences of escape from nonsense-mediated RNA decay. *Wiley Interdiscip. Rev. RNA* 11, e1560.
34. Baralle, F.E., and Giudice, J. (2017). Alternative splicing as a regulator of development and tissue identity. *Nat. Rev. Mol. Cell Biol.* 18, 437–451.
35. Lee, Y., and Rio, D.C. (2015). Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* 84, 291–323.
36. Dhir, A., and Buratti, E. (2010). Alternative splicing: role of pseudoexons in human disease and potential therapeutic strategies. *FEBS J.* 277, 841–855.
37. Vynios, D.H. (2014). Metabolism of cartilage proteoglycans in health and disease. *BioMed Res. Int.* 2014, 452315.
38. Dobkin, C., and Bank, A. (1983). A nucleotide change in IVS 2 of a beta-thalassemia gene leads to a cryptic splice not at the site of the mutation. *Prog. Clin. Biol. Res.* 134, 127–128.
39. Dobkin, C., and Bank, A. (1985). Reversibility of IVS 2 missplicing in a mutant human beta-globin gene. *J. Biol. Chem.* 260, 16332–16337.
40. Wong, C., Antonarakis, S.E., Goff, S.C., Orkin, S.H., Forget, B.G., Nathan, D.G., Giardina, P.J., and Kazazian, H.H., Jr. (1989). Beta-thalassemia due to two novel nucleotide substitutions in consensus acceptor splice sequences of the beta-globin gene. *Blood* 73, 914–918.
41. Treisman, R., Orkin, S.H., and Maniatis, T. (1983). Specific transcription and RNA splicing defects in five cloned beta-thalassemia genes. *Nature* 302, 591–596.
42. van der Wal, E., Bergsma, A.J., van Gestel, T.J.M., In 't Groen, S.L.M., Zaehres, H., Araújo-Bravo, M.J., Schöler, H.R., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2017). GAA Deficiency in Pompe Disease Is Alleviated by Exon Inclusion in iPSC-Derived Skeletal Muscle Cells. *Mol. Ther. Nucleic Acids* 7, 101–115.
43. Bennett, C.F. (2019). Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* 70, 307–321.
44. van der Wal, E., Bergsma, A.J., Pijnburg, J.M., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2017). Antisense Oligonucleotides Promote Exon Inclusion and Correct the Common c.-32-13T>G GAA Splicing Variant in Pompe Disease. *Mol. Ther. Nucleic Acids* 7, 90–100.

45. Aung-Htut, M.T., Ham, K.A., Tchan, M., Johnsen, R., Schnell, F.J., Fletcher, S., and Wilton, S.D. (2020). Splice modulating antisense oligonucleotides restore some acid-alpha-glucosidase activity in cells derived from patients with late-onset Pompe disease. *Sci. Rep.* 10, 6702.
46. Goina, E., Peruzzo, P., Bembi, B., Dardis, A., and Buratti, E. (2017). Glycogen Reduction in Myotubes of Late-Onset Pompe Disease Patients Using Antisense Technology. *Mol. Ther.* 25, 2117–2128.
47. Bond, C.S., Clements, P.R., Ashby, S.J., Collyer, C.A., Harrop, S.J., Hopwood, J.J., and Guss, J.M. (1997). Structure of a human lysosomal sulfatase. *Structure* 5, 277–289.
48. Pfam. Protein: ARSB_HUMAN (P15848). <http://pfam.xfam.org/protein/P15848>.
49. den Dunnen, J.T., Dalgleish, R., Maglott, D.R., Hart, R.K., Greenblatt, M.S., McGowan-Jordan, J., Roux, A.F., Smith, T., Antonarakis, S.E., and Taschner, P.E. (2016). HGVS Recommendations for the Description of Sequence Variants: 2016 Update. *Hum. Mutat.* 37, 564–569.

Supplemental Information

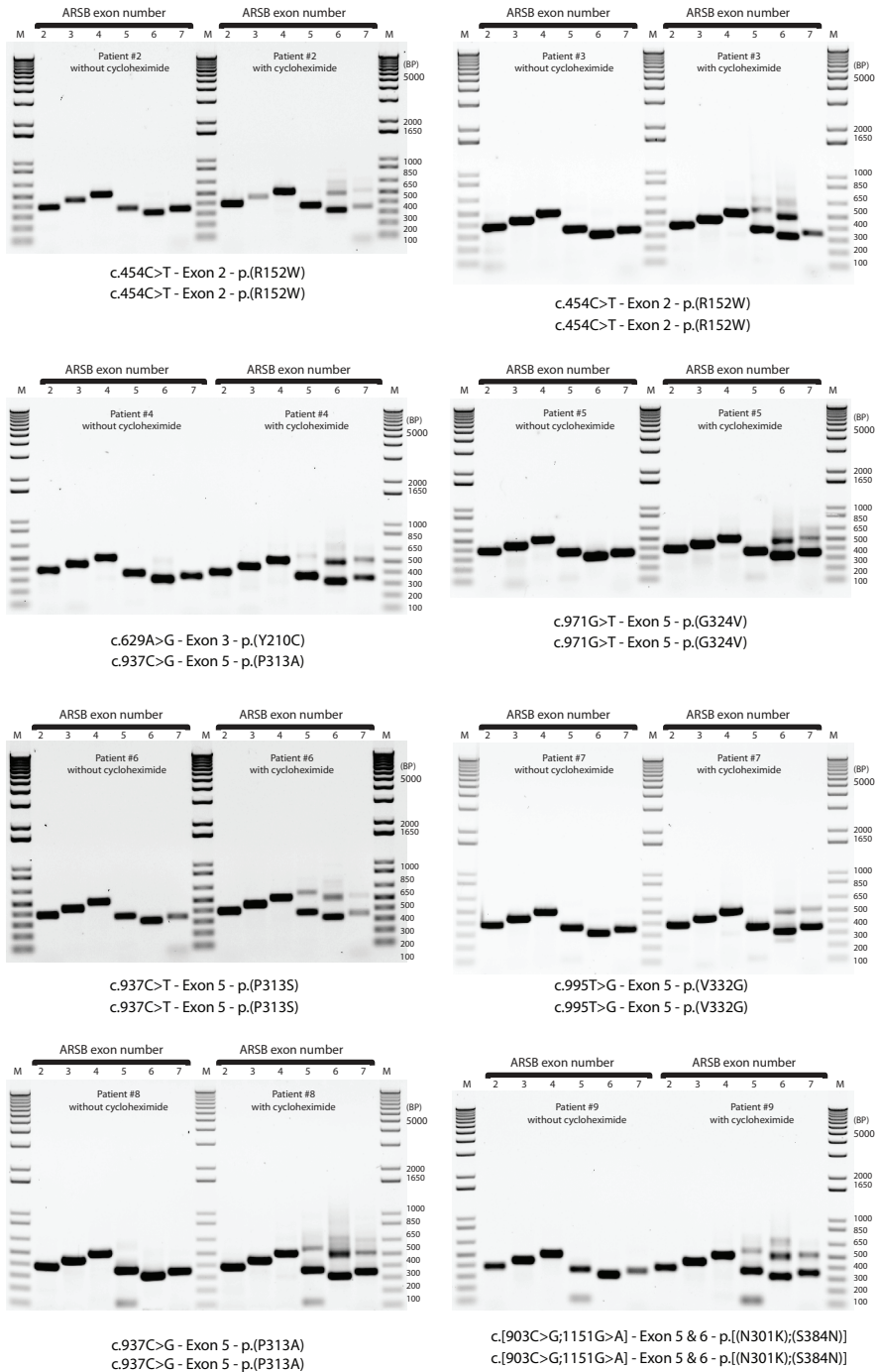


Figure S1. Analysis of patient #2 to #9 by flanking exon RT-PCR. The disease-associated variants and their location are indicated.

Figure S2. In silico prediction of splice donors and acceptors used by pseudo exons in ARSB intron 5 and intron 6.

Supplemental Information

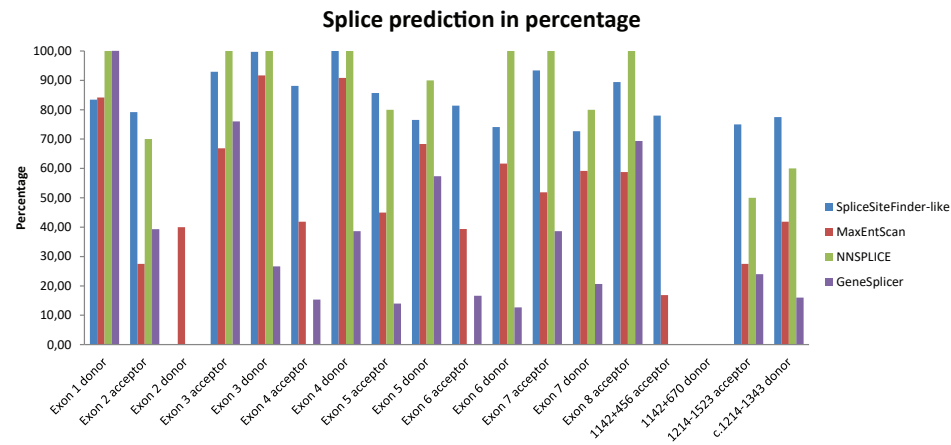


Figure S3. In silico splicing prediction of the strengths of canonical splice sites and cryptic splice sites utilized for pseudo exon inclusion.

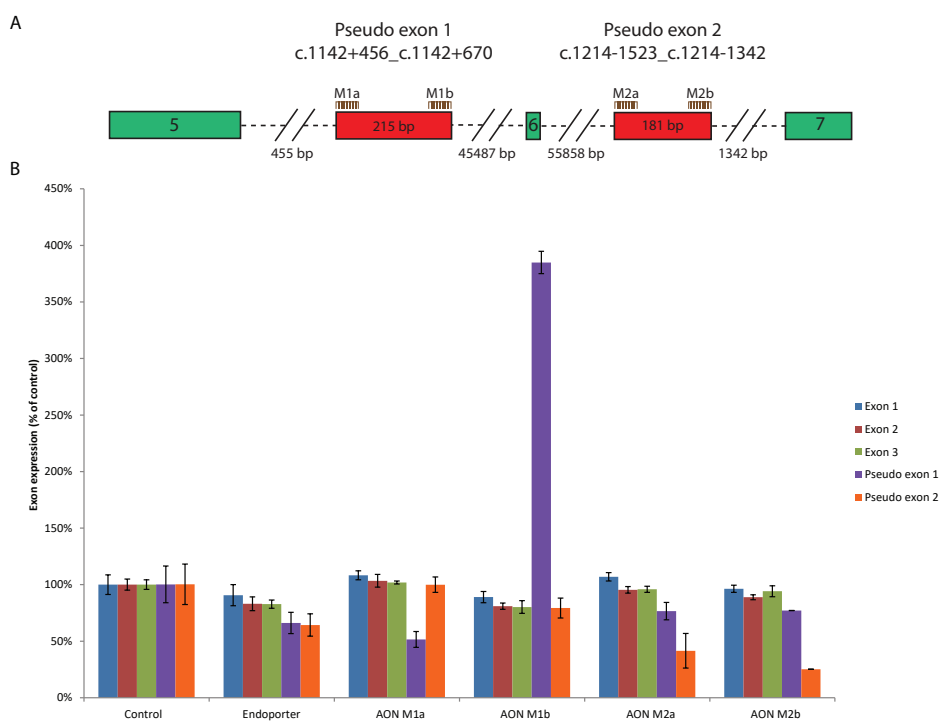


Figure S4. Antisense oligonucleotide (AON) treatment to block pseudo exon inclusion. **A:** Strategy for the AON treatment. PMO AONs (indicated in brown and shown in Tabel S1) were designed to block the splice acceptor and splice donor sites of pseudo exons 1 and 2 (indicated in red). **B:** Exon internal RT-qPCR analysis after AON treatment of healthy control fibroblasts. Treatment was for 24 hours. Although pseudo exon inclusion was decreased after AON treatment for pseudo exon 1 with M1a and for pseudo exon 2 with M2a and M2b, expression of canonical ARSB mRNA (measured for exons 1 to 3) did not increase. M1b promoted pseudo exon 1 inclusion by almost 4-fold for unknown reasons and decreased the overall expression slightly. GAPDH was used for normalization. Error bars indicate SD (n = 3).

Supplemental Information

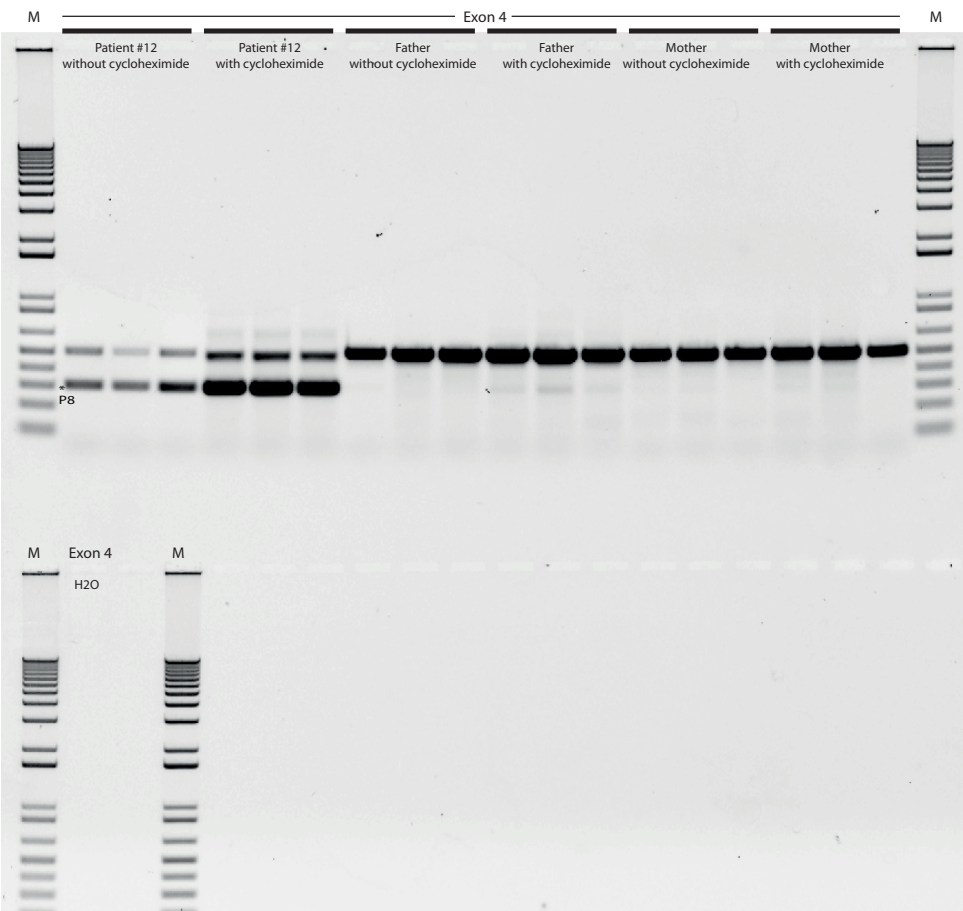
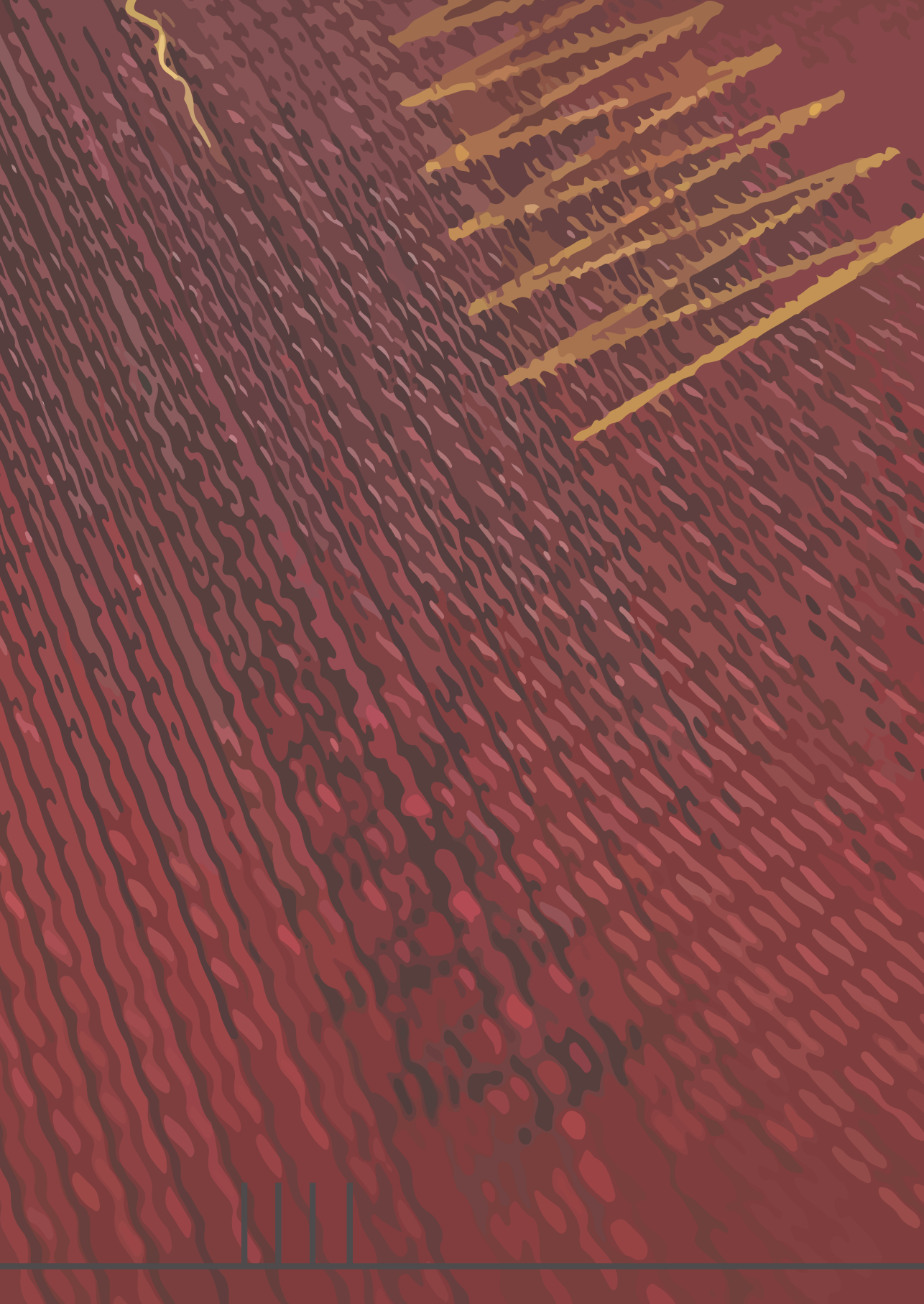


Figure S5. Splicing analysis of ARSB exon 4 of patient #12 and parents in triplicate.

Supplemental Information

Supplemental Table 1. Primers used for the RT-PCR, RT-qPCR and sequence reactions.

| RT-qPCR | | |
|------------------------------|---------------------------|-------------------|
| Name | Sequence (5'-3') | Product size (nt) |
| GAPDH FW | ATGGGGAAGGTGAAGGTCG | 70 |
| GAPDH RV | TAAAAGCAGCCCTGGTGACC | |
| Exon 1 Fw | CAGACGACCTAGGCTGGAAC | 101 |
| Exon 1 Rv | GTAGTTGTCCAGGAGCACCC | |
| Exon 2 Fw | CCTGCCCCAGCTCCTAAAAG | 108 |
| Exon 2 Rv | GTATCAAATCCTCGGCGGGT | |
| Exon 3 Fw | ACGCTCTGAATGTCACACGA | 129 |
| Exon 3 Rv | GTGGATGGTTAGTTATGAGGGCT | |
| Exon 4 Fw | TCTCCAGTCTGTGCATGAGC | 106 |
| Exon 4 Rv | AAGGGACACCATTCTGCAT | |
| Exon 5 Fw | GGCCCCCTTCGAGGAAGAAAA | 118 |
| Exon 5 Rv | GAGATGTGGATGAGCTCCCG | |
| Exon 6 Fw | AAGCCCATCCCCAGAATTG | 64 |
| Exon 6 Rv | CGGTGAAGAGTCCACGAAGT | |
| Exon 7 Fw | ATGGCTCCAGCAAAGGATGA | 102 |
| Exon 7 Rv | GCCCGTGAGGAGTTTCCAAT | |
| Exon 8 Fw | CACCAACCAAGACCCTCTGG | 118 |
| Exon 8 Rv | TGGTAGAACTGTAGGCGGGA | |
| RT-PCR | | |
| Name | Sequence (5'-3') | Product size (nt) |
| Exon 1-3 Fw | GGGTGCTCCTGGACAACACTAC | 379 |
| Exon 1-3 Rv | CCTGTTGCAACTTCTTCGCC | |
| Exon 2-4 Fw | ATGGCACCTGGGAATGTACC | 444 |
| Exon 2-4 Rv | GTGTTGTTCCAGAGCCCACT | |
| Exon 3-5 Fw | ACGCTCTGAATGTCACACGA | 514 |
| Exon 3-5 Rv | GTTGGCAGCCAGTCAGAGAT | |
| Exon 4-6 Fw | AAAAAGCAGTGGGCTCTGGA | 361 |
| Exon 4-6 Rv | CGGTGAAGAGTCCACGAAGT | |
| Exon 5-7 Fw | CAGAAGGGCGTGAAGAACCG | 314 |
| Exon 5-7 Rv | CCCGTGAGGAGTTTCCAATTTC | |
| Exon 6-8 Fw | ACTTCGTGGACTCTTCACCG | 348 |
| Exon 6-8 Rv | AGTACACGGGGACTGAGTGT | |
| M13 sequence primers (5'-3') | | |
| M13 Rv | CAGGAAACAGCTATGAC | |
| M13 Fw | GTAAAACGACGGCCAG | |
| PMO AONs (5'-3') | | |
| M1a | TTCTCAGCATCTAGAAGAAGGTATG | |
| M1b | ACGATTTCCAAATCTGAGTCTGGGT | |
| M2a | AGCAGCAAATCTTTAGCACCAAAGA | |
| M2b | CCTTATCCTCACCCAAGCAGGACTT | |



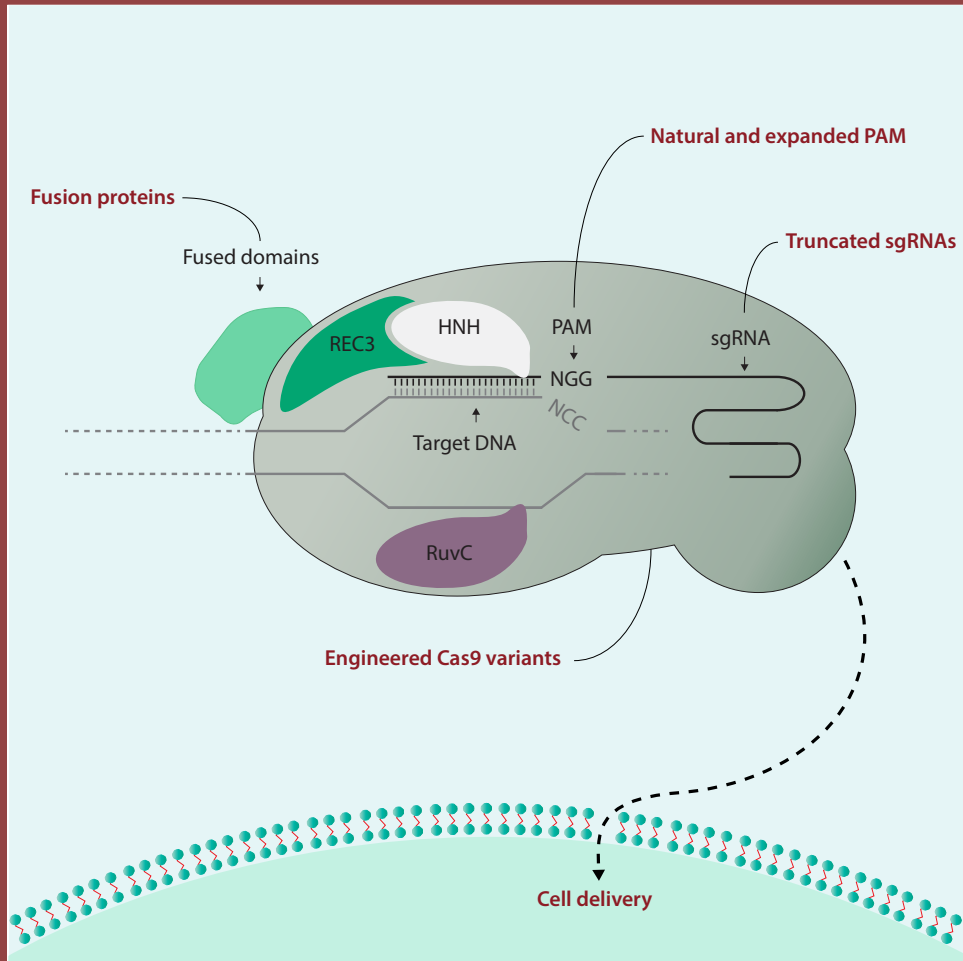
CHAPTER 4

SHARPENING THE MOLECULAR SCISSORS: ADVANCES IN GENE-EDITING TECHNOLOGY

Mike Broeders*, Pablo Herrero-Hernandez*, Martijn P.T. Ernst, Ans T. van der Ploeg,
W.W.M. Pim Pijnappel

*These authors contributed equally to this work
iScience. 2020 Jan 24;23(1):100789. doi: 10.1016/j.isci.2019.100789.
Epub 2019 Dec 19.

Graphical abstract



Keywords: gene editing; gene therapy; genome modification;
CRISPR-Cas; TALEN; ZFN

Sharpening the Molecular Scissors: Advances in Gene-Editing Technology

Mike Broeders,^{1,2,3,4} Pablo Herrero-Hernandez,^{1,2,3,4} Martijn P.T. Ernst,^{1,2,3} Ans T. van der Ploeg,^{1,3} and W.W.M. Pim Pijnappel^{1,2,3,*}

¹Department of Pediatrics, Erasmus University Medical Center, 3015 GD Rotterdam, Netherlands

²Department of Clinical Genetics, Erasmus University Medical Center, 3015 GD Rotterdam, Netherlands

³Center for Lysosomal and Metabolic Diseases, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

⁴These authors contributed equally

*Correspondence: w.pijnappel@erasmusmc.nl

<https://doi.org/10.1016/j.isci.2019.100789>

The ability to precisely modify human genes has been made possible by the development of tools such as meganucleases, zinc finger nucleases, TALENs, and CRISPR/Cas. These now make it possible to generate targeted deletions, insertions, gene knock outs, and point variants; to modulate gene expression by targeting transcription factors or epigenetic machineries to DNA; or to target and modify RNA. Endogenous repair mechanisms are used to make the modifications required in DNA; they include non-homologous end joining, homology-directed repair, homology-independent targeted integration, microhomology-mediated end joining, base-excision repair, and mismatch repair. Off-target effects can be monitored using *in silico* prediction and sequencing and minimized using Cas proteins with higher accuracy, such as high-fidelity Cas9, enhanced-specificity Cas9, and hyperaccurate Cas9. Alternatives to Cas9 have been identified, including Cpf1, Cas12a, Cas12b, and smaller Cas9 orthologs such as CjCas9. Delivery of gene-editing components is performed *ex vivo* using standard techniques or *in vivo* using AAV, lipid nanoparticles, or cell-penetrating peptides. Clinical development of gene-editing technology is progressing in several fields, including immunotherapy in cancer treatment, antiviral therapy for HIV infection, and treatment of genetic disorders such as β -thalassemia, sickle cell disease, lysosomal storage disorders, and retinal dystrophy. Here we review these technological advances and the challenges to their clinical implementation.

INTRODUCTION

In recent years, various platforms for genetically engineering somatic and pluripotent stem cells have been developed. They include zinc finger nucleases (ZFN), transcription activator-like effector nucleases (TALENs), meganucleases (MNs), and clustered regularly interspaced short palindromic repeats (CRISPR) in combination with CRISPR-associated protein (CRISPR/Cas). From agriculture to biomedical science, these platforms are being explored in various fields and, more recently, in the first clinical trials (Barrangou and Doudna, 2016; Naldini, 2015; Rodriguez-Rodriguez et al., 2019; Zhan et al., 2019).

Several developments are taking place in parallel. First, technological improvements and variations on gene-editing strategies are being reported with unsurpassed speed. Second, many possible applications are being developed to address a wide variety of biomedical questions. Third, the first platforms for gene editing are entering the clinical testing stage. Here, we follow these themes to present an overview of these recent developments, focusing on the likely clinical implementation of gene-editing strategies as well as discussing recent technological advances, and aspects such as safety, efficacy, and delivery that are relevant to clinical implementation. We provide a short overview of the progress gene editing is making toward applications in the clinic.

Technological Advances

Basics of CRISPR/Cas

Genome editing depends on the ability to generate specific pre-designed alterations in the genome. Inducing double-strand breaks (DSBs), single-strand breaks (SSBs) (also termed “nicks”), or specific base changes result in the activation of endogenous repair mechanisms that can be used to alter the genome. MNs, ZFNs, and TALENs were the first tools for genome editing in mammalian cells. MNs are naturally occurring endonucleases and can be re-targeted to new sites. Gene editing with ZFNs and TALENs results from the fusion of the FokI nuclease domain to the DNA-binding modules of zinc finger proteins (in the case of ZFNs) or transcription activator-like effector proteins (TALEs) (in the case of TALENs). However, the design and construction of these gene-editing tools can be labor-intensive, and their efficiencies for performing gene-editing varies. The discovery of CRISPR/Cas as a new gene-editing platform introduced a fast, cheap, and relatively efficient genome-editing method that revolutionized genome engineering.

The mechanism of CRISPR/Cas is based on its role in adaptive immunity in prokaryotes, in which short stretches of invading foreign nucleic acids, so-called protospacers, are incorporated into the CRISPR locus of the bacterial or archaeal genome. After acquisition, CRISPR RNA (crRNA) is generated from the protospacers at the CRISPR locus. This crRNA can bind to complementary foreign nucleic acid and directs the Cas protein to recognize invading sequences. A second RNA known as the trans-activating CRISPR RNA (tracrRNA) is transcribed from a genomic locus upstream of the CRISPR locus and forms a complex with the crRNA. The crRNA:tracrRNA complex associates with a Cas protein (the nuclease) and creates an active ribonucleoprotein (RNP) complex that targets foreign nucleic acids for degradation (Mojica and Montoliu, 2016; Maeder and Gersbach, 2016).

For genome editing, the tracrRNA and crRNA are fused into a single guide RNA (sgRNA), which binds complementarily to a DNA target and guides the Cas protein to the desired target site, creating a DSB (Jinek et al., 2012). The target sequence is based on the presence of a protospacer adjacent motif (PAM), which is an absolute prerequisite for Cas protein to induce a DSB. The first publications on CRISPR arrays date back as early as 1987 (Ishino et al., 1987), Cas genes were discovered in 2002 (Jansen et al., 2002), and CRISPR/Cas was shown to cleave bacteriophage and plasmid DNA *in vivo* at specific sites in 2010 (Garneau et al., 2010). However, it was not until 2012 that two groups (Gasiunas et al., 2012; Jinek et al., 2012) adapted this system into a gene-editing tool. A review by Fernández and colleagues provides an extensive history of genome editing tools (Fernandez et al., 2017).

Its accessibility and relatively low costs have brought CRISPR/Cas a large number of applications in research worldwide and have catalyzed further research on understanding its mechanism of action, improving its functional capacities, and extending its biomedical applications. The CRISPR/Cas9 system originally applied has downsides that include a lower specificity than other gene-editing tools and a relatively large cargo size that hampers delivery to cells via vectors with limited cargo-size capacity (Fernandez et al., 2017; Guha and Edgell, 2017; Guha et al., 2017; Gupta and Musunuru, 2014; Zych et al., 2018). However, as we discuss below, innovative research has produced many adaptations to the original system that enhance its versatility and improve properties such as specificity and efficacy (Zhang et al., 2016; Wu et al., 2018).

Versatility of CRISPR/Cas-Mediated Gene Editing

As nucleases merely induce breaks in the DNA, the introduction of specific alterations in the genetic code requires the exploitation of distinct DNA-repair mechanisms. Intelligent engineering of the original genome-editing tools has broadened the scope of this toolkit, making it possible to achieve a great number of applications. As a result, CRISPR/Cas can be applied to interfere at multiple steps of gene-expression processes and can target processes at genomic and transcriptomic levels. Approaches include gene knock-out, precise correction of disease-associated variants, insertion of a cDNA in a safe harbor (a location in the genome—such as the AAVS1 locus—where there is no risk of insertional mutagenesis) (van der Wal et al., 2018; Sadelain et al., 2011); and manipulation of gene-expression regulatory elements such as promoter activity (Baliou et al., 2018) or splicing (Bergsma et al., 2018; Smith et al., 2018).

Targeted Deletion and Gene Knock-Out. The DNA repair pathway that is most commonly used to create deletions and specific gene knock-outs is non-homologous end joining (NHEJ). NHEJ repairs blunt or incompatible double strands, either by direct ligation or through mediation by microhomology of 5-25 nucleotides that flank the DSB to facilitate end joining (Chang et al., 2017a). NHEJ represents the major double-strand break repair system in mammalian cells and is active throughout the cell cycle (Chang et al., 2017a; Ranjha et al., 2018). Repair can be precise, leaving the target site intact for recleaving by the Cas nuclease. However, NHEJ-mediated DNA repair is rather error prone, as it can introduce random insertions or deletions of base pairs (indels) that will destroy the target site. These errors can result in a frameshift that inactivates gene products through mRNA decay (Ranjha et al., 2018). This effect is exploited in gene editing to create targeted gene knock-outs. Examples of clinical applications include knocking out disease-promoting genes such as oncogenes or restoring a reading frame by interfering with splice sites such as in Duchenne muscular dystrophy (Amoasii et al., 2018; Tabebordbar et al., 2016; Nelson et al., 2016).

Targeted Gene Editing and Knock-In. Traditionally, the homology-directed repair (HDR) pathway is used to achieve a precise knock-in. HDR recognizes DSBs and utilizes a homologous template—which,

under physiological conditions, is the sister chromatid—to repair the defect (Ranjha et al., 2018). For precision gene editing, this pathway is exploited by introducing a donor template that is used in the repair of the DSB rather than the sister chromatid. This donor template has the desired alterations, which are flanked by 3' and 5' homology arms on both sites of the DSB. In this manner, it is possible to accurately correct a point variant (as in cystic fibrosis) (Schwank et al., 2013); a tandem repeat (as in Fragile X syndrome) (Xie et al., 2016); large inserts or deletions (as in Duchenne dystrophy) (Li et al., 2015); or other genetic defects.

Due to the use of a donor template, HDR is a very precise repair pathway that is less error-prone than NHEJ. However, HDR is less active in cells and therefore much less efficient in genome editing than NHEJ, which limits its clinical potential. To favor its activation over NHEJ induction, its efficiency would have to be increased. This has cast new light on research into the regulation of DNA repair pathways (Mateos-Gomez et al., 2017; Schimmel et al., 2017; Zelensky et al., 2017). Several strategies using small molecules to inhibit NHEJ have been followed with varying success (Chu et al., 2015; Maruyama et al., 2015; Song et al., 2016; Yu et al., 2015; Pinder et al., 2015). Recently, cold shock was found to increase HDR in cells *in vitro* (Guo et al., 2018a). As HDR is highly suppressed in the G1 phase and its activity is limited to the S and G2/M phases of the cell cycle, it is restricted to dividing cells (Ranjha et al., 2018), which, since many cells in the human body are non-dividing, greatly restricts the clinical potential of HDR-based gene editing. Recently, Orthwein et al. have shown reactivation of HDR in the G1 phase via the PALB2-BRCA1/CUL3/Keap1 pathway. Although this reactivation might allow therapeutic targeting of non-dividing cell types, the limited efficiency seen to date would indicate the need for further research to demonstrate that therapeutically relevant HDR levels can be achieved through G1 reactivation (Orthwein et al., 2015).

More recently, other pathways have been used to achieve targeted knock-in. Through a technique named homology-independent targeted integration (HITI), NHEJ can also be harnessed to create knock-ins (Suzuki and Izpisua Belmonte, 2018). After a donor vector containing the desired transgene flanked by a CRISPR target site has been introduced into the cell, the donor vector and the genomic target are cleaved by Cas9, generating blunt ends on both genomic target and donor vector. These blunt ends are utilized to induce NHEJ-mediated end-to-end ligation, allowing the integration of the donor sequence into the target location (Sawatsubashi et al., 2018). As the activity of NHEJ is higher than that of HDR, this approach can be more effective and can be used in non-dividing cell types, because NHEJ remains active in all phases of the cell cycle (Suzuki and Izpisua Belmonte, 2018).

Several disadvantages limit the clinical potential of NHEJ-mediated knock-in. First, the transgene is inserted in a random direction. Second, due to potential off-target effects of CRISPR/Cas9 system, the use of a donor template may give rise to nonspecific insertions of this template. Last, NHEJ can introduce random indels, possibly disrupting the target location. However, clever vector design and target site selection can minimize the number of non-specific insertions and insertions in the wrong orientation (Suzuki and Izpisua Belmonte, 2018; Sawatsubashi et al., 2018).

Besides the classical NHEJ and HDR pathways, another DNA repair pathway has emerged in the last two decades: microhomology-mediated end joining (MMEJ) (also known as alternative end joining (a-EJ)) (Frit et al., 2014). In MMEJ, microhomologies exist upstream and downstream of the DSB site on the two DNA strands. DSBs with microhomologies can result in annealing of the microhomologies and subsequent repair by MMEJ, which can result in short deletions (Kim et al., 2018b). This pathway has been used to integrate a gene of interest by using a precise integration into target chromosome (PITCh) system in cultured cells, zebrafish, silkworm, and frogs (Sakuma et al., 2016; Nakade et al., 2014; Hisano et al., 2015). With a reported knock-in efficiency that is 2.5 times higher than that of HR-assisted gene knock-in (Nakade et al., 2014), and with activity during all phases of the cell cycle (Taleei and Nikjoo, 2013; Truong et al., 2013), this pathway opens up the possibility of more efficiently targeted gene knock-in in different phases of the cell cycle. Further work will be necessary to demonstrate the applicability and the feasibility of MMEJ-mediated knock-ins for precision genome editing.

Base Editing in DNA. Incorrectly repaired DSBs and DSB at off-target sites are potentially pathogenic. Attempts to circumvent the generation of DSBs, and thereby to lower the risks associated with gene editing, have resulted in the development of the base-editing technique (Hess et al., 2017; Molla and Yang, 2019), which exploits the natural function of cytidine deaminases to convert cytidine to uridine in DNA. Eventually, the uridine is converted to thymidine by DNA duplication or via the DNA repair mechanisms

base excision repair (BER) or mismatch repair (MMR) (Hess et al., 2017). Two different base-editing systems, BE (Komor et al., 2016) and target-AID (Nishida et al., 2016), have been developed by coupling cytidine deaminases to a catalytically deficient Cas9 protein (dCas9). Recently, a system with an incorporated adenosine deaminase was also developed. This enzyme hydrolyses adenosine into inosine, the base that pairs with cytidine. Inosine is thus replicated as guanine, thereby rendering an A to G change by DNA duplication (Gaudelli et al., 2017). These methods are clinically very interesting, as they open up opportunities for correcting monogenetic diseases in ways that reduce the risk of off-target effects. However, as most diseases are caused by various different disease-associated variants, this approach has to be tailored to each unique variant (Lessard et al., 2017). Another application of base editing is the introduction of a transcription-termination sequence to disrupt a gene in a highly specific manner (Kuscu et al., 2017; Billon et al., 2017).

Prime Editing. Recently, Anzalone et al. developed prime editing, a novel strategy for introducing deletions, inserting new genetic content or generating any of the 12 possible base-to-base conversions (Anzalone et al., 2019). For prime editing nicking Cas9 variant is fused to a reverse transcriptase, which, together with a prime editing extended guide RNA (pegRNA), makes gene editing possible. Prime editing offers a new range of possibilities in genome editing, with greater flexibility than the base editors, a greater reported efficiency than HDR, and no introduction of a DSB. Further research will determine the advantages and limitations of this promising novel concept.

Transient Modifications. All previous modifications result in permanent modifications at the genomic level. As an alternative approach, gene editing has been used to introduce transient modifications.

CRISPR interference (CRISPRi) was developed to inhibit gene transcription. Depending on the nature of the targeted locus, this approach is designed to interfere with transcription initiation or elongation (Bikard et al., 2013; Qi et al., 2013). The fusion of dCas9 to transcriptional activators or repressors has further sophisticated the gene-editing toolkit for regulating transcription by CRISPR/Cas systems (Mahas et al., 2018).

The CRISPR/Cas system has also been engineered to modify epigenetic states by coupling dCas9 to several epigenetic modifiers such as P300 Core (Hilton et al., 2015), KRAB (Thakore et al., 2015), LSD1, Tet1, and Dnmt3 (Liao et al., 2017). Possible applications of this technique include the reversal of pathological epigenetic changes in conditions such as Fragile X syndrome, caused by silencing of the *FMR1* gene, which is associated with hypermethylation of the CGG expansion in the region encoding the *FMR1* 5' UTR. Recently, Liu and colleagues showed that by using dCas9-Tet1 the CGG expansion can be demethylated, leading to reactivation of *FMR1* (Liu et al., 2018b). Future research is needed to test whether this is a valuable approach to treating fragile X syndrome and other epigenetic diseases.

Various studies over the years have shown that it is possible to target RNA with Cas13, an analogous member of the Cas family that is also referred to as C2c2 (O'Connell et al., 2014; East-Seletsky et al., 2016; Abudayyeh et al., 2016; Batra et al., 2017). One of these studies described the development of the RNA-editing platform known as RNA Editing for Programmable A to I Replacement (REPAIR) (Cox et al., 2017). For REPAIR a catalytically inactive Cas13 nuclease is coupled to a modified deaminase domain of adenosine deaminase that acts on RNA type 2 (ADAR2), which swaps A bases for I in RNA sequences. Additional modifications in the Cas13-ADAR2 fusion increased the targeted specificity of the REPAIR system almost a thousand-fold. In another study, protein engineering and characterization of different Cas13d orthologs generated a ribonuclease effector called CasRx (Konermann et al., 2018). The CasRx system has been shown to be highly effective in transcript knockdown or repression and in splice isoform manipulation.

Potentially, transcript editing could play a role in the therapy of diseases that cause transient changes in gene expression, such as local inflammation. As no definitive changes are made in the genome, but rather at transcriptional level, these approaches might represent a safer therapeutic strategy. However, the transient nature of these modifications will require repeated administration of the therapeutic agents.

Challenges in Bringing Genome Editing to the Clinic: Safety, Efficacy, and Delivery

Many conditions have to be met before genome-editing techniques can be considered for clinical development. Their efficiency and delivery must be great enough to attain a clinically significant result. Adverse events produced by permanent unintended variants should be minimized. Below, we highlight the

targeting specificities and efficacies of natural and engineered variants and the broad diversity of strategies for delivering gene-editing tools.

Specificity

Mechanisms of Undesired On-Target and Off-Target Effects. The development of precise gene-editing tools to treat genetic disorders in the clinic requires careful consideration of the medical implications of permanent modifications in the genome. Nucleases should provide sufficient targeted specificity to prevent potentially detrimental effects derived from DSBs and the subsequent DNA repair mechanism. These effects can include undesired small insertions and deletions, point variants, aberrant chromosomal rearrangements, and large deletions in edited cells (Kosicki and Bradley, 2018; Hsu et al., 2013); they occur at or around the targeted location ("on-target"), as well as at more distant locations in the genome ("off-target").

Influence of p53. p53 has been reported to cause a targeting selection bias in engineered cells: Cas9-gene targeting showed higher editing efficiency in cells with an altered p53 gene (Haapaniemi et al., 2018; Ihry et al., 2018), and transient p53 inhibition was shown to increase editing efficiency in human pluripotent stem cells, retinal pigment epithelial cells, and, more recently, in hematopoietic stem and progenitor cells (HSPCs) (Schirolli et al., 2019). However, due to the role of p53 in multiple DNA damage-response mechanisms, transient p53 inhibition may also leave cells more vulnerable to off-target mutagenesis. This idea has been challenged by others, who analyzed a large number of datasets derived from CRISPR screens (Brown et al., 2019; Ihry et al., 2019; Mair et al., 2019). These studies failed to show differences between p53-deficient and wild-type cells with respect to the enrichment of essential genes, thereby arguing against selection for clones that have a defective DNA damage-response pathway in CRISPR screens. Recently, guidelines were proposed for the performance of CRISPR screens with respect to monitoring and ensuring the quality of the screening performance (Brown et al., 2019).

R-Loop Formation. In addition to the risk of cleavage at an undesired location, the formation of R-loop-linked mutagenesis poses a risk. CRISPR/Cas cleavage is initiated by forming an R-loop, i.e., RNA-guided DNA unwinding to form an RNA-DNA hybrid with a displaced DNA strand inside the Cas protein (Jiang et al., 2016; Szczelkun et al., 2014). Although this R-loop is vital for stable binding and cleavage by the Cas protein, recent evidence in yeast indicates that the R-loop itself promotes mutagenesis on both on- and off-target sites (Laughery et al., 2019). This implies that undesirable mutagenesis may occur in many applications based on dCas9-fusions that were previously presumed to be safe, such as epigenome editing, transcriptional repression, and activation.

Effects on Transcription and Translation. Another risk is imposed by the introduction of indels via NHEJ, which can have unanticipated impacts on the regulation of RNA products and the translation of the protein it encodes. This includes promotion of internal ribosomal entry and alternative spliced mRNAs that lead to alternative products with a gain-of-function or a partially functional protein (Thomas et al., 2019; Mou et al., 2017; Tuladhar et al., 2019). mRNA decay can also trigger transcriptional adaptation. Via this genetic compensation mechanism, the expression of targeted or related genes can be upregulated independently or through protein feedback loops (El-Brolosy et al., 2019). Because the underlying mechanisms involved in splicing regulation and genetic compensation remain poorly understood, it remains a challenge to anticipate the translational and transcriptional changes induced by NHEJ-mediated indels.

Off-Target Effects of Base Editing. Because base editors do not produce DSBs in the genome, they do not pose risks of unintended damage triggered by DSBs. However, recent studies have reported the production of other types of off-target effects by adenine and cytosine base editors. In mouse embryos, variants using the APOBEC1 cytosine deaminase have been reported to generate multiple single-nucleotide variants (SNVs) at frequencies 20 times higher than the spontaneous mutation rate (Zuo et al., 2019). Other studies observed similar results, not only with the cytosine deaminases but also with the adenine editors based on the TadA deaminase, albeit at much lower frequencies (Kim et al., 2019). To reduce the rate of mutagenesis, Kim and colleagues coupled TadA with engineered Cas9 variants, successfully reducing their off-target activity (Kim et al., 2019). In addition, tens of thousands of SNVs were recently identified in RNA transcripts induced by both adenine and cytosine editors (Grunewald et al., 2019a; Zhou et al., 2019). These studies raised concerns about potentially detrimental variations not only in the genome but also in the transcriptome. Appropriate monitoring strategies will therefore be needed to evaluate the DNA and RNA of

engineered cells that are created with base editors. To enable the safe use of base editors in the clinic, further research on novel variants or on recently engineered deaminases (Grunewald et al., 2019b) may reduce their off-target effects.

Monitoring Undesired On-Target and Off-Target Effects. Off-target mutations are more difficult to detect than on-target variants, as they can be present anywhere in the genome. The most common workflow applied to date is to design gRNA sequences with minimal predicted off-target effects and perform targeted sequencing to those sites with a high predicted value. Many tools are publicly available, the choice depending on the model organism of interest (Listgarten et al., 2018). A recent comprehensive study by Allen and colleagues investigated the outcome of more than 40,000 Cas9-edited DNA sequences using different cell types (Allen et al., 2018), showing how different cell types preferentially use specific repair mechanisms for certain DNA sequences. This research led to the development of a prediction tool (FORECasT) that could be used to predict off-target effects.

Common monitoring methods for detecting on-target mutations are based on polymerase chain reaction (PCR) (reviewed in Zischewski et al. [2017]). Each assay is designed according to the mutation introduced into the host genome, and the method selected must detect the difference in the heteroduplex DNA in order to determine the editing outcome. Mismatch cleavage assays, such as T7E1 or Surveyor, are commonly used to identify indels in bulk and single-cell preparations. Although these methods are simple and effective in detecting small indels, they underrate the detection of large deletions or insertions. Additional methods that may provide detailed information about introduced mutations are based on Sanger sequencing.

Considering our limited knowledge of the mechanisms involved in CRISPR/Cas off-target activity, unbiased methods should be included that do not rely on predictions. Although whole-genome sequencing offers an unbiased high-throughput assessment of unintended variants, it is expensive and only provides information on bulk genomes. When applied to cell clones that have been edited *ex vivo*, whole-genome sequencing provides a valuable option. However, when applied to cell populations without clonal expansion, whole-genome sequencing is less suitable, as it may fail to detect low-abundant events that might eventually lead to oncogenic transformation. Other methods for detecting off-target effects are based on the fact that nucleases generate breaks that are repaired by endogenous DNA repair mechanisms. These have been exploited to develop techniques for quantifying the DSBs that have occurred *in vitro* (BLESS, GUIDE-seq, Digenome-seq) (Zischewski et al., 2017).

The selection of the most suitable detection method will depend on the sensitivity, throughput, limitations, and cost of the technique, and ultimately on the editing strategy (*ex vivo* or *in vivo*) and delivery system used (viral or non-viral). A recent study by Akcakaya and colleagues describes a robust evaluation of off-targets using *in vitro* assessment of potential off-targets via CIRCLE-seq, followed by *in vivo* targeted deep sequencing of engineered organs (VIVO) (Akcakaya et al., 2018). In a more recent study, Wienert and colleagues developed an unbiased monitoring strategy (DISCOVER-Seq) to evaluate potential off-target DSBs in cells and tissues (Wienert et al., 2019). Because these approaches are based on pre-repair mechanisms, additional detection methods will be required to evaluate their editing outcome and cytotoxic effect. In human diseases, robust standardized assessment guidelines will therefore be essential for precise genome engineering.

Increasing Specificity. Research has made great advances in the characterization of novel CRISPR/Cas9 orthologs and homologs from various species and in the generation of engineered enhanced-specificity Cas enzymes and sgRNAs. The most commonly used CRISPR/Cas system in research exploits the Cas9 protein derived from *streptococcus pyogenes*. However, many other orthologs offer different targeting abilities and specificities (Karvelis et al., 2017). Other natural homologs discovered recently include Cas12a (Cpf1) and Cas12b (C2c1), which can also be engineered for enhanced DNA specificity (Wu et al., 2018).

Recently, many engineered forms of Cas9 proteins with improved and broad targeting specificities have been developed. For example, Cas9 nickases (Ran et al., 2013) or Fok I fused to dead Cas9 nucleases (Tsay et al., 2014) use two different sgRNAs to perform ssDNA breaks, which can significantly reduce the off-target effects of WT Cas9 variants. Cas9 nickases have been used to generate paired single-strand DNA breaks in donor plasmids and at genomic target sites, thereby increasing gene-targeting efficiency,

specificity, and fidelity in human cells when compared with nicking targeting DNA alone (Chen et al., 2017b).

More recently, new engineered Cas9 proteins with enhanced targeted specificity have raised the possibility of using smart-designed nucleases for precise genome editing. The high-fidelity Cas9 (Cas9-HF1) (Kleinstiver et al., 2016) was designed by altering the composition of four residues involved in non-specific interactions of Cas9 with its target DNA. These modifications reduced the generation of mismatches with minimal loss of on-target efficiency. Using GUIDE-seq to monitor the frequency of off-target effects showed that editing with Cas9-HF1 resulted in only one single mismatch compared with the 65 detected for WT Cas9 using eight different sgRNAs toward four human genes in U2OS cells. The enhanced specificity Cas9 (eSpCas9) (Slaymaker et al., 2016) was created by modifying positively charged residues involved in the unwinding of the non-complementary DNA strand. Weakening this interaction reduced Cas9 off-target activity 10 times compared with that of WT Cas9, while maintaining on-target efficiency in human embryonic kidney cells. The evolved Cas9 (evoCas9) was generated through directed evolution in yeast (Casini et al., 2018), which involved the screening of random variants to identify beneficial variants in the REC3 domain involved in the recognition of the sgRNA and DNA heteroduplex. The combination of four beneficial variants generated the evoCas9 nuclease: a high-fidelity Cas9 variant with minimal loss of on-target activity and a reported 98.7% reduction in off-target activity relative to that in WT Cas9 when analyzed via GUIDE-seq (Casini et al., 2018).

The lack of mechanistic insight regarding target recognition of Cas9-HF1 and eSpCas9 led to the development of the hyper-accurate Cas9 (HypaCas9) (Chen et al., 2017a). It was shown that raising the energy requirements for the conformational activation of the HNH domain, which acts as a Cas9 editing checkpoint, reduced off-target activity in these variants. HypaCas9 was designed by modifying four amino acids involved in this process. Genome-wide specificity of this variant was significantly better than that of the WT Cas9 and equivalent to that of the Cas9-HF1 and eSpCas9 nucleases. On-target activity was at least >70% that of WT Cas9 when tested in U2OS cells.

Other Cas engineering strategies focused on increasing the targeting capacities of CRISPR/Cas nucleases by altering their PAM recognition sites without losing on-target specificity (Figure 1) (Kleinstiver et al., 2015; Hu et al., 2018). Directed evolution was used to create different variants of Cas9 nucleases with altered PAM targeting properties; the VQR and VRER variants demonstrated enhanced specificity in human cells. Other variants include xCas9, which recognizes a broader range of PAM sequences than natural Cas9. Recently, Nishimasu and colleagues designed a SpCas9-NG nuclease with a targeting capacity similar to that of the xCas9 variant. However, comparative studies showed that the editing efficiency of the SpCas9-NG variant was greater than that of xCas9. To convert C-to-T bases at NG PAM sites in human cells, the authors further combined SpCas9-NG with a cytidine deaminase, thereby providing new engineered Cas9 variants with many capacities beyond improved specificity (Nishimasu et al., 2018). Other homologs, such as Cas12a, have also been engineered to act on different PAM sequences instead of acting on their natural binding sites (Gao et al., 2017).

In addition, sgRNAs have been altered to enhance their targeting specificities. By modifying their length, secondary structure, or chemical composition, several studies demonstrated that off-target mutagenesis can be reduced (Fu et al., 2014; Hendel et al., 2015; Kocak et al., 2019). However, these changes may also influence their on-target efficiency. Subsequent studies found that partial replacement of RNA nucleotides with DNA could further reduce the off-target activity while retaining on-target efficiency (Yin et al., 2018). The optimal targeting specificity of sgRNAs will ultimately depend on the *in silico* design. Many software programs are publicly available for this purpose. However, it is important to consider the genetic variations between different individuals, as it can substantially alter the on-target and off-targeting activities of precise genome engineering (Lessard et al., 2017).

Spatiotemporal Control of CRISPR/Cas Activity. The capacity of nucleases to generate permanent genome changes has prompted researchers to seek novel methods of exerting spatiotemporal control over CRISPR/Cas activity. Such control is important: when targeting DNA to reduce undesirable off-target effects, any intervention in a desired tissue or organ should be brief.

The spatial localization of precision gene-editing tools can be manipulated by using different delivery vehicles. For instance, different serotypes of viral particles such as lentivirus or adeno-associated virus (AAV)

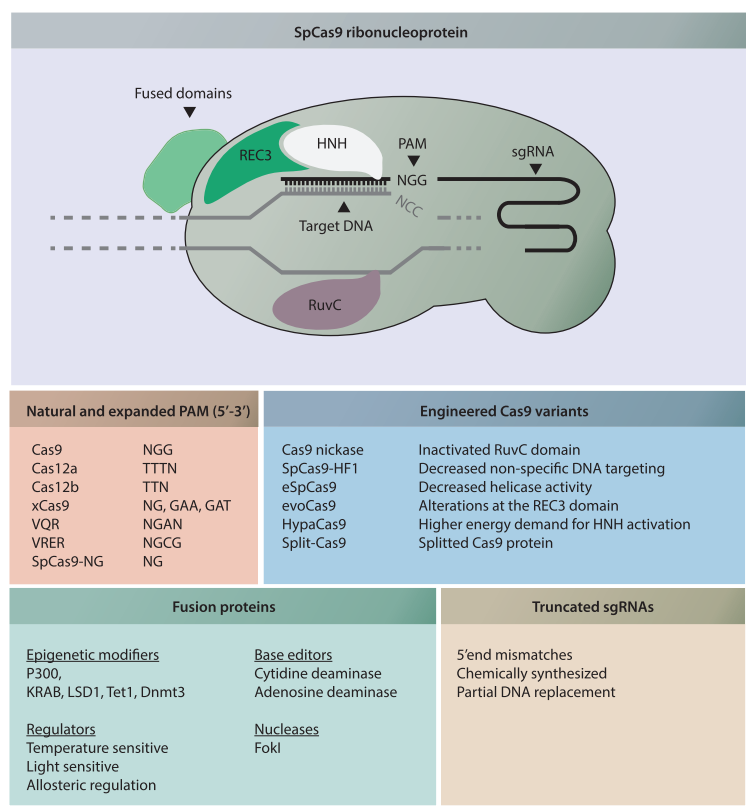


Figure 1. SpCas9 Ribonucleoprotein Variants with Altered Targeting Capacities
Engineered SpCas9 nuclease variants with potential therapeutic advantages. Variants with expanded PAM recognition sequences include natural Cas9 and its well-described homologs Cas12a and Cas12b. Designed engineered nucleases with improved specificity include the split-Cas9 variant, which was created to improve the packaging of Cas9 proteins in AAV vehicles. Cas9 nucleases have also been fused with multiple functional domains to allow for targeted epigenetic modifications, single nucleotide modifications, single-strand nicking activity, or temporal regulation of CRISPR/Cas activity. Truncated versions of sgRNAs have been successfully used to increase targeting specificity. SpCas9, *streptococcus pyogenes* Cas9.

have higher affinities for specific cell types. By using tissue-specific promoters, the expression of CRISPR/Cas components can be limited to target tissues or organs. Other organs can be targeted through local delivery with limited distribution into other tissues, such as the eye. However, for efficient targeting, diseases involving larger organs or tissues (such as muscle or skin) might require systemic administration.

Off-target effects are also magnified by extended gene-editing activity. Several approaches have therefore been developed to influence the spatiotemporal activity of CRISPR/Cas systems. Many natural Cas9 enzymes have been engineered into switchable Cas9 nucleases, either to be light sensitive, to be controlled by allosteric regulation, or to express specific localization signals (Richter et al., 2017). However, the clinical potential of these engineered forms can be limited, as some switchable nucleases respond to toxic regulators such as doxycycline or UV light. Nevertheless, the characterization of novel natural inhibitors for Cas9 and Cas12a variants holds great promise for the therapeutic regulation of CRISPR/Cas activity in a clinical setting (Pawluk et al., 2016; Watters et al., 2018; Jiang et al., 2019; Uribe et al., 2019). Another approach that holds great potential for controlling the CRISPR/Cas system and off-target activity is the use of anti-CRISPRs. Discovered in phages to overcome CRISPR immunity, these anti-CRISPRs can be used, repurposed, and structurally engineered to control the CRISPR/Cas9 system (Pawluk et al., 2018).

Immunogenicity and Safety. Gene targeting must be performed with minimal toxicity and immunogenicity. The immunogenic properties of gene-editing components have recently been described in several studies (Charlesworth et al., 2019; Kim et al., 2018a; Wagner et al., 2019; Chew et al., 2016). CRISPR/Cas systems are protein complexes derived from bacteria and archaea, some of which, such as *S. aureus* (SaCas9) or *S. pyogenes* (SpCas9) are common infectious agents in the human population. The risks associated with immune responses against Cas9 proteins (or other Cas homologs) and sgRNAs should be considered in clinical trials. A recent study by Charlesworth and colleagues on the presence of preexisting antibodies reported that 78% of donors involved in the study presented antibodies against SaCas9 and 58% presented antibodies against SpCas9. In addition, analysis of blood samples indicated that 78% had anti-SaCas9 T cells and 67% had anti-SpCas9 T cells (Charlesworth et al., 2019). The detection of anti-Cas9 T cells might pose a serious challenge to precision genomic engineering, as exposure to a cytotoxic T cell response might eliminate engineered cells. Because the assays used in this study were not intended to investigate cellular immune responses against CRISPR/Cas components, this issue should be addressed in future studies.

Gene-editing tools used *in vivo* will be exposed to the host immune system. Potential immune reactions will be influenced by the type of delivery vehicle used. For instance, viral-based vectors could lead to long-term expression of gene-editing tools. These could lead to sustained nuclease activity, which, depending on the variant used, might in turn lead to extensive damage to the human genome and a prolonged immune response. In addition, immune responses to viral-based vectors such as AAV are known for their capacity to neutralize the therapy, including an induction of a cytotoxic T cell response that eliminates cells that have been targeted by the AAV vector (Mingozzi and High, 2011, 2013). This is highly relevant, as AAV is a naturally occurring virus against which ~40% or an even higher percentage of people already carry antibodies to AAV capsid proteins.

To minimize the effects of potential immune reactions, researchers can follow various strategies. In one approach, gene-editing activity can be spatiotemporally controlled, which has great potential for clinical use (see above for the section on *Spatiotemporal control of CRISPR/Cas activity*). Gene-editing components can for instance be directed to immune-privileged organs such as the eye or to tolerogenic organs such as the liver. Strategies involving immune modulation might prevent side effects derived from bacterial Cas proteins or from antigens derived from the delivery vehicles. An appropriate evaluation of candidate approaches and potential detrimental effects will be essential for the safe use of precision genome engineering in the clinic.

Delivery Strategies

The development of efficient and safe delivery systems is one of the most challenging aspects of introducing precise genome editing to the clinic. The delivery approach to be used will be influenced by the type of gene-editing material. So far, CRISPR/Cas systems have been successfully administered in naked or encapsulated plasmid DNA, mRNA, or functional (ribonucleo)protein complexes both *in vivo* and *ex vivo* (Figure 2).

In Vivo Delivery. In many clinical and pre-clinical studies, adeno-associated viruses are commonly used delivery vehicles *in vivo*. Different AAV serotypes provide increased delivery efficiencies for specific cell types, thereby allowing tissue/organ targeting (Colella et al., 2018). *In vivo* delivery of gene-editing machinery using AAV has been successfully used in multiple animal models of metabolic diseases (Pankowicz et al., 2016; Yin et al., 2016; Villiger et al., 2018; Rossidis et al., 2018), human immunodeficiency virus (HIV) infection (Yin et al., 2017a), muscle dystrophies (Amoasii et al., 2018), brain disease (Nishiyama et al., 2017), retinal disorders (Suzuki et al., 2016; Huang et al., 2017; Maeder et al., 2019), degenerative disorders (Beyret et al., 2019; Santiago-Fernandez et al., 2019), and diabetes and kidney malignancies (Liao et al., 2017). More recently, it has been used to upregulate the expression of endogenous genes in haploinsufficiencies such as obesity (Matharu et al., 2018).

Due to its high efficiency, viral delivery through AAVs offers promising results for precision genome-editing medicine. However, AAVs elicit immune responses that may limit the therapeutic potential of genome-engineering tools. If the gene-editing strategy should be administered repeatedly over time, this is especially relevant, as patients will develop antibodies against the AAV virus after the first administration, precluding any subsequent treatment using AAV as delivery vehicle. Furthermore, as a significant proportion of the

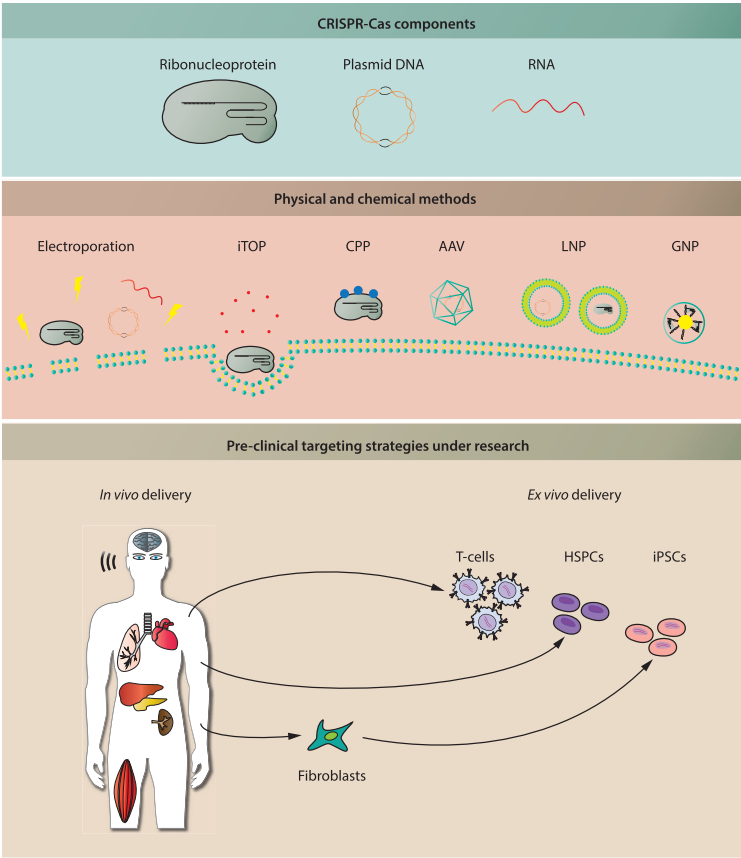


Figure 2. Delivery Strategies Used in Preclinical Studies

CRISPR/Cas gene-editing tools have been delivered as ribonucleoprotein complexes, plasmid DNA, or RNA. They can be delivered as naked components using chemical or physical methods or in delivery vehicles including virus and/or different types of nanocomplexes. There are two main strategies for delivering precision gene-editing platforms. In multiple organs and tissues in animal models, *in vivo* delivery strategies have been successfully used, some of which have been clinically approved for application in humans. *Ex vivo* delivery strategies are being extensively used in T cell engineering, hematopoietic stem-cell gene editing, and iPSCs modeling. iTOP, induced transduction by osmocytosis and propanebetaine; CPP, cell-penetrating peptide; LNP, lipid nanoparticle; GNP, gold nanoparticle; HSPCs, hematopoietic stem and progenitor cells; iPSCs, induced pluripotent stem cells.

population has pre-existing antibodies against the AAV virus, it is not eligible for AAV-based treatment (Louis Jeune et al., 2013). However, some of these limitations may be overcome by combined immunosuppressive therapies. AAV delivery can also drive long-term transgene expression and might result in low incidence of transgene integration. This promotes long-lasting gene editing in tissues and thus poses a high risk of off-target events. Methods for inactivating Cas activity would be required to prevent the long-lasting introduction of double-strand breaks *in vivo* and to reduce the risks of chromosomal abnormalities.

The large size of Cas proteins hampers their packaging into AAV plasmids. To circumvent this, researchers have successfully treated eye diseases using smaller Cas9 orthologs derived from *campylobacter jejuni* (CjCas9) (Kim et al., 2017). The study in question described how researchers were able to package the DNA sequence of the CjCas9, sgRNAs, and a donor template in a single AAV vector. Others have designed a flexible AAV-split-Cas9 platform to compact the Cas9 DNA sequence into AAV vectors (Chew et al., 2016). By manipulating the WT protein sequence of the Cas9 nuclease into fused functional domains, it was possible to shorten the newly engineered Cas9 DNA sequence by more than 2 kb.

Over recent years, different types of nanomaterial have been developed with encouraging results in multiple diseases (reviewed in Li et al. [2018]). Unlike AAVs, lipid-based nanoparticles are able to transfer Cas9 plasmids or proteins without the risk of genomic integration; some of them have received FDA approval for therapeutic use (Glass et al., 2018). In mice, lipid nanoparticles are currently used to deliver Cas9 components locally to the brain (Wang et al., 2016) or inner ear (Gao et al., 2018) and systemically to the liver (Yin et al., 2017b; Finn et al., 2018). However, for the delivery of Cas9 and a donor template for HDR, both components must be encapsulated in distinct lipid nanoparticles, which might affect its editing efficiency *in vivo*. More recently, newly developed gold nanoparticles have been successfully used in rodent models to simultaneously deliver Cas9 ribonucleoproteins and donor templates to treat Duchenne muscular dystrophy (Lee et al., 2017) and fragile X syndrome (Lee et al., 2018). Park and colleagues have shown successful genome editing of post-mitotic neurons in different Alzheimer disease mouse models using amphiphilic nanocomplexes generated using the R7L10 peptide together with the Cas9 nuclease and the sgRNAs (Park et al., 2019a).

Ex Vivo Delivery. Many vehicles are being exploited to deliver precision gene editing *ex vivo*. Depending on the type of cell or stem cell used for *ex vivo* editing, the gene-editing machinery may be delivered via electroporation, microinjection, chemical methods such as cell-penetrating peptides and nanoparticles (Wu et al., 2018), and virus-based vehicles such as AAVs or lentiviruses (Figure 2). The efficacy with which various types of immune cells can be engineered *ex vivo* has been investigated in multiple preclinical models, greatly stimulating research in several fields, particularly hematology and cancer therapeutics (Bak et al., 2018; Huang et al., 2018). Pluripotent stem cells such as induced pluripotent stem cells (iPSCs) have also been used extensively for genome engineering *ex vivo* for cell-based regenerative approaches and for disease modeling (Jang and Ye, 2016). iPSCs offer great potential for disease modeling, as they can be differentiated into any cell type that is relevant for the disease, such as cardiomyocytes (Brandao et al., 2017; Devalla and Passier, 2018), skeletal muscle cells (van der Wal et al., 2018; Magli and Perlingeiro, 2017; Chal and Pourquie, 2017), neuronal cells (Bordoni et al., 2018; Compagnucci et al., 2014), hepatocytes (Fiorotto et al., 2019; Hannoun et al., 2016), and many other cell types. Isogenic controls that correct for genetic background effects can be generated in iPSCs using CRISPR/Cas; this correction for background effects is considered important due to the large variation in many parameters among individuals. The risk of acquiring variants during the reprogramming of somatic cells into iPSCs has imposed the need for caution in the use of iPSC-derived cells for cell-based therapy. In addition, the development of affordable clinical treatment is inhibited by the cost associated with the quality control of patient-specific iPSCs. As a possible solution, the generation of iPSC-cell banks covering the majority of HLA isotypes known globally is currently in progress. This would make a validated iPSC line available for cell-based therapy for almost each individual patient without the need to generate patient-specific iPSCs (Ben Jehuda et al., 2018; Mandai et al., 2017).

Preclinical Studies and Clinical Trials

Therapeutic gene editing by ZFN, TALEN, or CRISPR/Cas9 platforms is already being explored in a number of clinical trials registered at clinicaltrials.gov. As we discuss briefly below, these target cancer, genetic disorders, and HIV/AIDS (Figure 3).

Cancer

In HPV-related cervical cancer, CRISPR/CAS9, ZFNs, or TALENs are applied topically to precancerous lesions in order to disrupt viral oncogenes *E6* or *E7* *in vivo* (Ding et al., 2014; Hu et al., 2014, 2015; Kennedy et al., 2014; Shankar et al., 2017; Zhen et al., 2014). In addition to induction of (viral) oncogene knock-out, gene editing for the treatment of various types of cancer is also being investigated in cellular immunotherapies or adoptive cell therapy (ACT). To overcome evasion of immune surveillance by cancer cells, *PD-1*, an immune checkpoint molecule, was knocked out *ex vivo* in autologous T cells using CRISPR/Cas9 before re-infusion into patients (Beane et al., 2015; Menger et al., 2016; Su et al., 2016; Zhao et al., 2018). *PD-1* knock-out have been deployed to improve redirected T cells (Guo et al., 2018b; Hu et al., 2019a, 2019b; Lu et al., 2019; Ouchi et al., 2018; Rupp et al., 2017; Su et al., 2017), which were genetically altered to express a transgenic T cell receptor (tTCR) or a chimeric antigen receptor (CAR) targeted at a disease-associated antigen. To enable ACT for T cell malignancies expressing CD7, CRISPR/Cas9 has also been used *ex vivo* to knock out the *CD7* gene in anti-CD7 CAR-T cells, preventing these CAR-T cells from killing each other (Cooper et al., 2018; Gomes-Silva et al., 2017). Finally, to circumvent the necessity for custom-made autologous therapy for each patient, *ex vivo* gene editing by CRISPR/CAS9 or TALENs in CAR-T cells manufactured from

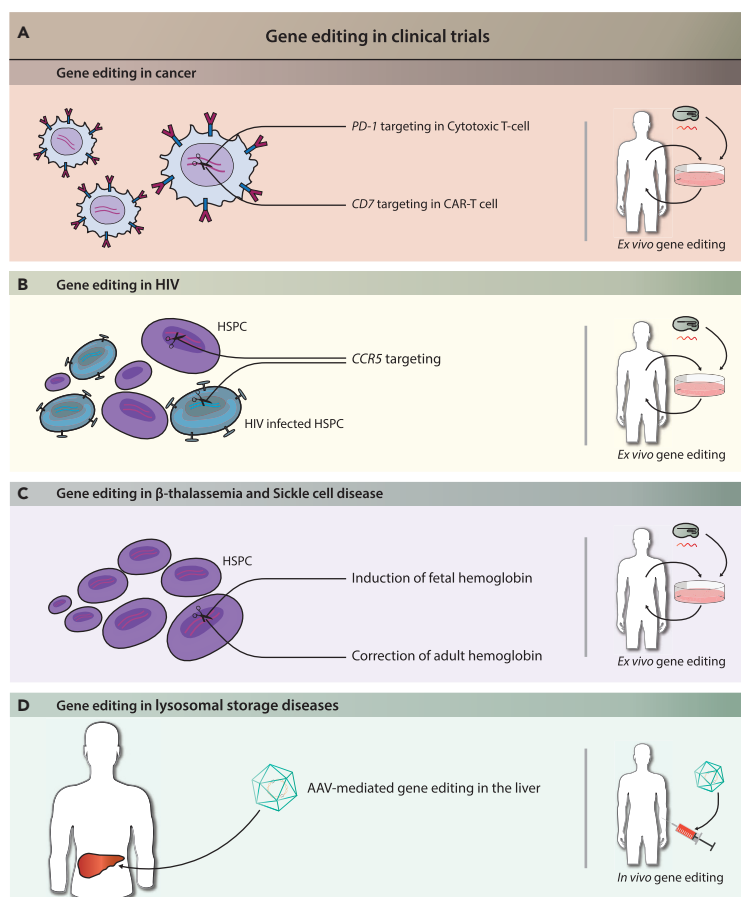


Figure 3. Examples of Gene-Editing Strategies in Current Clinical Trials

(A) In cancer, gene-editing targets *PD-1* or *CD7* in T cells to enhance immune responses.

(B) In patients with HIV, *CCR5* is targeted in HSPCs to prevent HIV entry.

(C) In β -thalassemia and sickle cell disease, induction of fetal hemoglobin or correction of adult hemoglobin in HSPCs is used.

(D) AAV-mediated gene editing in the liver provides circulating enzymes in lysosomal storage diseases.

donor-derived T cells is being applied to generate universal CAR-T cells (Torikai and Cooper, 2016; Yang et al., 2015). This involves the disruption of the endogenous TCR to prevent graft-versus-host disease and of HLA components to prevent graft rejection by the host immune system.

HIV Infection/AIDS

In patients infected with HIV, T cells or HSPCs are harvested and the *CCR5* gene is knocked out *in vitro* by CRISPR/Cas9 or ZFNs before reinfusion into the patient (Holt et al., 2010; Perez et al., 2008; Tebas et al., 2014; Wang et al., 2014; Xu et al., 2017; Yu et al., 2018). As *CCR5* is essential for HIV invasion of T cells, this intervention should prevent spreading of the virus into the engineered cells (Brelot and Chakrabarti, 2018).

Genetic Disorders

Sickle cell disease (SCD) and β -thalassemia are caused by disease-associated variants in the *HBB* gene, which encodes the beta subunit of hemoglobin. One strategy involves inducing the production of fetal

hemoglobin (HbF), which can compensate for malformed adult hemoglobin (in SCD) or reduced levels of adult hemoglobin (in β -thalassemia). This can be accomplished by *ex vivo* disruption of the intronic erythroid-specific enhancer of *BCL11A* in HSPCs by CRISPR/Cas9 or ZFN (Chang et al., 2017b; Psatha et al., 2018). After reinfusion, these HSPCs will produce erythrocytes with reduced *BCL11A* levels. As *BCL11A* represses transcription from *HBG* (Liu et al., 2018a), which encodes the gamma subunit of HbF, the level of the gamma subunit and thus HbF will increase. In another approach, one clinical trial is currently investigating direct gene restoration by CRISPR/Cas9 *ex vivo* in iPSC-derived HSPCs for the treatment of β -thalassemia (Cai et al., 2018; Martin et al., 2019; Park et al., 2019b; Wattanapanitch et al., 2018).

Clinical trials have started for the monogenic disorders hemophilia B and mucopolysaccharidosis (MPS) types I and II. Hemophilia B is caused by a deficiency of clotting factor IX, whereas MPS I and II are lysosomal storage disorders caused by deficiency of enzymes involved in the lysosomal degradation of glycosaminoglycans. These trials use systemic injection of AAV-expressing ZFN to achieve the targeted integration of a functional copy of the deficient gene into the albumin locus in liver cells, which secrete the enzyme into the circulation (Laoharawee et al., 2018; Sharma et al., 2015).

Finally, gene editing is being investigated in one of the subtypes of Leber's congenital amaurosis (LCA), a congenital retinal dystrophy. The subtype LCA10 is most commonly caused by an intronic variant in the *CEP290* gene that generates a cryptic splice site, resulting in defective protein production (Xu et al., 2018). A CRISPR/Cas9-based gene-editing strategy has been developed that removes the cryptic splice site; it is administered *in vivo* via subretinal injection using AAV as delivery vehicle (Maeder et al., 2019; Ruan et al., 2017).

FUTURE CHALLENGES

Among the most challenging aspects of gene editing to date are its delivery, specificity, and efficacy. The delivery strategy that is most appropriate for gene-editing intervention will depend on the cell-targeting approach that is most suitable for treating a specific disease. *Ex vivo* strategies can be conducted in a controlled environment, are amenable to efficient engineering using multiple methods, and may be a more straightforward way of bringing precision gene-editing medicine to the clinic. For example, when applied to stem cells (including adult stem cells or iPSCs) *ex vivo* gene editing can be subjected to quality control for genotoxicity before it is decided to engraft the cells in question. The use of *ex vivo* protein transduction ensures a transient exposure to nucleases, as opposed to the long-term exposure with *in vivo* approaches that use DNA constructs. However, *ex vivo* strategies face other issues associated with *ex vivo* cell manipulation. Engineered cells have to engraft efficiently into host individuals and should evade immune rejections if they are derived from a different donor. On the other hand, *in vivo* strategies will face the challenge of uncontrolled off-target events. In such cases, the selection of a suitable delivery vehicle, of a highly specific gene-targeting tool, and of possibly a strategy for spatiotemporal control, will all be especially relevant to preventing unintended variants and reducing potential immune reactions. *In vivo* gene-editing strategies that use persistent CAS9 expression—such as AAV—as a delivery vehicle may pose a risk of long-term exposure to gene-editing events, increasing the risk for off-target effects. Despite these challenges, the ongoing iteration of the CRISPR/Cas system into versions with improved specificity and efficacy holds great promise for a wide range of clinical applications. It will be important to remain cautious and, when the time comes for clinical testing, to evaluate the advantages against the possible risks. To fully understand the long-term effects of potential new treatments involving precise genomic engineering, thorough pre-clinical work is required.

ACKNOWLEDGMENTS

This work has received funding from Texnet, Zeldzame Ziekten Fonds/WE Foundation, Metakids, and Stofwisselkracht. The collaboration project is co-funded by the PPP Allowance made available by Health ~ Holland, Top Sector Life Sciences & Health, to stimulate public-private partnerships (project numbers LSHM16008 and LSHM19015). The collaboration project is co-initiated by the Prinses Beatrix Spierfonds.

AUTHOR CONTRIBUTIONS

M.B., P.H.H., M.E., and W.W.M.P.P. wrote and revised the manuscript. M.B., P.H.H., and M.E. performed the graphical abstract and figures. A.T.v.d.P. contributed to the review and editing of the manuscript. A.T.v.d.P. and W.W.M.P.P. secured funding.

DECLARATION OF INTERESTS

A.T.v.d.P. has provided consulting services for various industries in the field of Pompe disease under an agreement between these industries and Erasmus MC, Rotterdam, The Netherlands.

REFERENCES

- Abudayyeh, O.O., Gootenberg, J.S., Konermann, S., Joung, J., Slaymaker, I.M., Cox, D.B., Shmakov, S., Makarova, K.S., Semenova, E., Minakhin, L., et al. (2016). C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* 353, aaf5573.
- Akcakaya, P., Bobbin, M.L., Guo, J.A., Lopez, J.M., Clemens, M.K., Garcia, S.P., Fellows, M.D., Porritt, M.J., Firth, M.A., Carreras, A., et al. (2018). In vivo CRISPR-Cas gene editing with no detectable genome-wide off-target mutations. *Nature* 561, 416–419.
- Allen, F., Crepaldi, L., Alsinet, C., Strong, A.J., Kleshchynikov, V., De Angeli, P., Palenikova, P., Kosicki, M., Bassett, A.R., Harding, H., et al. (2018). Mutations generated by repair of Cas9-induced double strand breaks are predictable from surrounding sequence. *Nat. Biotechnol.* 37, 64–72.
- Amosii, L., Hildyard, J.C.W., Li, H., Sanchez-Ortiz, E., Mireault, A., Caballero, D., Harron, R., Stathopoulou, T.R., Massey, C., Shelton, J.M., et al. (2018). Gene editing restores dystrophin expression in a canine model of Duchenne muscular dystrophy. *Science* 362, 86–91.
- Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157.
- Bak, R.O., Dever, D.P., and Porteus, M.H. (2018). CRISPR/Cas9 genome editing in human hematopoietic stem cells. *Nat. Protoc.* 13, 358–376.
- Baliou, S., Adamaki, M., Kyriakopoulos, A.M., Spandidos, D.A., Panayiotidis, M., Christodoulou, I., and Zoumpoulis, V. (2018). Role of the CRISPR system in controlling gene transcription and monitoring cell fate (Review). *Mol. Med. Rep.* 17, 1421–1427.
- Barrangou, R., and Doudna, J.A. (2016). Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* 34, 933–941.
- Batra, R., Nelles, D.A., Pirie, E., Blue, S.M., Marina, R.J., Wang, H., Chaim, I.A., Thomas, J.D., Zhang, N., Nguyen, V., et al. (2017). Elimination of toxic microsatellite repeat expansion RNA by RNA-targeting Cas9. *Cell* 170, 899–912.e10.
- Beane, J.D., Lee, G., Zheng, Z., Mendel, M., Abate-Daga, D., Bharathan, M., Black, M., Gandhi, N., Yu, Z., Chandran, S., et al. (2015). Clinical scale zinc finger nuclease-mediated gene editing of PD-1 in tumor infiltrating lymphocytes for the treatment of metastatic melanoma. *Mol. Ther.* 23, 1380–1390.
- Ben Jehuda, R., Shemer, Y., and Binah, O. (2018). Genome editing in induced pluripotent stem cells using CRISPR/Cas9. *Stem Cell Rev. Rep.* 14, 323–336.
- Bergsma, A.J., van der Wal, E., Broeders, M., van der Ploeg, A.T., and Pim Pijnappel, W.W.M. (2018). Alternative splicing in genetic diseases: improved diagnosis and novel treatment options. *Int. Rev. Cell Mol. Biol.* 335, 85–141.
- Beyret, E., Liao, H.K., Yamamoto, M., Hernandez-Benitez, R., Fu, Y., Erikson, G., Reddy, P., and Izpisua Belmonte, J.C. (2019). Single-dose CRISPR-Cas9 therapy extends lifespan of mice with Hutchinson-Gilford progeria syndrome. *Nat. Med.* 25, 419–422.
- Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res.* 41, 7429–7437.
- Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., and Ciccio, A. (2017). CRISPR-Mediated base editing enables efficient disruption of Eukaryotic genes through induction of STOP codons. *Mol. Cell* 67, 1068–1079.e4.
- Bordon, M., Rey, F., Fantini, V., Pansarasa, O., Di Giulio, A.M., Carelli, S., and Cereda, C. (2018). From neuronal differentiation of iPSCs to 3D neuro-organoids: modelling and therapy of neurodegenerative diseases. *Int. J. Mol. Sci.* 19, 3972.
- Brandao, K.O., Tabel, V.A., Atsma, D.E., Mummery, C.L., and Davis, R.P. (2017). Human pluripotent stem cell models of cardiac disease: from mechanisms to therapies. *Dis. Model. Mech.* 10, 1039–1059.
- Brelot, A., and Chakrabarti, L.A. (2018). CCR5 revisited: how mechanisms of HIV entry govern AIDS pathogenesis. *J. Mol. Biol.* 430, 2557–2589.
- Brown, K.R., Mair, B., Soste, M., and Moffat, J. (2019). CRISPR screens are feasible in TP53 wild-type cells. *Mol. Syst. Biol.* 15, e8679.
- Cai, L., Bai, H., Mahairaki, V., Gao, Y., He, C., Wen, Y., Jin, Y.C., Wang, Y., Pan, R.L., Qasba, A., et al. (2018). A universal approach to correct various HBB gene mutations in human stem cells for gene therapy of beta-thalassemia and sickle cell disease. *Stem Cells Transl. Med.* 7, 87–97.
- Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demicheli, F., et al. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36, 265–271.
- Chal, J., and Pourquie, O. (2017). Making muscle: skeletal myogenesis in vivo and in vitro. *Development* 144, 2104–2122.
- Chang, H.H.Y., Pannunzio, N.R., Adachi, N., and Lieber, M.R. (2017a). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* 18, 495–506.
- Chang, K.H., Smith, S.E., Sullivan, T., Chen, K., Zhou, Q., West, J.A., Liu, M., Liu, Y., Vieira, B.F., Sun, C., et al. (2017b). Long-Term engraftment and fetal globin induction upon BCL11A gene editing in bone-marrow-derived CD34(+) hematopoietic stem and progenitor cells. *Mol. Ther. Methods Clin. Dev.* 4, 137–148.
- Charlesworth, C.T., Deshpande, P.S., Dever, D.P., Camarena, J., Lemgart, V.T., Cromer, M.K., Vakulskas, C.A., Collingwood, M.A., Zhang, L., Bode, N.M., et al. (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* 25, 249–254.
- Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017a). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407–410.
- Chen, X., Janssen, J.M., Liu, J., Maggio, I., T'Jong, A.E.J., Mikkers, H.M.M., and Gonçalves, M.A.F.V. (2017b). In trans paired nicking triggers seamless genome editing without double-stranded DNA cutting. *Nat. Commun.* 8, 657.
- Chew, W.L., Tabebordbar, M., Cheng, J.K.W., Mali, P., Wu, E.Y., Ng, A.H.M., Zhu, K., Wagers, A.J., and Church, G.M. (2016). A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* 13, 868–874.
- Chu, V.T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K., and Kuhn, R. (2015). Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat. Biotechnol.* 33, 543–548.
- Colella, P., Ronzitti, G., and Mingozzi, F. (2018). Emerging issues in AAV-mediated in vivo gene therapy. *Mol. Ther. Methods Clin. Dev.* 8, 87–104.
- Compagnucci, C., Nizzardo, M., Corti, S., Zanni, G., and Bertini, E. (2014). In vitro neurogenesis: development and functional implications of iPSC technology. *Cell Mol. Life Sci.* 71, 1623–1639.
- Cooper, M.L., Choi, J., Staser, K., Ritchey, J.K., Devenport, J.M., Eckardt, K., Rettig, M.P., Wang, B., Eisenberg, L.G., Ghobadi, A., et al. (2018). An "off-the-shelf" fratricide-resistant CAR-T for the treatment of T cell hematologic malignancies. *Leukemia* 32, 1970–1983.
- Cox, D.B.T., Gootenberg, J.S., Abudayyeh, O.O., Franklin, B., Kellner, M.J., Joung, J., and Zhang, F. (2017). RNA editing with CRISPR-Cas13. *Science* 358, 1019–1027.
- Devalla, H.D., and Passier, R. (2018). Cardiac differentiation of pluripotent stem cells and implications for modeling the heart in health and disease. *Sci. Transl. Med.* 10, eaah5457.
- Ding, W., Hu, Z., Zhu, D., Jiang, X., Yu, L., Wang, X., Zhang, C., Wang, L., Ji, T., Li, K., et al. (2014). Zinc finger nucleases targeting the human papillomavirus E7 oncogene induce E7 disruption and a transformed phenotype in

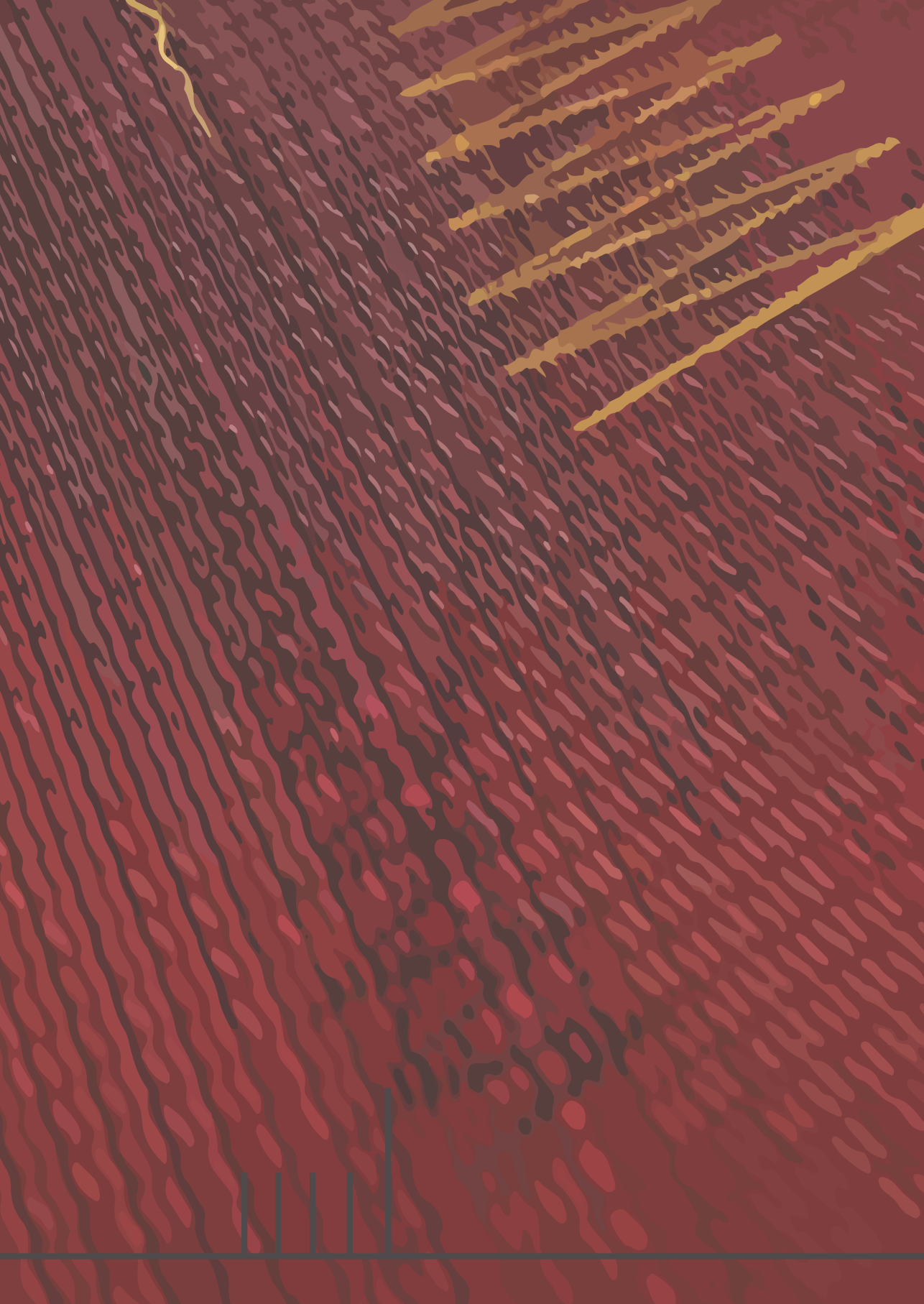
- HPV16/18-positive cervical cancer cells. *Clin. Cancer Res.* 20, 6495–6503.
- East-Seletsky, A., O'Connell, M.R., Knight, S.C., Burstein, D., Cate, J.H., Tian, R., and Doudna, J.A. (2016). Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* 538, 270–273.
- El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Gunther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.L., et al. (2019). Genetic compensation triggered by mutant mRNA degradation. *Nature* 568, 193–197.
- Fernandez, A., Josa, S., and Montoliu, L. (2017). A history of genome editing in mammals. *Mamm. Genome* 28, 237–246.
- Finn, J.D., Smith, A.R., Patel, M.C., Shaw, L., Younis, M.R., van Heteren, J., Dirstine, T., Ciullo, C., Lescarbeau, R., Seitzer, J., et al. (2018). A single administration of CRISPR/Cas9 lipid nanoparticles achieves robust and persistent in vivo genome editing. *Cell Rep.* 22, 2455–2468.
- Fiorotto, R., Amenduni, M., Mariotti, V., Fabris, L., Spirli, C., and Strazabosco, M. (2019). Liver diseases in the dish: iPSC and organoids as a new approach to modeling liver diseases. *Biochim. Biophys. Acta Mol. Basis Dis.* 1865, 920–928.
- Frit, P., Barboule, N., Yuan, Y., Gomez, D., and Calso, P. (2014). Alternative end-joining pathway(s): bricolage at DNA breaks. *DNA Repair (Amst.)* 17, 81–97.
- Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., Joung, J.K., Unit, P., Hospital, M.G., Biology, I., Hospital, M.G., and Hospital, M.G. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284.
- Gao, L., Cox, D.B.T., Yan, W.X., Manteiga, J.C., Schneider, M.W., Yamano, T., Nishimura, H., Nureki, O., Crosetto, N., and Zhang, F. (2017). Engineered Cpf1 variants with altered PAM specificities increase genome targeting range. *Nat. Biotechnol.* 35, 789–792.
- Gao, X., Tao, Y., Lamas, V., Huang, M., Yeh, W.H., Pan, B., Hu, Y.J., Hu, J.H., Thompson, D.B., Shu, Y., et al. (2018). Treatment of autosomal dominant hearing loss by in vivo delivery of genome editing agents. *Nature* 553, 217–221.
- Garneau, J.E., Dupuis, M.E., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadan, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71.
- Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U S A* 109, E2579–E2586.
- Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A/T to G/C in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
- Glass, Z., Lee, M., Li, Y., and Xu, Q. (2018). Engineering the delivery system for CRISPR-based genome editing. *Trends Biotechnol.* 36, 173–185.
- Gomes-Silva, D., Srinivasan, M., Sharma, S., Lee, C.M., Wagner, D.L., Davis, T.H., Rouce, R.H., Bao, G., Brenner, M.K., and Mamontkin, M. (2017). CD7-edited T cells expressing a CD7-specific CAR for the therapy of T-cell malignancies. *Blood* 130, 285–296.
- Grunewald, J., Zhou, R., Garcia, S.P., Iyer, S., Lareau, C.A., Aryee, M.J., and Joung, J.K. (2019a). Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base. *Nature* 569, 433–437.
- Grunewald, J., Zhou, R., Iyer, S., Lareau, C.A., Garcia, S.P., Aryee, M.J., and Joung, J.K. (2019b). CRISPR DNA base editors with reduced RNA off-target and self-editing activities. *Nat. Biotechnol.* 37, 1041–1048.
- Guha, T.K., and Edgell, D.R. (2017). Applications of alternative nucleases in the age of CRISPR/Cas9. *Int. J. Mol. Sci.* 18, 2565.
- Guha, T.K., Wai, A., and Hausner, G. (2017). Programmable genome editing tools and their regulation for efficient genome engineering. *Comput. Struct. Biotechnol. J.* 15, 146–160.
- Guo, Q., Mintier, G., Ma-Edmonds, M., Storton, D., Wang, X., Xiao, X., Kienle, B., Zhao, D., and Feder, J.N. (2018a). Cold shock increases the frequency of homology directed repair gene editing in induced pluripotent stem cells. *Sci. Rep.* 8, 2080.
- Guo, X., Jiang, H., Shi, B., Zhou, M., Zhang, H., Shi, Z., Du, G., Luo, H., Wu, X., Wang, Y., et al. (2018b). Disruption of PD-1 enhanced the anti-tumor activity of chimeric antigen receptor T cells against hepatocellular carcinoma. *Front. Pharmacol.* 9, 1118.
- Gupta, R.M., and Musunuru, K. (2014). Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Invest.* 124, 4154–4161.
- Haapaniemi, E., Botla, S., Persson, J., Schmierer, B., and Taipale, J. (2018). CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat. Med.* 1.
- Hannoun, Z., Steichen, C., Dianat, N., Weber, A., and Dubart-Kupferschmitt, A. (2016). The potential of induced pluripotent stem cell derived hepatocytes. *J. Hepatol.* 65, 182–199.
- Hendel, A., Bak, R.O., Clark, J.T., Kennedy, A.B., Ryan, D.E., Roy, S., Steinfeld, I., Lunstad, B.D., Kaiser, R.J., and Alec, B. (2015). Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* 33, 985–989.
- Hess, G.T., Tycko, J., Yao, D., and Bassik, M.C. (2017). Methods and applications of CRISPR-mediated base editing in Eukaryotic genomes. *Mol. Cell* 68, 26–43.
- Hilton, I.B., D'Ipollito, A.M., Vockley, C.M., Thakore, P.I., Crawford, G.E., Reddy, T.E., and Gersbach, C.A. (2015). Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers. *Nat. Biotechnol.* 33, 510–517.
- Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., Yamamoto, T., and Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci. Rep.* 5, 8841.
- Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C., and Cannon, P.M. (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat. Biotechnol.* 28, 839–847.
- Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832.
- Hu, Z., Yu, L., Zhu, D., Ding, W., Wang, X., Zhang, C., Wang, L., Jiang, X., Shen, H., He, D., et al. (2014). Disruption of HPV16-E7 by CRISPR/Cas system induces apoptosis and growth inhibition in HPV16 positive human cervical cancer cells. *Biomed. Res. Int.* 2014, 612823.
- Hu, Z., Ding, W., Zhu, D., Yu, L., Jiang, X., Wang, X., Zhang, C., Wang, L., Ji, T., Liu, D., et al. (2015). TALEN-mediated targeting of HPV oncogenes ameliorates HPV-related cervical malignancy. *J. Clin. Invest.* 125, 425–436.
- Hu, J.H., Miller, S.M., Geurts, M.H., Tang, W., Chen, L., Sun, N., Zeina, C.M., Gao, X., Rees, H.A., Lin, Z., and Liu, D.R. (2018). Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature* 556, 57–63.
- Hu, B., Zou, Y., Zhang, L., Tang, J., Niedermann, G., Firat, E., Huang, X., and Zhu, X. (2019a). Nucleofection with plasmid DNA for CRISPR/Cas9-Mediated inactivation of programmed cell death protein 1 in CD133-specific CAR T cells. *Hum. Gene Ther.* 30, 446–458.
- Hu, W., Zi, Z., Jin, Y., Li, G., Shao, K., Cai, Q., Ma, X., and Wei, F. (2019b). CRISPR/Cas9-mediated PD-1 disruption enhances human mesothelin-targeted CAR T cell effector functions. *Cancer Immunol. Immunother.* 68, 365–377.
- Huang, X., Zhou, G., Wu, W., Duan, Y., Ma, G., Song, J., Xiao, R., Vandenberghe, L., Zhang, F., D'Amore, P.A., and Lei, H. (2017). Genome editing abrogates angiogenesis in vivo. *Nat. Commun.* 8, 4–11.
- Huang, C.-h., Lee, K.-c., and Doudna, J.A. (2018). Applications of CRISPR-Cas enzymes in cancer therapeutics and detection. *Trends Cancer* 4, 499–512.
- Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J., Sondey, M., Ye, C., et al. (2018). p53 inhibits CRISPR–Cas9 engineering in human pluripotent stem cells. *Nat. Med.* 1–8.
- Ihry, R.J., Salick, M.R., Ho, D.J., Sondey, M., Kommineni, S., Paula, S., Raymond, J., Henry, B., Frias, E., Wang, Q., et al. (2019). Genome-Scale CRISPR screens identify human pluripotency-specific genes. *Cell Rep.* 27, 616–630.e6.
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia*

- coli, and identification of the gene product. *J. Bacteriol.* 169, 5429–5433.
- Jang, Y.Y., and Ye, Z. (2016). Gene correction in patient-specific iPSCs for therapy development and disease modeling. *Hum. Genet.* 135, 1041–1058.
- Jansen, R., Embden, J.D., Gastra, W., and Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43, 1565–1575.
- Jiang, F., Taylor, D.W., Chen, J.S., Kornfeld, J.E., Zhou, K., Thompson, A.J., Nogales, E., and Doudna, J.A. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* 351, 867–871.
- Jiang, F., Liu, J.J., Osuna, B.A., Xu, M., Berry, J.D., Rauch, B.J., Nogales, E., Bondy-Denomy, J., and Doudna, J.A. (2019). Temperature-Responsive competitive inhibition of CRISPR-Cas9. *Mol. Cell* 73, 601–610.e5.
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821.
- Karvelis, T., Gasunas, G., and Siksnys, V. (2017). Harnessing the natural diversity and in vitro evolution of Cas9 to expand the genome editing toolbox. *Curr. Opin. Microbiol.* 37, 88–94.
- Kennedy, E.M., Kornepati, A.V., Goldstein, M., Bogerd, H.P., Poling, B.C., Whisnant, A.W., Kastan, M.B., and Cullen, B.R. (2014). Inactivation of the human papillomavirus E6 or E7 gene in cervical carcinoma cells by using a bacterial CRISPR/Cas RNA-guided endonuclease. *J. Virol.* 88, 11965–11972.
- Kim, E., Koo, T., Park, S.W., Kim, D., Kim, K., Cho, H.Y., Song, D.W., Lee, K.J., Jung, M.H., Kim, S., et al. (2017). In vivo genome editing with a small Cas9 orthologue derived from *Campylobacter jejuni*. *Nat. Commun.* 8, 1–12.
- Kim, S., Koo, T., Jee, H.-G., Cho, H.-Y., Lee, G., Lim, D.-G., Shin, H.S., and Kim, J.-S. (2018a). CRISPR RNAs trigger innate immune responses in human cells. *Genome Res.* 28, 367–373.
- Kim, S.I., Matsumoto, T., Kagawa, H., Nakamura, M., Hirohata, R., Ueno, A., Ohishi, M., Sakuma, T., Soga, T., Yamamoto, T., and Woltjen, K. (2018b). Microhomology-assisted scarless genome editing in human iPSCs. *Nat. Commun.* 9, 939.
- Kim, D., Kim, D.E., Lee, G., Cho, S.I., and Kim, J.S. (2019). Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat. Biotechnol.* 37, 430–435.
- Kleistiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.R.J., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485.
- Kleistiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N., Zheng, Z., Joong, J.K., Unit, P., Biology, I., Hospital, M.G., and Kong, H. (2016). High-fidelity CRISPR-Cas9 variants with undetectable genome-wide off-targets. *Nature* 529, 490–495.
- Kocak, D.D., Josephs, E.A., Bhandarkar, V., Adkar, S.S., Kwon, J.B., and Gersbach, C.A. (2019). Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37, 657–666.
- Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
- Konermann, S., Lotfy, P., Brideau, N.J., Oki, J., Shokhiev, M.N., and Hsu, P.D. (2018). Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* 173, 665–676.e14.
- Kosicki, M., and Bradley, A. (2018). Repair of CRISPR-Cas9-induced double-stranded breaks leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771.
- Kuscu, C., Parlak, M., Tufan, T., Yang, J., Szlachta, K., Wei, X., Mammadov, R., and Adli, M. (2017). CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat. Methods* 14, 710–712.
- Laoharawee, K., DeKaveler, R.C., Podetz-Pedersen, K.M., Rohde, M., Sproul, S., Nguyen, H.O., Nguyen, T., St Martin, S.J., Ou, L., et al. (2018). Dose-dependent prevention of metabolic and neurologic disease in murine MPS II by ZFN-mediated in vivo genome editing. *Mol. Ther.* 26, 1127–1136.
- Laughery, M.F., Mayes, H.C., Pedroza, I.K., and Wyrick, J.J. (2019). R-loop formation by dCas9 is mutagenic in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 47, 2389–2401.
- Lee, K., Conboy, M., Park, H.M., Jiang, F., Kim, H.J., Dewitt, M.A., Mackley, V.A., Chang, K., Rao, A., Skinner, C., et al. (2017). Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA in vivo induces homology-directed DNA repair. *Nat. Biomed. Eng.* 1, 889–901.
- Lee, B., Lee, K., Panda, S., Gonzales-Rojas, R., Chong, A., Bugay, V., Park, H.M., Brenner, R., Murthy, N., and Lee, H.Y. (2018). Nanoparticle delivery of CRISPR into the brain rescues a mouse model of fragile X syndrome from exaggerated repetitive behaviours. *Nat. Biomed. Eng.* 2, 497–507.
- Lessard, S., Francioli, L., Alfoldi, J., Tardif, J.-C., Ellinor, P.T., MacArthur, D.G., Lettre, G., Orkin, S.H., and Canver, M.C. (2017). Human genetic variation alters CRISPR-Cas9 on- and off-targeting specificity at therapeutically implicated loci. *Proc. Natl. Acad. Sci. U S A* 114, E11257–E11266.
- Li, H.L., Fujimoto, N., Sasakawa, N., Shirai, S., Ohkame, T., Sakuma, T., Tanaka, M., Amano, N., Watanabe, A., Sakurai, H., et al. (2015). Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports* 4, 143–154.
- Li, L., Hu, S., and Chen, X. (2018). Non-viral delivery systems for CRISPR/Cas9-based genome editing: challenges and opportunities. *Biomaterials* 171, 207–218.
- Liao, H.K., Hatanaka, F., Araoka, T., Reddy, P., Wu, M.Z., Sui, Y., Yamauchi, T., Sakurai, M., O’Keefe, D.D., Nunez-Delgado, E., et al. (2017). In vivo target gene activation via CRISPR/Cas9-Mediated trans-epigenetic modulation. *Cell* 171, 1495–1507.e15.
- Listgarten, J., Weinstein, M., Kleinstiver, B.P., Sousa, A.A., Joung, J.K., Crawford, J., Gao, K., Hoang, L., Elibol, M., Doench, J.G., and Fusi, N. (2018). Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs. *Nat. Biomed. Eng.* 2, 38–47.
- Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018a). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e17.
- Liu, X.S., Wu, H., Krzisch, M., Wu, X., Graef, J., Muffat, J., Hniz, D., Li, C.H., Yuan, B., Xu, C., et al. (2018b). Rescue of fragile X syndrome neurons by DNA methylation editing of the FMR1 gene. *Cell* 172, 979–991.e6.
- Louis Jeune, V., Joergensen, J.A., Hajjar, R.J., and Weber, T. (2013). Pre-existing anti-adenosine-associated virus antibodies as a challenge in AAV gene therapy. *Hum. Gene Ther. Methods* 24, 59–67.
- Lu, S., Yang, N., He, J., Gong, W., Lai, Z., Xie, L., Tao, L., Xu, C., Wang, H., Zhang, G., et al. (2019). Generation of cancer-specific cytotoxic PD-1(+) T cells using liposome-encapsulated CRISPR/Cas system with dendritic/tumor fusion cells. *J. Biomed. Nanotechnol.* 15, 593–601.
- Maeder, M.L., and Gersbach, C.A. (2016). Genome-editing technologies for gene and cell therapy. *Mol. Ther.* 24, 430–446.
- Maeder, M.L., Stefanidakis, M., Wilson, C.J., Baral, R., Barrera, L.A., Bounoutos, G.S., Bumcrot, D., Chao, H., Ciulla, D.M., DaSilva, J.A., et al. (2019). Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* 25, 229–233.
- Magli, A., and Perlingeiro, R.R.C. (2017). Myogenic progenitor specification from pluripotent stem cells. *Semin. Cell Dev. Biol.* 72, 87–98.
- Mahas, A., Neal Stewart, C., Jr., and Mahfouz, M.M. (2018). Harnessing CRISPR/Cas systems for programmable transcriptional and post-transcriptional regulation. *Biotechnol. Adv.* 36, 295–310.
- Mair, B., Tomic, J., Masud, S.N., Tonge, P., Weiss, A., Usaj, M., Tong, A.H.Y., Kwan, J.J., Brown, K.R., Titus, E., et al. (2019). Essential gene profiles for human pluripotent stem cells identify uncharacterized genes and substrate dependencies. *Cell Rep.* 27, 599–615.e12.
- Mandai, M., Watanabe, A., Kurimoto, Y., Hirami, Y., Morinaga, C., Daimon, T., Fujiwara, M., Akimaru, H., Sakai, N., Shibata, Y., et al. (2017). Autologous induced stem-cell-derived retinal cells for macular degeneration. *N. Engl. J. Med.* 376, 1038–1046.
- Martin, R.M., Ikeda, K., Cromer, M.K., Uchida, N., Nishimura, T., Romano, R., Tong, A.J., Lemgart, V.T., Camarena, J., Pavel-Dinu, M., et al. (2019). Highly efficient and marker-free genome editing of human pluripotent stem cells by CRISPR-Cas9

- RNP and AAV6 donor-mediated homologous recombination. *Cell Stem Cell* 24, 821–828.e5.
- Maruyama, T., Dougan, S.K., Truttmann, M.C., Bilate, A.M., Ingram, J.R., and Ploegh, H.L. (2015). Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat. Biotechnol.* 33, 538–542.
- Mateos-Gomez, P.A., Kent, T., Deng, S.K., McDevitt, S., Kashkina, E., Hoang, T.M., Pomerantz, R.T., and Sfeir, A. (2017). The helicase domain of Poltheta counteracts RPA to promote alt-NHEJ. *Nat. Struct. Mol. Biol.* 24, 1116–1123.
- Matharu, N., Rattanasopha, S., Tamura, S., Maliskova, L., Wang, Y., Bernard, A., Hardin, A., Eckalbar, W.L., Vaisse, C., and Ahtivu, N. (2018). CRISPR-mediated activation of a promoter or enhancer rescues obesity caused by haploinsufficiency. *Science* 0629, eaau0629.
- Menger, L., Sledzinska, A., Bergerhoff, K., Vargas, F.A., Smith, J., Poirat, L., Pule, M., Herero, J., Peggs, K.S., and Quesada, S.A. (2016). TALEN-Mediated inactivation of PD-1 in tumor-reactive lymphocytes promotes intratumoral T-cell persistence and rejection of established tumors. *Cancer Res.* 76, 2087–2093.
- Mingozzi, F., and High, K.A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat. Rev. Genet.* 12, 341–355.
- Mingozzi, F., and High, K.A. (2013). Immune responses to AAV vectors: overcoming barriers to successful gene therapy. *Blood* 122, 23–36.
- Mojica, F.J., and Montolieu, L. (2016). On the origin of CRISPR-Cas technology: from prokaryotes to mammals. *Trends Microbiol.* 24, 811–820.
- Molla, K.A., and Yang, Y. (2019). CRISPR/Cas-Mediated base editing: technical considerations and practical applications. *Trends Biotechnol.* 37, 1121–1142.
- Mou, H., Smith, J.L., Peng, L., Yin, H., Moore, J., Zhang, X.O., Song, C.Q., Sheel, A., Wu, Q., Ozata, D.M., et al. (2017). CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion. *Genome Biol.* 18, 108.
- Nakade, S., Tsubota, T., Sakane, Y., Kume, S., Sakamoto, N., Obara, M., Daimon, T., Sezutsu, H., Yamamoto, T., Sakuma, T., and Suzuki, K.T. (2014). Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat. Commun.* 5, 5560.
- Naldini, L. (2015). Gene therapy returns to centre stage. *Nature* 526, 351–360.
- Nelson, C.E., Hakim, C.H., Ousterout, D.G., Thakore, P.I., Moreb, E.A., Castellanos Rivera, R.M., Madhavan, S., Pan, X., Ran, F.A., et al. (2016). In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* 351, 403–407.
- Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., Mochizuki, M., Miyabe, A., Araki, M., Hara, K.Y., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, aaf8729.
- Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayeh, O.O., Gootenberg, J.S., Mori, H., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science* 361, 1259–1262.
- Nishiyama, J., Mikuni, T., and Yasuda, R. (2017). Virus-Mediated genome editing via homology-directed repair in mitotic and postmitotic cells in mammalian brain. *Neuron* 96, 755–768.e5.
- O'Connell, M.R., Oakes, B.L., Sternberg, S.H., East-Seletsky, A., Kaplan, M., and Doudna, J.A. (2014). Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* 516, 263–266.
- Orthwein, A., Noordermeer, S.M., Wilson, M.D., Landry, S., Enchev, R.I., Sherker, A., Munro, M., Pinder, J., Salsman, J., Delliare, G., et al. (2015). A mechanism for the suppression of homologous recombination in G1 cells. *Nature* 528, 422–426.
- Ouchi, Y., Patil, A., Tamura, Y., Nishimasu, H., Negishi, A., Paul, S.K., Takemura, N., Satoh, T., Kimura, Y., Kurachi, M., et al. (2018). Generation of tumor antigen-specific murine CD8⁺ T cells with enhanced anti-tumor activity via highly efficient CRISPR/Cas9 genome editing. *Int. Immunol.* 30, 141–154.
- Pankowicz, F.P., Barzi, M., Legras, X., Hubert, L., Mi, T., Tomolonis, J.A., Ravishanker, M., Sun, Q., Yang, D., Borowiak, M., et al. (2016). Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia. *Nat. Commun.* 7, 1–6.
- Park, H., Oh, J., Shim, G., Cho, B., Chang, Y., Kim, S., Baek, S., Kim, H., Shin, J., Choi, H., et al. (2019a). In vivo neuronal gene editing via CRISPR-Cas9 amphiphilic nanocomplexes alleviates deficits in mouse models of Alzheimers disease. *Nat. Neurosci.* 22, 524–528.
- Park, S.H., Lee, C.M., Dever, D.P., Davis, T.H., Camarena, J., Sfrifa, W., Zhang, Y., Paikari, A., Chang, A.K., Porteus, M.H., et al. (2019b). Highly efficient editing of the beta-globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* 47, 7955–7972.
- Pawluk, A., Amrani, N., Zhang, Y., Garcia, B., Hidalgo-Reyes, Y., Lee, J., Edraki, A., Shah, M., Sontheimer, E.J., Maxwell, K.L., and Davidson, A.R. (2016). Naturally occurring off-switches for CRISPR-Cas9. *Cell* 167, 1829–1838.e9.
- Pawluk, A., Davidson, A.R., and Maxwell, K.L. (2018). Anti-CRISPR: discovery, mechanism and function. *Nat. Rev. Microbiol.* 16, 12–17.
- Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L., et al. (2008). Establishment of HIV-1 resistance in CD4⁺ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* 26, 808–816.
- Pinder, J., Salsman, J., and Delliare, G. (2015). Nuclear domain 'knock-in' screen for the evaluation and identification of small molecule enhancers of CRISPR-based genome editing. *Nucleic Acids Res.* 43, 9379–9392.
- Psatha, N., Reik, A., Phelps, S., Zhou, Y., Dalas, D., Yannaki, E., Levasseur, D.N., Urnov, F.D., Holmes, M.C., and Papayannopoulou, T. (2018). Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with beta-thalassemia major. *Mol. Ther. Methods Clin. Dev.* 10, 313–326.
- Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183.
- Ran, F.A., Hsu, P.D., Lin, C.-y., Gootenberg, J.S., Trevino, A., Scott, D.A., Inoue, A., and Matoba, S. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 154, 1380–1389.
- Ranjha, L., Howard, S.M., and Cejka, P. (2018). Main steps in DNA double-strand break repair: an introduction to homologous recombination and related processes. *Chromosoma* 127, 187–214.
- Richter, F., Fonfara, I., Gelfert, R., Nack, J., Charpentier, E., and Möglich, A. (2017). Switchable Cas9. *Curr. Opin. Biotechnol.* 48, 119–126.
- Rodriguez-Rodriguez, D.R., Ramirez-Solis, R., Garza-Elizondo, M.A., Garza-Rodriguez, M.L., and Barrera-Saldana, H.A. (2019). Genome editing: a perspective on the application of CRISPR/Cas9 to study human diseases (Review). *Int. J. Mol. Med.* 43, 1559–1574.
- Rossidis, A.C., Stratigis, J.D., Chadwick, A.C., Hartman, H.A., Ahn, N.J., Li, H., Singh, K., Coons, B.E., Li, L., Lv, W., et al. (2018). In utero CRISPR-mediated therapeutic editing of metabolic genes. *Nat. Med.* 24, 1513–1518.
- Ruan, G.X., Barry, E., Yu, D., Lukason, M., Cheng, S.H., and Scaria, A. (2017). CRISPR/Cas9-Mediated genome editing as a therapeutic approach for leber congenital amaurosis 10. *Mol. Ther.* 25, 331–341.
- Rupp, L.J., Schumann, K., Roybal, K.T., Gate, R.E., Ye, C.J., Lim, W.A., and Marson, A. (2017). CRISPR/Cas9-mediated PD-1 disruption enhances anti-tumor efficacy of human chimeric antigen receptor T cells. *Sci. Rep.* 7, 737.
- Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2011). Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* 12, 51–58.
- Sakuma, T., Nakade, S., Sakane, Y., Suzuki, K.T., and Yamamoto, T. (2016). MMEJ-assisted gene knock-in using TALENs and CRISPR/Cas9 with the PITCH systems. *Nat. Protoc.* 11, 118–133.
- Santiago-Fernandez, O., Osorio, F.G., Quesada, V., Rodriguez, F., Basso, S., Maeso, D., Rolas, L., Barkaway, A., Nourshargh, S., Folgueras, A.R., et al. (2019). Development of a CRISPR/Cas9-based therapy for Hutchinson-Gilford progeria syndrome. *Nat. Med.* 25, 423–426.
- Sawatsubashi, S., Joko, Y., Fukumoto, S., Matsumoto, T., and Sugano, S.S. (2018). Development of versatile non-homologous end joining-based knock-in module for genome editing. *Sci. Rep.* 8, 593.

- Schimmel, J., Kool, H., van Schendel, R., and Tijsterman, M. (2017). Mutational signatures of non-homologous and polymerase theta-mediated end-joining in embryonic stem cells. *EMBO J.* 36, 3634–3649.
- Shirol, G., Conti, A., Ferrari, S., Della Volpe, L., Jacob, A., Albano, L., Beretta, S., Calabria, A., Vavassori, V., Gasparini, P., et al. (2019). Precise gene editing preserves hematopoietic stem cell function following transient p53-mediated DNA damage response. *Cell Stem Cell* 24, 551–565.e8.
- Schwank, G., Koo, B.K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. *Cell Stem Cell* 13, 653–658.
- Shankar, S., Prasad, D., Sanawar, R., Das, A.V., and Pillai, M.R. (2017). TALEN based HPV-E7 editing triggers necrotic cell death in cervical cancer cells. *Sci. Rep.* 7, 5500.
- Sharma, R., Anguela, X.M., Doyon, Y., Wechsler, T., DeKaveler, R.C., Sproul, S., Paschon, D.E., Miller, J.C., Davidson, R.J., Shivak, D., et al. (2015). In vivo genome editing of the albumin locus as a platform for protein replacement therapy. *Blood* 126, 1777–1784.
- Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88.
- Smith, J.L., Mou, H., and Xue, W. (2018). Understanding and repurposing CRISPR-mediated alternative splicing. *Genome Biol.* 19, 184.
- Song, J., Yang, D., Xu, J., Zhu, T., Chen, Y.E., and Zhang, J. (2016). RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat. Commun.* 7, 10548.
- Su, S., Hu, B., Shao, J., Shen, B., Du, J., Du, Y., Zhou, J., Yu, L., Zhang, L., Chen, F., et al. (2016). CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Sci. Rep.* 6, 20070.
- Su, S., Zou, Z., Chen, F., Ding, N., Du, J., Shao, J., Li, L., Fu, Y., Hu, B., Yang, Y., et al. (2017). CRISPR-Cas9-mediated disruption of PD-1 on human T cells for adoptive cellular therapies of EBV positive gastric cancer. *Oncotarget* 6, e1249558.
- Suzuki, K., and Izpisua Belmonte, J.C. (2018). In vivo genome editing via the HITI method as a tool for gene therapy. *J. Hum. Genet.* 63, 157–164.
- Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., Wu, J., Zhu, J., Kim, E.J., Hatanaka, F., Yamamoto, M., Araoka, T., Li, Z., et al. (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature* 540, 144–149.
- Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gascienas, G., Karvelis, T., Pschera, P., Siksnys, V., and Seidel, R. (2014). Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc. Natl. Acad. Sci. U S A* 111, 9798–9803.
- Tabebordbar, M., Zhu, K., Cheng, J.K.W., Chew, W.L., Widrick, J.J., Yan, W.X., Maesner, C., Wu, E.Y., Xiao, R., Ran, F.A., et al. (2016). In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* 351, 407–411.
- Taleei, R., and Nikjoo, H. (2013). Biochemical DSB-repair model for mammalian cells in G1 and early S phases of the cell cycle. *Mutat. Res.* 756, 206–212.
- Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Surosky, R.T., Giedlin, M.A., Nichol, G., et al. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370, 901–910.
- Thakore, P.I., D'Ippolito, A.M., Song, L., Safi, A., Shivakumar, N.K., Kabadi, A.M., Reddy, T.E., Crawford, G.E., and Gersbach, C.A. (2015). Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat. Methods* 12, 1143–1149.
- Thomas, M., Burgio, G., Adams, D.J., and Iyer, V. (2019). Collateral damage and CRISPR genome editing. *PLoS Genet.* 15, e1007994.
- Torikai, H., and Cooper, L.J. (2016). Translational implications for off-the-shelf immune cells expressing chimeric antigen receptors. *Mol. Ther.* 24, 1178–1186.
- Truong, L.N., Li, Y., Shi, L.Z., Hwang, P.Y., He, J., Wang, H., Razavian, N., Berns, M.W., and Wu, X. (2013). Microhomology-mediated End Joining and Homologous Recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc. Natl. Acad. Sci. U S A* 110, 7720–7725.
- Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., Joung, K., Unit, P., and Hospital, M.G. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576.
- Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D.H., Kim, J., et al. (2019). CRISPR-Cas9-based mutagenesis frequently provokes on-target mRNA misregulation. *Nat. Commun.* 10, 4056.
- Uribe, R.V., van der Helm, E., Misiakou, M.A., Lee, S.W., Kol, S., and Sommer, M.O.A. (2019). Discovery and characterization of Cas9 inhibitors disseminated across seven bacterial phyla. *Cell Host Microbe* 25, 233–241.e5.
- Villiger, L., Grisch-Can, H.M., Lindsay, H., Ringnald, F., Pogliano, C.B., Allegri, G., Fingerhut, R., Häberle, J., Matos, J., Robinson, M.D., et al. (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat. Med.* 24, 1519–1525.
- Wagner, D.L., Amani, L., Wendering, D.J., Reinke, P., Volk, H.-D., and Schmuck-Henneresse, M. (2019). High prevalence of Streptococcus pyogenes Cas9-reactive T cells within the adult human population. *Nat. Med.* 25, 242–248.
- van der Wal, E., Herrero-Hernandez, P., Wan, R., Broeders, M., In 't Groen, S.L.M., van Gestel, T.J.M., van IJcken, W.F.J., Cheung, T.H., van der Ploeg, A.T., Schaaf, G.J., and Pijnappel, W. (2018). Large-Scale expansion of human iPSC-derived skeletal muscle cells for disease modeling and cell-based therapeutic strategies. *Stem Cell Reports* 10, 1975–1990.
- Wang, W., Ye, C., Liu, J., Zhang, D., Kimata, J.T., and Zhou, P. (2014). CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PLoS One* 9, e115987.
- Wang, M., Zuris, J.A., Meng, F., Rees, H., Sun, S., Deng, P., Han, Y., Gao, X., Pouli, D., Wu, Q., et al. (2016). Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. *Proc. Natl. Acad. Sci. U S A* 113, 2868–2873.
- Wattanapanitch, M., Damkham, N., Potirait, P., Trakamsanga, K., Janan, M., U-Pratya, Y., Kheolamaj, P., Klincumhorn, N., and Issaragrisil, S. (2018). One-step genetic correction of hemoglobin E/beta-thalassemia patient-derived iPSCs by the CRISPR/Cas9 system. *Stem Cell Res. Ther.* 9, 46.
- Watters, K.E., Fellmann, C., Bai, H.B., Ren, S.M., and Doudna, J.A. (2018). Systematic discovery of natural CRISPR-Cas12a inhibitors. *Science* 5138, 1–8.
- Wienert, B., Wyman, S.K., Richardson, C.D., Yeh, C.D., Akcakaya, P., Porritt, M.J., Morlock, M., Vu, J.T., Kazane, K.R., Watry, H.L., et al. (2019). Unbiased detection of CRISPR off-targets in vivo using DISCOVER-Seq. *Science* 364, 286–289.
- Wu, W.Y., Lebbink, J.H.G., Kanaar, R., Geijsen, N., and van der Oost, J. (2018). Genome editing by natural and engineered CRISPR-associated nucleases. *Nat. Chem. Biol.* 14, 642–651.
- Xie, N., Gong, H., Suhl, J.A., Chopra, P., Wang, T., and Warren, S.T. (2016). Reactivation of FMRI by CRISPR/Cas9-Mediated deletion of the expanded CGG-repeat of the fragile X chromosome. *PLoS One* 11, e0165499.
- Xu, L., Yang, H., Gao, Y., Chen, Z., Xie, L., Liu, Y., Liu, Y., Wang, X., Li, H., Lai, W., et al. (2017). CRISPR/Cas9-Mediated CCR5 ablation in human hematopoietic stem/progenitor cells confers HIV-1 resistance in vivo. *Mol. Ther.* 25, 1782–1789.
- Xu, C.L., Cho, G.Y., Sengillo, J.D., Park, K.S., Mahajan, V.B., and Tsang, S.H. (2018). Translation of CRISPR genome surgery to the bedside for retinal diseases. *Front. Cell Dev. Biol.* 6, 46.
- Yang, Y., Jacoby, E., and Fry, T.J. (2015). Challenges and opportunities of allogeneic donor-derived CAR T cells. *Curr. Opin. Hematol.* 22, 509–515.
- Yin, H., Song, C.-q., Dorkin, J.R., Zhu, L.J., Li, Y., Wu, Q., Park, A., Yang, J., Suresh, S., Bizhanova, A., et al. (2016). Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat. Biotechnol.* 34, 328–333.
- Yin, C., Zhang, T., Qu, X., Zhang, Y., Putatunda, R., Xiao, X., Li, F., Xiao, W., Zhao, H., Dai, S., et al. (2017a). In vivo excision of HIV-1 provirus by saCas9 and multiplex single-guide RNAs in animal models. *Mol. Ther.* 25, 1168–1186.
- Yin, H., Song, C.-q., Suresh, S., Wu, Q., Walsh, S., Rhyh, L.H., Mintzer, E., Bolukbasi, M.F., Zhu, L.J., Kauffman, K., et al. (2017b). Structure-guided chemical modification of guide RNA enables

- potent non-viral in vivo genome editing. *Nat. Biotechnol.* 35, 1179–1187.
- Yin, H., Song, C.Q., Suresh, S., Kwan, S.Y., Wu, Q., Walsh, S., Ding, J., Bogorad, R.L., Zhu, L.J., Wolfe, S.A., et al. (2018). Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat. Chem. Biol.* 14, 311–316.
- Yu, C., Liu, Y., Ma, T., Liu, K., Xu, S., Zhang, Y., Liu, H., La Russa, M., Xie, M., Ding, S., and Qi, L.S. (2015). Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell* 16, 142–147.
- Yu, A.Q., Ding, Y., Lu, Z.Y., Hao, Y.Z., Teng, Z.P., Yan, S.R., Li, D.S., and Zeng, Y. (2018). TALENs-mediated homozygous CCR5Delta32 mutations endow CD4+ U87 cells with resistance against HIV1 infection. *Mol. Med. Rep.* 17, 243–249.
- Zelensky, A.N., Schimmel, J., Kool, H., Kanaar, R., and Tijsterman, M. (2017). Inactivation of Pol theta and C-NHEJ eliminates off-target integration of exogenous DNA. *Nat. Commun.* 8, 66.
- Zhan, T., Rindtorff, N., Betge, J., Ebert, M.P., and Boutros, M. (2019). CRISPR/Cas9 for cancer research and therapy. *Semin. Cancer Biol.* 55, 106–119.
- Zhang, J.H., Adikaram, P., Pandey, M., Genis, A., and Simonds, W.F. (2016). Optimization of genome editing through CRISPR-Cas9 engineering. *Bioengineered* 7, 166–174.
- Zhao, Z., Shi, L., Zhang, W., Han, J., Zhang, S., Fu, Z., and Cai, J. (2018). CRISPR knock out of programmed cell death protein 1 enhances anti-tumor activity of cytotoxic T lymphocytes. *Oncotarget* 9, 5208–5215.
- Zhen, S., Hua, L., Takahashi, Y., Narita, S., Liu, Y.H., and Li, Y. (2014). In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9. *Biochem. Biophys. Res. Commun.* 450, 1422–1426.
- Zhou, C., Sun, Y., Yan, R., Liu, Y., Zuo, E., Gu, C., Han, L., Wei, Y., Hu, X., Zeng, R., et al. (2019). Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature* 571, 275–278.
- Zischewski, J., Fischer, R., and Bortesi, L. (2017). Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol. Adv.* 35, 95–104.
- Zuo, E., Sun, Y., Wei, W., Yuan, T., Ying, W., Sun, H., Yuan, L., Steinmetz, L.M., Li, Y., and Yang, H. (2019). Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science* 364, 289–292.
- Zych, A.O., Bajor, M., and Zagodzón, R. (2018). Application of genome editing techniques in immunology. *Arch. Immunol. Ther. Exp. (Warsz)* 66, 289–298.



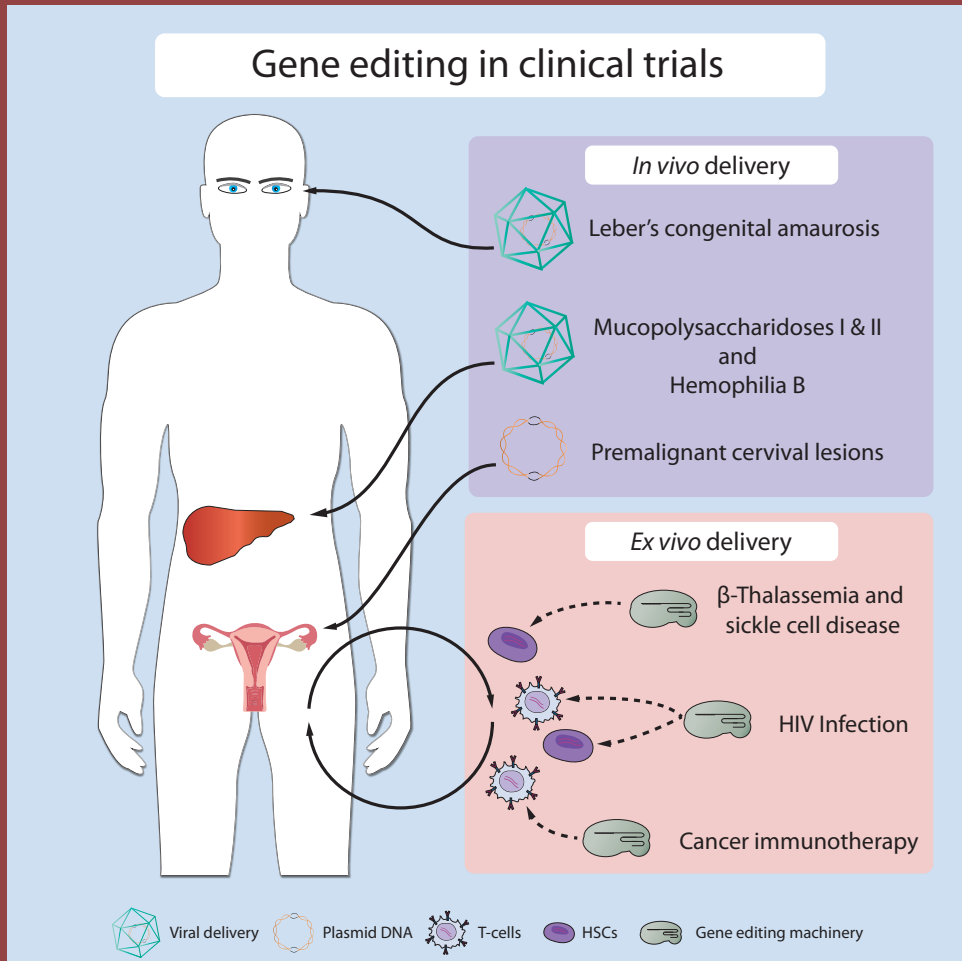
CHAPTER 5

READY FOR REPAIR? GENE EDITING ENTERS THE CLINIC FOR THE TREATMENT OF HUMAN DISEASE

Martijn P.T. Ernst, **Mike Broeders***, Pablo Herrero-Hernandez*, Esmee Oussoren,
Ans T. van der Ploeg, W.W.M. Pim Pijnappel

*These authors contributed equally to this work
Mol Ther Methods Clin Dev. 2020 Jul 3;18:532-557. doi: 10.1016/j.omtm.2020.06.022.
eCollection 2020 Sep 11.

Graphical abstract



keywords: gene editing; gene therapy; genome modification; medicine;
CRISPR-Cas; TALEN; ZFN; clinical trials

Molecular Therapy
Methods & Clinical Development

Ready for Repair? Gene Editing Enters the Clinic for the Treatment of Human Disease

Martijn P.T. Ernst,^{1,2,3} Mike Broeders,^{1,2,3,4} Pablo Herrero-Hernandez,^{1,2,3,4} Esme Oussoren,^{1,3}
Ans T. van der Ploeg,^{1,3} and W.W.M. Pim Pijnappel^{1,2,3}

¹Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands; ²Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands; ³Center for Lysosomal and Metabolic Diseases, Erasmus University Medical Center, 3015 GE Rotterdam, the Netherlands

We present an overview of clinical trials involving gene editing using clustered interspaced short palindromic repeats (CRISPR)-CRISPR-associated protein 9 (Cas9), transcription activator-like effector nucleases (TALENs), or zinc finger nucleases (ZFNs) and discuss the underlying mechanisms. In cancer immunotherapy, gene editing is applied *ex vivo* in T cells, transgenic T cell receptor (tTCR)-T cells, or chimeric antigen receptor (CAR)-T cells to improve adoptive cell therapy for multiple cancer types. This involves knockouts of immune checkpoint regulators such as PD-1, components of the endogenous TCR and histocompatibility leukocyte antigen (HLA) complex to generate universal allogeneic CAR-T cells, and CD7 to prevent self-destruction in adoptive cell therapy. In cervix carcinoma caused by human papillomavirus (HPV), E6 and E7 genes are disrupted using topically applied gene editing machinery. In HIV infection, the CCR5 co-receptor is disrupted *ex vivo* to generate HIV-resistant T cells, CAR-T cells, or hematopoietic stem cells. In β -thalassemia and sickle cell disease, hematopoietic stem cells are engineered *ex vivo* to induce the production of fetal hemoglobin. AAV-mediated *in vivo* gene editing is applied to exploit the liver for systemic production of therapeutic proteins in hemophilia and mucopolysaccharidoses, and in the eye to restore splicing of the CEP920 gene in Leber's congenital amaurosis. Close consideration of safety aspects and education of stakeholders will be essential for a successful implementation of gene editing technology in the clinic.

Conventional Gene Therapy

Traditionally, gene therapy relies on viral-based delivery of a protein-coding gene that either semi-randomly integrates into the genome (for retroviruses and lentiviruses) or remains as extrachromosomal DNA copy (for adeno-associated virus [AAV]).¹⁻³ These forms of gene therapy usually use overexpression of a protein that is missing or mutated in human disease. Lentiviral gene therapy has the advantage of being highly efficient and causing long-term efficacy. A drawback of lentiviral gene therapy is the lack of control of the location at which the virus integrates into the host genome, with the risk of insertional mutagenesis. By optimizing the lentiviral backbone and by controlling the number of viral copies, it has been demonstrated in multiple clinical trials that lentiviral gene therapy is safe provided that it is used with the proper precautions.^{2,4} AAV-mediated gene therapy does not rely on integration into the host genome but instead involves delivery of a

DNA episome to the nucleus. It is therefore considered to have a lower risk of genotoxicity compared to lentiviral gene therapy. However, episomal copies of AAV DNA are lost upon cell division, resulting in loss of efficacy. This restricts AAV gene therapy to nondividing cells. In addition, pre-existing immunity to AAV capsid proteins occurs in a significant percentage of the human population and precludes eligibility for the treatment.⁵ Acquired immunity after a single AAV-mediated gene therapy treatment occurs invariably in patients and precludes eligibility for a second treatment. In both forms of gene therapy, cDNA overexpression can only be used when dosage effects of the transgene product do not apply. Although the desired average number of gene copies can be approached via the viral titer, it is not possible to precisely control this using viral-based overexpression.

Basics of Gene Editing

Developments in recent years have enabled the seamless engineering of the human genome using a variety of tools collectively termed gene editing. Precision gene editing strategies allow alteration of the genome of cells at specific loci to generate targeted genomic changes, which are being exploited for multiple applications in medicine. We first introduce the basics of gene editing and then summarize the major challenges for their clinical implementation. Gene editing tools that are currently under investigation in clinical trials include zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs), and clustered interspaced short palindromic repeats (CRISPR) in combination with CRISPR-associated protein (Cas). For a detailed comparison between these tools, we refer to previously published reviews.^{6,7} In short, target site recognition occurs by sequence-specific DNA-binding proteins (in the case of ZFNs and TALENs) or by a short stretch of RNA termed single guide RNA (sgRNA; in the case of CRISPR-Cas). Current clinical applications of gene editing rely on the introduction of double-strand DNA breaks (DSBs), mediated by Fok-I (in the case of ZFNs or TALENs) or by Cas nucleases (in the case of CRISPR-Cas) and the introduction of desired genomic alterations through the cell's endogenous DNA repair mechanisms.

<https://doi.org/10.1016/j.omtm.2020.06.022>.

⁴These authors contributed equally to this work.

Correspondence: W.W.W. Pim Pijnappel, Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands.

E-mail: w.pijnappel@erasmusmc.nl

Molecular Therapy Methods & Clinical Development

Two major DNA repair pathways are being exploited to conduct targeted genomic changes in clinical trials: (1) gene editing through homology-directed repair (HDR) used to replace a pathogenic variant or insert foreign DNA elements to restore the wild-type (WT) expression of a missing (or truncated) gene; and (2) non-homologous end joining (NHEJ) used to remove DNA elements leading to aberrant expression of genes or to gain a therapeutic function.

In contrast to traditional strategies for gene therapy, gene editing provides more versatile tools for gene therapy, for example to precisely correct point variants,^{8,9} to place an extra, healthy gene copy at a safe genomic location of choice (a safe harbor: a location in the human genome at which integration of a gene is not harmful),^{10,11} or to disrupt a gene. This would, for example, enable the restoration of endogenous expression levels following precise correction of the disease-associated variant within the natural locus, which would be especially important for gene products for which a correct dosage is required. It would also increase control of integration sites of a cDNA by choosing appropriate safe harbor locations. Such locations also should provide efficient transcription of the transgene by providing a favorable epigenetic environment consisting of euchromatin. Examples of safe harbor locations in the human genome are the albumin, AAVS1, and the CCR5 loci.

On-Target or Off-Target?

Although the technology for gene editing is rapidly evolving, there are still important challenges for its clinical implementation. First, undesired editing of genomic regions can occur as a side effect of gene editing.⁷ This can be off-target, i.e., the introduction of a DNA break outside the genomic region of choice due to the targeting of the gene editing machinery to a chromosomal location that carries sequence similarity to the region of interest. In this scenario, genes or regulatory regions other than the targeted gene can be modified, resulting in undesired downstream effects. Undesired events may include insertions, deletions, and chromosomal translocations.^{12,13} Undesired variants can also be generated on-target, i.e., unintended modification of the genomic region of interest. In this scenario, regulatory elements within the gene of interest may be unintentionally changed. This may include elements involved in promoter activity, splicing, mRNA stability, protein translation, or microRNA (miRNA) genes (that are often present in introns or untranslated regions). The CRISPR-Cas9 system is inherently more prone to off-target effects compared to ZFNs or TALENs, because target site recognition in CRISPR-Cas9 relies on RNA-DNA interaction of only short stretches, and the RNA-DNA interaction allows some mismatches. In contrast, ZFNs and TALENs depend on highly specific protein-DNA interactions that allow fewer mismatches.¹⁴ This has promoted much research directed toward enhancing the performance of CRISPR-Cas-based gene editing with respect to specificity and nuclease activity (see below). Methods to detect undesired events in gene editing often rely on *in silico* predictions, followed by analyses of predicted off-target events. This is not necessarily sufficient for clinical application, and unbiased analysis based on next-generation sequencing is expected to become an important tool in the future. For a more exten-

sive discussion on off-target effects, see Broeders et al.,⁷ Kim et al.,¹⁵ Manghwar et al.,¹⁶ and Pattanayak et al.¹⁷

Delivery of Gene Editing

The delivery of gene editing tools is a crucial aspect when it comes to clinical implementation. Two routes can be distinguished: *ex vivo* and *in vivo* delivery.^{18,19} In *ex vivo* delivery, autologous or allogeneic cells are modified by gene editing outside the patient, and gene-modified cells are transplanted into the patient. Any route of administration of gene editing machinery can be applied *ex vivo*, such as transfection, nucleofection, or (viral) transduction. *Ex vivo* gene editing allows quality control prior to treatment. In particular, undesired off-target and on-target events can be monitored. Note that quality control can be performed on bulk generations of cells. Rare undesired events that occur in only a few cells and that might cause cellular transformation will be difficult to detect. Alternatively, this method involves an extra complication: the engraftment of (stem) cells. For example, maintaining engraftment potential and viability of the cell of interest can be challenging. Clinically, the most advanced forms of *ex vivo* gene editing involve T cells and hematopoietic stem cells (HSCs). In *in vivo* gene editing, gene editing tools are applied directly to the organism. Vehicles for delivery include AAV, lipid nanoparticles (LNPs), gold nanoparticles (GNPs), or cell-penetrating peptides (CPPs). The delivery method in *in vivo* gene editing is crucial for its safety.²⁰ When gene editing components are delivered *in vivo* via vehicles that remain present for an extended period, for example via AAV, there is a cumulative risk of undesired genotoxic events that can last for the time that the AAV remains present, which has been estimated to last for a period of 10 years or longer.¹ In contrast, when delivered as RNA or protein, there is only short-term exposure and a reduced risk of genotoxicity.

For *in vivo* gene editing, immunity against the delivery vehicle and the gene editing components are important considerations.²¹ Both pre-existing and acquired immunity should be considered. The AAV delivery vehicle is subject to pre-existing immunity in a significant proportion of the population.¹ In addition, preexisting immunity to Cas9 protein from several species has been reported in several studies. This may neutralize the therapy or induce adverse events.^{21–23}

In summary, the safety and efficacy of gene editing technology for the treatment of human disease depend on multiple factors, including the choice of the gene editing method, being either *ex vivo* or *in vivo*, the gene editing technique, target site selection, delivery method, and target tissue.

Gene Editing 2.0: Preclinical Developments

Technological developments are ongoing to improve gene editing tools with respect to specificity, efficiency, and versatility. These have been extensively described by us and others in recent reviews^{7,24–26} and are only briefly mentioned here.

First, variations of the original CRISPR-Cas9 method have been designed. These include the following: homology-independent targeted integration (HITI) for generating a knockin via NHEJ without

involvement of HDR;²⁷ microhomology-mediated end joining (MMEJ)-dependent knockin, which is based on the presence of short stretches of homology that are utilized by the MMEJ DNA repair pathway;²⁸ base editing;²⁹ a mismatch repair- or base excision repair-dependent pathway in which a natural cytidine or adenosine deaminase (ADA) is coupled to a catalytically dead Cas9 (dCas9) to convert cytidine to uridine (which is replicated as thymidine), or to convert adenine to inosine, which is replicated as guanine; and prime editing,³⁰ in which a Cas9 nicking variant is used that introduces single stranded DNA breaks and that is coupled to reverse transcriptase to enable a wide variety of genomic changes. Second, other natural and engineered Cas9 variants have been identified and developed with distinct and/or enhanced targeting properties, including Cas12a (Cpf1), Cas12b (C2c1), FokI fused to dCas9,³¹ Cas9-HF1,³² eSpCas9,³³ evoCas9,³⁴ and HypaCas9.³⁵ Third, Cas9 variants with distinct protospacer-adjacent motif (PAM) recognition sites have been generated, including VQR and VRER variants, xCas9, and SpCas9-NG.³⁶ And fourth, sgRNAs have been modified with respect to their length, structure, and chemistry to reduce off-target properties.^{37–39} These promising developments need future work to evaluate their suitability for clinical testing.

Scope of This Review

Whereas there have been numerous applications of gene editing in preclinical studies, information on clinical applications of gene editing is scattered in the literature. In this review, we present a comprehensive overview of current clinical trials using gene editing strategies for the treatment of human disease, and include selected preclinical examples. For more extensive overviews of preclinical studies, we refer to excellent reviews.^{40,41} In addition, in this review, we focus on gene editing in somatic cells, and we refer to other recent reviews and opinion articles for editing the germline.^{42–44} Thus far, precision gene editing has entered the clinic for the treatment of cancer immunotherapy, viral infections, and inherited hematologic, metabolic, and eye disorders (Table 1). These trials along with the underlying strategies are described in more detail below.

Gene Editing in Cancer Immunotherapy

Adoptive cell therapy (ACT) is a cellular form of cancer immunotherapy involving T cells with anti-tumor activity⁴⁵ that are expanded *ex vivo*, *ex vivo* genetically engineered or not, and applied to the patient via the circulation. Three major types of lymphocytes are used in ACT: (1) tumor-infiltrating lymphocytes (TILs), which are T cells that are isolated from tumors; and peripheral blood T lymphocytes that are (2) selected for tumor reactivity and expanded *ex vivo* before reinfusion or (3) genetically modified *ex vivo* with a transgenic T cell receptor (tTCR) or a chimeric antigen receptor (CAR) to target tumor cells.⁴⁶ ACT has been combined with *ex vivo* gene editing in a number of clinical trials, as discussed below.

Immune Checkpoint Knockout

Immune checkpoints are immune modulatory signals that can dampen the amplitude and quality of the immune response. Their

physiological function is to prevent overstimulation of the immune system in order to maintain self-tolerance. A hallmark of cancer cells is their ability to exploit immune checkpoints to evade attack by the immune system. Cancer cells or their microenvironment can achieve this by activating immune checkpoints via overexpression of ligands or receptors that regulate the function of T cells.^{47,48} In this way, cancer cells escape immune surveillance. To exploit this property of cancer cells for anti-cancer therapy, monoclonal antibodies have been developed that block natural immune checkpoints (present on T cells) or their ligands (present on cancer cells or in their microenvironment). This has revolutionized the field of anti-cancer therapy.⁴⁹ Examples include PD-1 and PD-L1 inhibitors, which have shown impressive results for treating different types of cancer at an advanced stage,^{50,51} especially melanoma.⁵² PD-1, encoded by the *PDCDI* gene, is a cell-surface receptor expressed on cytotoxic T cells that downregulates T cell activity upon interaction with its ligand PD-L1, which is overexpressed on malignant cells and cells in the tumor microenvironment.⁴⁸ In spite of general good tolerability, systemic administration of immune checkpoint inhibitors can result in autoimmune phenomena, referred to as immune-related adverse events (IRAEs).⁵³ IRAEs occur in up to 70% of patients receiving PD-1 and PD-L1 inhibitors^{50,51} and have been described in multiple organ systems. Steroids might be used to manage IRAEs, but the extent of interference with immunotherapy is unknown.⁵³

Knocking out immune checkpoint molecules in tumor-specific T cells is a promising strategy for ACT to circumvent systemic effects of checkpoint inhibition (Figure 1). When applied to total T cells harvested from patients, knocking out immune checkpoint molecules should render these less susceptible to immune inhibitory signals upon reinfusion. However, such an approach involves a heterogeneous T cell population rather than tumor-specific T cells. One solution to this problem would be to increase tumor specificity of circulating T cells *in vitro* by exposure to tumor-associated antigens.⁵⁴

Due to the impressive clinical results from checkpoint inhibitors and TILs to treat melanoma, this type of cancer was chosen in the initial preclinical studies on applying immune checkpoint knockout (KO) in ACT using *ex vivo* gene editing. Promising *in vitro* results were reported from co-cultures of human tumor-specific T cells in which PD-1 was disrupted with melanoma cell lines,^{55,56} and more recently by infusing PD-1 knockout T cells cells into mice that had been xenografted with human melanoma cells.⁵⁷ An improved cytotoxic effect of tumor-specific T cells following PD-1 knockout was also reported in preclinical studies of other cancer models, such as in a cultured gastric cancer cell line,⁵⁶ and in mice subcutaneously injected with either a fibrosarcoma cell line,⁵⁸ a multiple myeloma (MM) cell line,⁵⁹ or a liver cancer cell line.⁶⁰ Academic hospitals have been recruiting patients in clinical trials to investigate autologous, PD-1 knocked out T cells for the treatment of multiple types of cancer, including solid tumors arising from the esophagus,⁶¹ lung,⁶² prostate,⁶³ and Epstein-Barr-related neoplasms.⁶⁴ The publicly provided information is scarce. Presumably, as described for preclinical studies, these T cells have been manipulated *ex vivo* to enhance their tumor

Molecular Therapy
Methods & Clinical Development

| Table 1. Current Clinical Trials Involving Gene Editing | | | | | | |
|--|-------------|-------------------------|----------------------|----------|-------------|---------------------|
| Title | Tool | Status | Country | Delivery | ID | Ref. |
| Cancer Immunotherapy | | | | | | |
| PD-1 knockout engineered T cells for advanced esophageal cancer | CRISPR-Cas9 | completed | China | ex vivo | NCT03081715 | 61 |
| PD-1 knockout engineered t cells for metastatic non-small cell lung cancer | CRISPR-Cas9 | active, not recruiting | China | ex vivo | NCT02793856 | 62 |
| Therapeutic vaccine plus PD-1 knockout in prostate cancer treatment | CRISPR-Cas9 | recruiting | China | ex vivo | NCT03525652 | 63 |
| PD-1 knockout EBV-CTLs for advanced stage Epstein-Barr virus (EBV) associated malignancies | CRISPR-Cas9 | recruiting | China | ex vivo | NCT03044743 | 64 |
| CD19 CAR and PD-1 knockout engineered T cells for CD19 positive malignant B cell derived leukemia and lymphoma | N.S. | not yet recruiting | China | ex vivo | NCT03298828 | 82 |
| Study of PD-1 gene-knocked out mesothelin-directed CAR-T cells with the conditioning of PC in mesothelin positive multiple solid tumors | CRISPR-Cas9 | recruiting | China | ex vivo | NCT03747965 | 83 |
| CAR T and PD-1 knockout engineered T cells for esophageal cancer | N.S. | recruiting | China | ex vivo | NCT03706326 | 84 |
| Anti-MUC1 CAR T cells and PD-1 knockout engineered T cells for NSCLC | N.S. | recruiting | China | ex vivo | NCT03525782 | 85 |
| CRISPR (HPK1) edited CD19-specific CAR-T cells (XYF19 CAR-T Cells) for CD19+ leukemia or lymphoma | CRISPR-Cas9 | recruiting | China | ex vivo | NCT04037566 | 86 |
| Study of UCART19 in pediatric patients with relapsed/refractory B acute lymphoblastic leukemia (PALL) | TALEN | active, not recruiting | US/EU/UK | ex vivo | NCT02808442 | 103 |
| Dose escalation study of UCART19 in adult patients with relapsed/refractory B cell acute lymphoblastic leukaemia (CALM) | TALEN | active, not recruiting | US/EU/UK/Japan | ex vivo | NCT02746952 | 104 |
| Safety and efficacy of ALLO-501 anti-CD19 allogeneic CAR T cells in adults with relapsed/refractory large B cell or follicular lymphoma (ALPHA) | TALEN | recruiting | US | ex vivo | NCT03939026 | 105 |
| Safety and efficacy of ALLO-715 BCMA allogenic CAR T cells in in adults with relapsed or refractory multiple myeloma (UNIVERSAL) | TALEN | recruiting | US | ex vivo | NCT04093596 | 106 |
| A study to evaluate the long-term safety of patients with advanced lymphoid malignancies who have been previously administered with UCART19/ALLO-501 | TALEN | enrolling by invitation | US/EU/UK/Japan | ex vivo | NCT02735083 | 107 |
| A study evaluating UCART019 in patients with relapsed or refractory CD19+ leukemia and lymphoma | CRISPR-Cas9 | recruiting | China | ex vivo | NCT03166878 | 112 |
| A safety and efficacy study evaluating CTX110 in subjects with relapsed or refractory B cell malignancies | CRISPR-Cas9 | recruiting | US/Australia/Germany | ex vivo | NCT04035434 | 115 |
| A safety and efficacy study evaluating CTX120 in subjects with relapsed or refractory multiple myeloma | CRISPR-Cas9 | recruiting | US/Australia | ex vivo | NCT04244656 | 116 |
| CTA101 UCAR-T cell injection for treatment of relapsed or refractory CD19+ B cell acute lymphoblastic leukemia | CRISPR-Cas9 | recruiting | China | ex vivo | NCT04154709 | 117 |
| Phase I study of UCART22 in patients with relapsed or refractory CD22+ B cell acute lymphoblastic leukemia (BALLI-01) | TALEN | recruiting | US | ex vivo | NCT04150497 | 118 |

(Continued on next page)

Table 1. Continued

| Title | Tool | Status | Country | Delivery | ID | Ref. |
|--|----------------------|--------------------|---------|----------------|-------------|---------------------|
| CTA101 in the treatment of relapsed or refractory diffuse large B cell lymphoma | CRISPR-Cas9 | not yet recruiting | China | <i>ex vivo</i> | NCT04026100 | 119 |
| A feasibility and safety study of universal dual specificity CD19 and CD20 or CD22 CAR-T cell immunotherapy for relapsed or refractory leukemia and lymphoma | CRISPR-Cas9 | recruiting | China | <i>ex vivo</i> | NCT03398967 | 120 |
| Study evaluating safety and efficacy of UCART123 in patients with acute myeloid leukemia (AMEL-01) | TALEN | recruiting | US | <i>ex vivo</i> | NCT03190278 | 121 |
| Study evaluating safety and efficacy of UCART targeting CS1 in patients with relapsed/refractory multiple myeloma (MELANI-01) | TALEN | recruiting | US | <i>ex vivo</i> | NCT04142619 | 122 |
| Anti-CD19 U-CAR-T cell therapy for B cell hematologic malignancies | N.S. | not yet recruiting | China | <i>ex vivo</i> | NCT04264039 | 123 |
| Anti-CD7 U-CAR-T cell therapy for T/NK cell hematologic malignancies | N.S. | not yet recruiting | China | <i>ex vivo</i> | NCT04264078 | 124 |
| Efficacy and safety evaluation of BCMA-UCART | N.S. | recruiting | China | <i>ex vivo</i> | NCT03752541 | 125 |
| Safety and efficacy evaluation of CD19-UCART | N.S. | recruiting | China | <i>ex vivo</i> | NCT03229876 | 126 |
| The clinical study of CD19 UCAR-T cells in patients with B cell acute lymphoblastic leukemia (B-ALL) | N.S. | recruiting | China | <i>ex vivo</i> | NCT04166838 | 127 |
| NY-ESO-1-redirected CRISPR (TCRendo and PD1) edited t cells (NYCE T cells) | CRISPR-Cas9 | terminated | US | <i>ex vivo</i> | NCT03399448 | 133 |
| Study of CRISPR-Cas9 mediated PD-1 and TCR gene-knocked out mesothelin-directed CAR-T cells in patients with mesothelin positive multiple solid tumors | CRISPR-Cas9 | recruiting | China | <i>ex vivo</i> | NCT03545815 | 134 |
| Cell therapy for high risk T cell malignancies using CD7-specific CAR expressed on autologous T cells | CRISPR-Cas9 | not yet recruiting | US | <i>ex vivo</i> | NCT03690011 | 144 |
| Cervical Cancer | | | | | | |
| Study of molecular-targeted therapy using zinc finger nuclease in cervical precancerous lesions | ZFN | N.S. | China | <i>in vivo</i> | NCT02800369 | 160 |
| Study of targeted therapy using transcription activator-like effector nucleases in cervical precancerous lesions | TALEN | N.S. | China | <i>in vivo</i> | NCT03226470 | 161 |
| A safety and efficacy study of TALEN and CRISPR/Cas9 in the treatment of HPV-related cervical intraepithelial neoplasia | CRISPR-Cas9 TALEN | N.S. | China | <i>in vivo</i> | NCT03057912 | 162 |
| HIV Infection and AIDS | | | | | | |
| Autologous T cells genetically modified at the CCR5 gene by zinc finger nucleases SB-728 for HIV | ZFN | completed | US | <i>ex vivo</i> | NCT00842634 | 189 |
| Phase 1 dose escalation study of autologous t cells genetically modified at the CCR5 gene by zinc finger nucleases in HIV-infected patients | ZFN | completed | US | <i>ex vivo</i> | NCT01044654 | 190 |
| Repeat doses of SB-728mR-T after cyclophosphamide conditioning in HIV-infected subjects on HAART | ZFN | completed | US | <i>ex vivo</i> | NCT02225665 | 191 |
| A phase I study of T cells genetically modified at the CCR5 gene by zinc finger nucleases SB-728mR in HIV-infected patients | ZFN | completed | US | <i>ex vivo</i> | NCT02388594 | 192 |
| Dose escalation study of cyclophosphamide in HIV-infected subjects on HAART receiving SB-728-T | ZFN | completed | US | <i>ex vivo</i> | NCT01543152 | 193 |

(Continued on next page)

Molecular Therapy Methods & Clinical Development

Table 1. *Continued*

| Title | Tool | Status | Country | Delivery | ID | Ref. |
|--|-------------|-------------------------|-----------------|----------------|-------------|---------------------|
| CCR5-modified CD4 ⁺ T cells for HIV infection (TRAILBLAZER) | ZFN | recruiting | US | <i>ex vivo</i> | NCT03666871 | 194 |
| Study of autologous T cells genetically modified at the CCR5 gene by zinc finger nucleases in HIV-infected subjects | ZFN | completed | US | <i>ex vivo</i> | NCT01252641 | 195 |
| Long-term follow-up of HIV subjects exposed to SB-728-T or SB-728mR-T | ZFN | enrolling by invitation | US | <i>ex vivo</i> | NCT04201782 | 197 |
| Safety study of zinc finger nuclease CCR5-modified hematopoietic stem/progenitor cells in HIV-1 infected patients | ZFN | active, not recruiting | US | <i>ex vivo</i> | NCT02500849 | 203 |
| Safety of transplantation of CRISPR CCR5 modified CD34 ⁺ cells in HIV-infected subjects with hematological malignances | CRISPR-Cas9 | recruiting | China | <i>ex vivo</i> | NCT03164135 | 204 |
| CD4 CAR+ ZFN-modified T cells in HIV therapy | ZFN | active, not recruiting | US | <i>ex vivo</i> | NCT03617198 | 206 |
| β-thalassemia and Sickle Cell Disease | | | | | | |
| A safety and efficacy study evaluating CTX001 in subjects with transfusion-dependent β-thalassemia | CRISPR-Cas9 | recruiting | US/Canada/EU/UK | <i>ex vivo</i> | NCT03655678 | 263 |
| A study to assess the safety, tolerability, and efficacy of ST-400 for treatment of transfusion-dependent beta-thalassemia (TDT) | ZFN | active, not recruiting | US | <i>ex vivo</i> | NCT03432364 | 264 |
| A safety and efficacy study evaluating CTX001 in subjects with severe sickle cell disease | CRISPR-Cas9 | recruiting | US/Canada/EU | <i>ex vivo</i> | NCT03745287 | 265 |
| A study to assess the safety, tolerability, and efficacy of BIVV003 for autologous hematopoietic stem cell transplantation in patients with severe sickle cell disease (BIVV003) | ZFN | recruiting | US | <i>ex vivo</i> | NCT03653247 | 266 |
| A long-term follow-up study in subjects who received CTX001 | CRISPR-Cas9 | enrolling by invitation | US/EU | <i>ex vivo</i> | NCT04208529 | 267 |
| iHSCs with the gene correction of HBB intervent subjects with β-thalassemia mutations | CRISPR-Cas9 | not yet recruiting | N.S. | <i>ex vivo</i> | NCT03728322 | 280 |
| Hemophilia | | | | | | |
| Ascending dose study of genome editing by zinc finger nuclease therapeutic SB-FIX in subjects with severe hemophilia B | ZFN | active, not recruiting | US | <i>in vivo</i> | NCT02695160 | 289 |
| Mucopolysaccharidoses | | | | | | |
| Ascending dose study of genome editing by the zinc finger nuclease (ZFN) therapeutic SB-318 in subjects with MPS I | ZFN | active, not recruiting | US | <i>in vivo</i> | NCT02702115 | 319 |
| Ascending dose study of genome editing by the zinc finger nuclease (ZFN) therapeutic SB-913 in subjects with MPS II | ZFN | active, not recruiting | US | <i>in vivo</i> | NCT03041324 | 320 |
| Leber's Congenital Amaurosis | | | | | | |
| Single ascending dose study in participants with LCA10 | CRISPR-Cas9 | recruiting | US | <i>in vivo</i> | NCT03872479 | 329 |
| N.S., not specified. | | | | | | |

specificity, but this has not been specified. Recently, the results for PD-1-edited T cells in metastatic lung carcinoma patients were published.⁶⁵ Although no methods for increasing the tumor specificity of T cells was described, no severe adverse events were reported in 12 patients after a median follow-up time of 47.1 weeks. Despite the treatment, 10 patients progressed, and only 2 responded transiently.

Although not designed to investigate the therapeutic effect, these results were somewhat disappointing and are possibly caused by inadequate levels of tumor-specific T cells.

Another method of generating tumor-specific T cell clones is the *ex vivo* expansion of T cells that are isolated from tumor tissue,

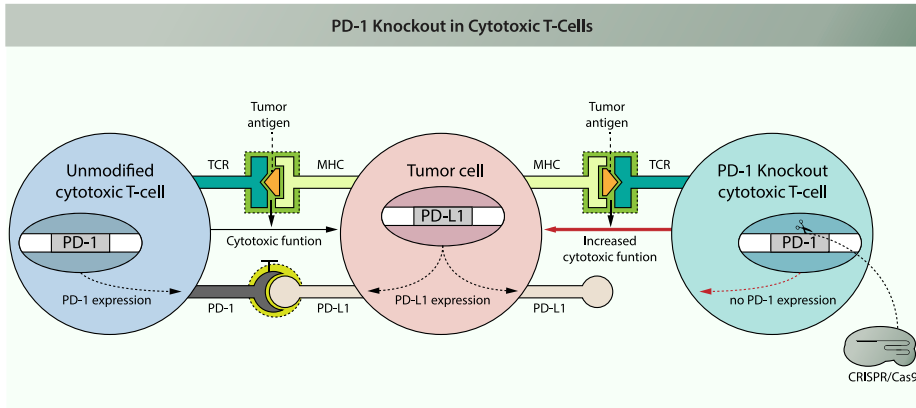


Figure 1. Effect of PD-1 Knockout in Cytotoxic T Cells

Cytotoxic T cells are able to recognize tumor cells via the T cell receptor (TCR). This receptor recognizes an antigen that is presented on potential target cells by the MHC. Binding results in T cell activation through signal transduction. The activated T cell will expand and exert its cytotoxic effector function on target cells, thus inducing apoptosis. If the target cell expresses PD-L1, it can interact with PD-1 that is expressed on the surface of the T cell. This will lead to activation of PD-1, one of the immune checkpoint molecules, resulting in inhibition of the T cell's cytotoxic activity. If PD-1 is disrupted in the cytotoxic T cell, PD-L1 expressed from the tumor cell can no longer interact with the T cell and inhibition of T cell cytotoxicity is prevented. In this scenario, PD-1 disruption prevents escape of tumor cells from attack by cytotoxic T cells. Red indicates the result of intervention.

so-called TILs. Although not yet clinically applied, PD-1 knockout in TILs has resulted preclinically in an improved anti-tumor effect *in vitro*⁵⁵ and *in vivo*.⁵⁸

Innate immune cells such as dendritic cells (DCs) and natural killer (NK) cells are also target cells for the development of immunotherapy against cancer.⁶⁶ Importantly, NK cells have also been shown to express several immune checkpoint inhibitors.⁶⁷ An example of recent preclinical developments is the knockout of the *NKp46* and *CIS* checkpoint genes in primary human NK cells.^{68,69} Although gene-edited innate immune cells have not yet reached clinical trials, these efforts illustrate the ongoing work that might promote their clinical development.

Immune Checkpoint Knockout in Genetically Engineered T Cells: tTCR-T and CAR-T cells

Besides the isolation of T cells with enhanced anti-tumor activity from patients, it is also possible to induce tumor specificity in T cells using genetic engineering (using viral transduction or gene editing). Such redirected T cells can be generated by forced expression of receptors with enhanced specificity for a tumor-associated antigen, such as tTCRs or CARs.^{70,71} tTCRs are transgenic forms of naturally occurring receptors isolated from tumor-specific T cells and depend on the major histocompatibility complex (MHC) for efficient antigen recognition.⁷² CARs are synthetic receptors that do not depend on MHC for efficient antigen binding.⁷³ To avoid negative regulation by tumors, immune checkpoint inhibition (using antibodies) or knock out (using gene editing) are also worthwhile strategies in tTCR-T cells and CAR-T cells.

The concept of immune checkpoint knockout in redirected T cells has been demonstrated *in vitro* and *in vivo*, both for tTCR-T cells⁷⁴ and CAR-T cells.^{75–79} Improved antitumor reactivity of redirected T cells after PD-1 disruption was observed in a range of preclinical cancer models, for example, models of melanoma,⁷⁴ hepatocellular carcinoma,⁷⁵ glioma,^{76,79} breast cancer,⁷⁷ and erythroleukemia.⁷⁸ In addition, encouraging clinical results have already been obtained by combining CAR-T cells with immune checkpoint inhibitors.^{80,81} Using gene editing, PD-1 knockout in CAR-T cells that were redirected against the B cell marker cluster of differentiation 19 (CD19)⁸² and membrane proteins mesothelin⁸³ and MUC1,^{84,85} which are upregulated in a range of malignancies, are investigated in clinical trials for the treatment of B cell leukemia/lymphoma,⁸² multiple mesothelin-positive solid tumors (such as pancreatic cancer, cholangiocarcinoma cancer, and ovarian cancer),⁸³ esophageal cancer,⁸⁴ and lung cancer.⁸⁵ One trial investigates the infusion of CAR-T cells carrying an HPK1 knockout in patients with relapsed or refractory CD19⁺ leukemia or lymphoma.⁸⁶ HPK1 is a protein kinase that was found to suppress the anti-tumor response of T cells by attenuating TCR signaling.⁸⁷ In addition, HPK1 exerts T cell inhibitory effects downstream of E prostanoic receptor activation by prostaglandin E2, a metabolic byproduct that is overproduced by cancers such as non-small-cell lung carcinomas.^{88,89} Mice with a kinase-dead HPK1 showed improved anti-tumor^{89,90} and antiviral responses.⁹⁰

Disruption of other molecules with immunomodulatory effects in ACT has been performed in preclinical studies, but no clinical trials are currently open. For example, infusion of cytotoxic T cells in which the immune checkpoint gene *CTLA-4* was disrupted resulted in decreased tumor growth compared to infusion of non-edited

Molecular Therapy

Methods & Clinical Development

counterparts in mice that were xenografted subcutaneously with bladder cancer cell lines⁹¹ or colon cancer cell lines.⁹² In addition, the anti-tumor effect of CAR-T cells against a human glioma cell line that was subcutaneously engrafted in mice was enhanced upon knockout of *DGK*,⁹³ which encodes an intracellular enzyme that negatively regulates TCR signaling.⁹⁴ In contrast, disruption of the immune checkpoint gene *LAG-3* in CAR-T cells did not result in an enhanced anti-tumor effect in mice subcutaneously engrafted with a human lymphoma cell line,⁹⁵ suggesting that the choice of immune checkpoint gene is important to design an efficient treatment.

Universal ACT

So far, we discussed autologous T cell therapies. However, this is not always feasible for every patient.⁹⁶ The establishment of universal, allogeneic ACT might be an attractive alternative, because such “off-the-shelf” therapy would overcome the high costs and experimental burden of manufacturing a custom-made autologous or histocompatibility leukocyte antigen (HLA)-matched allogeneic therapy for every patient. For such therapy, the risks of graft-versus-host disease (GvHD) and graft rejection by the patients’ immune system for universal ACT must be addressed. The strategies used involve knockout of the TCR to prevent GvHD, and knockout of human leukocyte antigen (HLA) genes to prevent graft rejection by the host immune system.^{97,98} Clinical studies and preclinical examples are discussed below.

In vitro studies showed that anti-CD19 CAR-T cells, which target B cells, tolerated ZFN-mediated knockout of the TCR, as assessed by cell proliferation and their ability to lyse target cells.⁹⁹ In addition, *in vivo* application of such cells demonstrated an anti-leukemic response in mice that were intravenously injected with a lymphoma cell line that was similar or better compared to non-edited cells.^{100,101} The feasibility of clinical implementation of such a strategy was illustrated by a study in which two therapy-refractory pediatric patients with acute lymphoblastic leukemia (ALL) were treated with allogeneic anti-CD19 CAR-T cells from unselected donors¹⁰² that had been engineered *in vitro* using TALENs in two ways. First, expression of the endogenous $\alpha\beta$ TCR was disrupted by targeting the constant region of the TCR α chain. Second, *CD52* was knocked out with the following rationale. *CD52* is expressed on T cells, and anti-*CD52* antibodies (alemtuzumab) are part of the conditioning regimen prior to allogeneic HSC transplantation to reduce the risk of graft rejection by the host’s lymphocytes. To prevent alemtuzumab from attacking anti-CD19 CAR-T cells, these cells were made resistant by knockout of *CD52*. Despite development of grade 2 GvHD in one of the patients, the results of this trial indicated an ongoing disease-free survival of the two patients of 12 and 18 months after the start of therapy.¹⁰² These results suggest that off-the-shelf allogeneic CAR-T cells therapy is feasible, and that adverse events such as GvHD are manageable. This exact strategy is adopted in clinical trials investigating universal CAR-T cells in pediatric or adult B cell ALL,^{103,104} B cell lymphoma,¹⁰⁵ and MM patients.¹⁰⁶ The long-term effects of two of these products are investigated in a separate trial.¹⁰⁷

To reduce the risk of graft rejection by the host immune system, HLA genes have been disrupted in donor T cells.^{108–111} Notably, CRISPR-

Cas9-mediated triple KO of the T cell receptor α constant (*TRAC*) locus, an HLA complex gene (*B2M*), and an immune checkpoint gene (*PDCD1*) was used to potentiate the anti-tumor effect of CAR-T cells against multiple targets in mouse models, for example in mice intravenously injected with a B cell ALL cell line,¹⁰⁹ in mice intraperitoneally injected with a lymphoma cell line,¹¹⁰ and in mice intracerebrally injected with a glioma cell line.¹¹¹ In one active clinical trial both the endogenous TCR and HLA complex are knocked out in anti-CD19 CAR-T cells for treating of B cell leukemia and lymphoma.¹¹²

In another concept, a tumor-targeting CAR or tTCR is inserted into the *TRAC* locus using CRISPR-Cas9-mediated HDR. This yields two effects: knockout of the endogenous TCR, and knockin of the CAR/tTCR. In a preclinical study, a CD19-directed CAR was inserted into the *TRAC* locus in human T cells by HDR using CRISPR-Cas9.¹¹³ When these CAR-T cells were administered to a mouse model of ALL, an improved anti-leukemic response was observed that resulted in prolonged survival compared to conventionally generated CAR-T cells.¹¹³ A similar strategy proved feasible for inserting a tTCR directed against the immunogenic cancer antigen NY-ESO-1 in the *TRAC* locus.¹¹⁴ This strategy is adopted in two clinical trials for patients with B cell malignancies¹¹⁵ or MM,¹¹⁶ in which the endogenous TCR is disrupted by knockin of an anti-CD19 or anti-BCMA CAR in the TCR locus of allogeneic T cells, respectively. In addition, the HLA complex is disrupted by knockout of the *B2M* gene.

Additional clinical studies are planned, in which infusion of universal CAR-T cells (engineered using TALENs or CRISPR-Cas9) will be investigated for the treatment of B cell ALL or lymphoma,^{117–120} acute myeloid leukemia (AML),¹²¹ and multiple myeloma.¹²² No molecular details are provided for these trials. Five more clinical trials are active or planned that will investigate universal CAR-T cells in hematological malignancies, but no information on the applied gene editing platform has been provided.^{123–127}

A challenging application in one of the aforementioned trials is the treatment of AML with ACT, because molecular targets of leukemic cells in AML are also expressed in HSCs. As a result, ACT will attack the host’s HSCs and impair hematopoiesis.¹²⁸ Indeed, severe myelotoxicity, leading to prolonged pancytopenia, was seen in preclinical studies using CAR-T cells directed at CD33¹²⁹ and CD123.¹³⁰ One possible strategy to circumvent this problem would be to co-transplant HSCs in which the target molecule is knocked out together with the CAR-T cells. As the CAR-T cells will attack the leukemic cells and unmodified recipient HSCs, the gene-edited donor HSCs will not be targeted anymore and will repopulate the bone marrow. This strategy has been proven feasible in a mouse model for AML, in which anti-CD33 CAR-T cells along with CD33-edited HSCs were used.¹³¹ The leukemic cells responded to anti-CD33 CAR-T cell treatment, while myelotoxicity was selectively mitigated in mice transplanted with CD33-edited HSCs. An ongoing clinical trial investigates universal CAR-T cells in refractory or relapsed AML, but it does not include a method to mitigate the possible myelotoxic effect of CAR-T cells.¹²¹

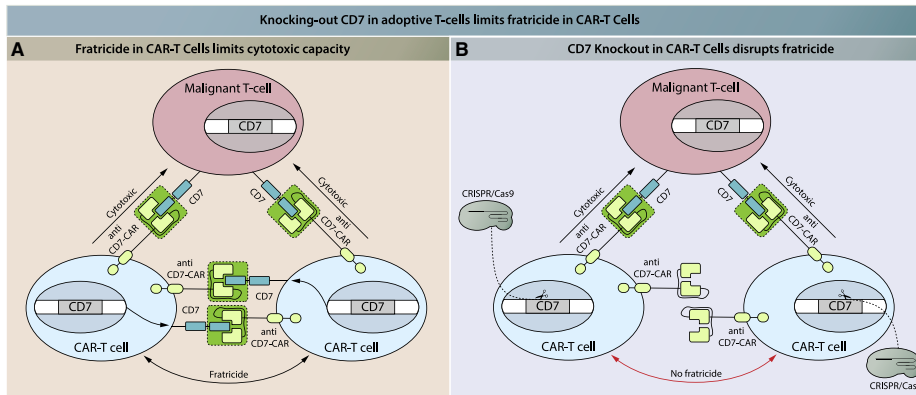


Figure 2. CD7 Knockout in Anti-CD7 CAR-T Cells Prevents Fratricide

(A) Anti-CD7 CAR-T cells recognize the CD7 antigen on (malignant) T cells via their chimeric antigen receptor, which triggers the CAR-T cell cytotoxic function and thus results in lysis of the target cell. CD7 is expressed on the surface of all T cells. As CAR-T cells also express CD7, CAR-T cells will recognize other CAR-T cells and lyse these, which is termed fratricide. (B) The gene encoding CD7 can be knocked out in anti-CD7 CAR-T cells, for example by CRISPR-Cas9. Without CD7, these CAR-T cells will not be recognized and lysed by other anti-CD7 CAR-T cells, thus preventing fratricide. Red indicates the result of intervention.

Endogenous TCR Knockout in Autologous ACT

Above we described the knockout of the endogenous TCR in allogeneic ACT products to prevent GvHD. However, there is also a rationale for knocking out endogenous TCR components in autologous ACT. The reason for this is that the endogenous TCR can interfere with the function of the tTCR/CAR, either by competing for cell surface expression, or by dimerization to form a novel hybrid compound TCR that might cause autoimmune reactions.¹³² Knockout of endogenous TCR components in autologous ACT cells is therefore adopted in two clinical trials with either tTCR-T cells redirected against NY-ESO-1 (in MM, melanoma, or subtypes of sarcoma)¹³³ or CAR-T cells redirected against mesothelin (in any mesothelin-positive solid tumor).¹³⁴ PD-1 is also knocked out in the tTCR-T cells and CAR-T cells in these trials. Initial results of the first trial have been published, and they indicated no major adverse events in the three patients that were included.¹³⁵ The patients suffered from advanced refractory malignancies, and the response to therapy was variable: one patient did not respond and died, while two patients showed initial disease stabilization, followed by disease progression after 30 or 100 days. Responses to follow-up treatment in these two patients were variable. Interestingly, the authors reported a relatively long half-life of tTCR-T cells at an average of 83.9 days. As other studies reported a half-life of roughly 1 week of non-edited NY-ESO-1 tTCR-T cells,^{136–138} the knockout of PD-1 and/or endogenous TCR components might have contributed to a slower decay of the tTCR-T cells.

A Special Case: ACT for T Cell Malignancies

It is particularly challenging to design an effective ACT using T cells for T cell malignancies. T cells should target molecules that are preferably expressed by malignant T cells but not by normal T cells. The difficulty in finding specific targets in malignant T cells results in self-

destruction of tTCR-T cells or CAR-T cells used in ACT.¹³⁹ This process, called fratricide, can interfere with ACT efficacy and has been observed in both CAR-T cells¹⁴⁰ and transgenic TCR-T cells.¹⁴¹ One possible solution to this problem is to knockout the target molecule in the adoptive T cells by gene editing. In this way, transgenic TCR-T or CAR-T cells will recognize and attack malignant T cells, but not each other. This strategy has been proven effective in circumventing fratricide in a preclinical setting,^{142,143} and it is currently applied in a clinical trial applied to CD7. CD7 is expressed on the cell surface of T cells, and in this trial anti-CD7 CAR-T cells are tested for the treatment of T cell leukemia/lymphoma. To prevent fratricide, CD7 was knocked out in CAR-T cells using CRISPR-Cas9 (Figure 2).¹⁴⁴ In addition, one previously mentioned clinical trial investigates universal anti-CD7 CAR-T cells in T cell malignancies, but knockout of CD7 in the CAR-T cells has not been mentioned.¹²⁴

Gene Editing in Viral Infection

Cervical Cancer

Cervical cancer is the third most prevalent type of cancer in women worldwide.¹⁴⁵ The most contributing etiological factor is human papillomavirus (HPV) infection via sexual intercourse, especially serotypes HPV-16 and HPV-18. Most HPV infections are cleared by the host immune system, but persistent infections can give rise to malignant transformation. Several vaccines have been developed for primary prevention of cervix carcinoma, with varying levels of population coverage worldwide.¹⁴⁶ Premalignant lesions are treated by local excision, while therapeutic modalities for invasive cervix carcinoma are dependent on the cancer stage and include surgery, radiotherapy, and chemotherapy.¹⁴⁷ In spite of these preventive and curative modalities, survival rates of cervical cancer range from 93% at early disease stage to 15% at disseminated disease stage.¹⁴⁸ New

Molecular Therapy Methods & Clinical Development

treatment modalities are crucial to increase survival rates of cervix carcinoma.

One such recent advance is RNA interference (RNAi)-mediated knockdown of the viral oncogenes E6 and E7, as these have been identified to drive and sustain HPV-related carcinogenesis.¹⁴⁹ In multiple studies, knockdown of E6 and E7 resulted in increased cell death in HPV-positive cell lines.^{150,151} However, multiple obstacles, such as the occurrence of escape mechanisms and insufficient efficiency, have prohibited RNA-based strategies from entering clinical trials so far.¹⁵² Guided gene knockout might partially overcome these limitations. First of all, RNAi only lowers target gene expression, whereas gene editing can completely disrupt or delete a gene, leaving no room for residual gene expression. Mutation of the target region, a known escape mechanism of RNA viruses, as observed in studies using RNAi-mediated knockdown, likely still applies to knockout strategies using gene editing. Another escape mechanism, which is expression of viral suppressors of RNAi, is expected not to apply to gene editing.¹⁵³ Investigating viral escape from strategies involving gene editing in cervical cancer caused by HPV will be an important aspect in future research. As is true for any cancer, it will be important to start treatment at the earliest stage possible and to use treatments that are highly efficient.

Gene editing for treating HPV infection has focused on E6 and E7. It is generally appealing to target viral genes, because these are exogenous sequences, reducing the chances of unintended off-target events in endogenous genes. Successful knockout of E6 and E7 genes has been achieved via ZFNs,¹⁵⁴ TALENs,^{155,156} and CRISPR-Cas9.^{157–159} The *in vitro* knockout of viral E6 or E7 sequences in HPV-infected cell line models caused inhibition of cell growth and cell viability, which is in line with results obtained from RNAi. In addition, gene-edited cells showed reduced capability to engraft in mice compared to unedited cells when transplanted subcutaneously.^{154,155,157} Results were consistent for targeting HPV-16 and HPV-18.¹⁵⁵ Furthermore, *in vivo* gene editing with topically applied TALEN components using polymer-complexed T512 plasmids in K14-HPV16 transgenic mice, a model system for cervical HPV-16 infection, resulted in reduced viral DNA loads and a reversal of histological malignant abnormalities.¹⁵⁵ As only the TALEN platform was topically applied *in vivo* in a cervical cancer mouse model,¹⁵⁵ the effects of topically applied gene editing tools on cervical cancer could not be compared. Based on these results, multiple clinical trials have been designed to investigate gene editing of precancerous cervical lesions, directed at the HPV genome. These clinical trials apply either ZFN,¹⁶⁰ TALEN,^{161,162} or CRISPR-Cas9¹⁶² gene editing platforms, which are administered either by topical gel or vaginal suppository.

In the future, topically applied gene editing tools might be investigated in combination with chemotherapy in metastasized cervix carcinoma. Preclinically, an additive anti-cancer effect of gene editing was already shown *in vitro* and *in vivo* in combination with cisplatin.¹⁶³ In addition, the potential of HPV targeting extends beyond the treatment of cervix carcinoma, as HPV-related cancers include other anogenital cancers such as vulvar, vaginal, anal,

and penile cancer, but also cancers in the head and neck region.¹⁶⁴ In preclinical studies, CRISPR-Cas9-based strategies have been tested for treating other chronic viral infections, such as hepatitis B virus,^{165–172} Epstein-Barr virus,^{173–176} and human immunodeficiency virus (HIV) (see section below). As these viral infections affect distinct tissues and/or have distinct modes of action, these might need tailored strategies for delivery to the required target. An overview of these gene editing strategies is provided in a review by de Buhr and Lebbink.¹⁷⁷

Gene Editing in HIV Infection and AIDS

HIV is a lentivirus that integrates its genome (after reverse transcription of its RNA into DNA) into the genome of host CD4⁺ T helper cells, forming a provirus. After the initial acute phase of infection, a pool of T cells remains latently infected. When the provirus becomes activated, host cells produce new viral particles and undergo cell death. This causes acquired immunodeficiency syndrome (AIDS) if the numbers of T helper cells drop to levels that are insufficient to effectively protect the host from infections or malignant transformations.¹⁷⁸ Currently, HIV infections are treated by antiretroviral therapy (ART) to reduce the risk of progression to AIDS. However, ART needs to be taken life-long, requires adherence to the treatment regimen, and can have side effects and incomplete efficacy.^{179,180} Although no curative treatment has been found to date, there are two documented cases of HIV patients who have been cured from HIV infection. The first patient, known as the Berlin patient, received two HSC transplantations for AML, and has remained HIV-negative since.^{181,182} His donor harbored a homozygous *CCR5* Δ32/Δ32 loss-of-function allele, which had previously been known to impair infection of T cells by HIV-1.¹⁸³ A similar second patient was identified recently.¹⁸⁴ In addition, genetic association studies have shown that *CCR5*Δ32 homozygotes are resistant to HIV infection, whereas heterozygotes display delayed progression of disease.^{185–187} It was therefore hypothesized that *ex vivo* disruption of *CCR5* in patient-derived T cells, followed by reinfusion, could mimic the curative outcome of the Berlin patient. *CCR5* was targeted by ZFNs in human primary CD4⁺ T cells, and biallelic gene disruption was achieved in 33% of modified cells *in vitro*.¹⁸⁸ In an HIV infection mouse model, injection of *CCR5* KO T cells resulted in decreased viral load and an increased T cell population compared to wild-type T cells.¹⁸⁸ Six out of a total of seven clinical trials assessing the infusion of autologous CD4⁺ *CCR5* knockout T cells using ZFNs have been completed,^{189–195} and results of one have been published.¹⁹⁶ In the study of Tebas et al.,¹⁹⁶ CD4⁺ *CCR5* KO T cell infusion proved to be safe in HIV patients. In addition, levels of blood HIV DNA decreased in most patients, although the trial was not designed to measure efficacy. One clinical trial is currently investigating the long-term effects of *CCR5*-edited T cells.¹⁹⁷

It is unclear how long engineered T cells can in principle protect against AIDS given their limited lifespan. Therefore, several groups are focusing on deleting *CCR5* in HSCs, as these have self-renewal capacity to remain present as stem cells and can give rise to all cells of

the hematopoietic lineage.¹⁹⁸ HSCs would for example also give rise to CD4⁺ myeloid cells, which are also susceptible to HIV infection.¹⁹⁹ CCR5 disruption by ZFNs was achieved in human CD34⁺ HSCs, and these cells were able to engraft in immunosuppressed or immunodeficient mice.^{200–202} In addition, infusion of CCR5-modified HSCs resulted in reduced plasma HIV levels in mouse models when compared to unmodified HSC infusions.²⁰² Currently, two clinical trials are recruiting patients to test this strategy using either ZFN²⁰³ or CRISPR-Cas9.²⁰⁴

The previous strategies involve supplying patients with HIV-resistant cells to diminish the effect of HIV on the immune system. Alternatively, CAR-T cells that are redirected toward HIV-related proteins can be applied to actively attack T cells that are infected by the virus.²⁰⁵ Via gene editing, CCR5 might be disrupted in the CAR-T cells to prevent HIV from infecting these cells. Multiple clinical trials are planned or ongoing for CAR-T cells as a treatment option for HIV. In one of those, ZFNs are applied to disrupt CCR5 in CAR-T cells.²⁰⁶

CCR5 disruption will not be efficacious in all patients, since CCR5 might be redundant for cell entry by certain HIV strains.^{207,208} Another disadvantage is the necessity of biallelic knockout of CCR5 to efficiently impair viral reproduction.^{198,209} An alternative is disruption of the HIV genome itself, which may be especially attractive since this is not an endogenous sequence and may therefore be less susceptible to off-target effects. Targeted disruption of the HIV genome, however, faces the challenge of mutational escape. Another challenge is that HIV-1 forms a stable reservoir in resting CD4⁺ T cells, which sustains the disease and causes the residual viremia in patients undergoing ART.²¹⁰ If the latent reservoir could be directly targeted or activated, HIV infection could possibly be cured without the requirement of myeloablative therapy and subsequent HSC transplantation. Multiple proof-of-principle studies have shown the feasibility of targeting HIV genomic sequences in infected cells *in vitro*,^{211–219} but the problems of mutational escape and targeting the HIV latent reservoir have not been solved to date.²²⁰

Alternatively, the strategies mentioned above could be realized via RNAi. CCR5 knockdown by short hairpin RNA (shRNA) in HSCs or T cells has been readily tested in preclinical studies and is the subject of a phase I/II clinical trial.²²¹ Targeting of HIV transcripts by RNAi has also been tested preclinically.²²¹ Besides mutational escape mentioned above, RNAi faces the additional challenge of transcriptional upregulation of the target in response to knockdown.²²²

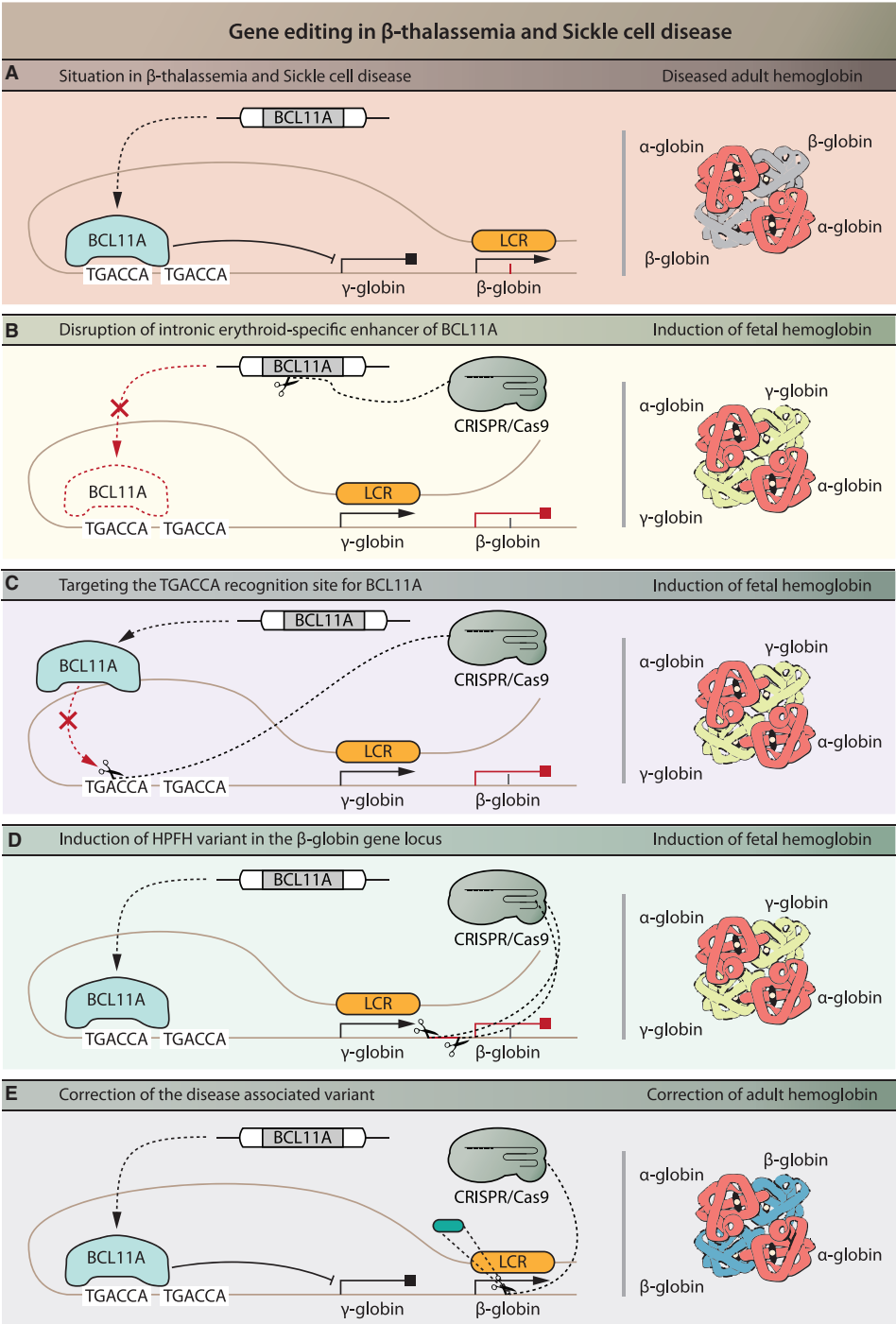
Exciting preclinical studies have shown the feasibility for applying gene editing to the engineering of B cells that produce antibodies specific to a number of viruses, including Rous sarcoma virus (RSV), influenza virus, Epstein-Barr virus (EBV), or HIV, all of which are viruses for which there is to date no vaccine available. In the example of HIV, broad neutralizing antibodies (bNAbs) have been detected in a small number of infected individuals at ~1–3 years after infection.²²³ These NAb protect against HIV infection. Primary human B cells have been successfully engineered using CRISPR-Cas9 to produce

NAb against HIV,²²⁴ and a proof of principle using engineered mouse B cells provided protection against infection with RSV.²²⁵

Gene Editing in Hematological Disorders *β-thalassemia and Sickle Cell Disease*

β-thalassemia is an autosomal recessive disease with more than 200 known disease-associated variants in the gene coding for the hemoglobin β chain (*HBB*), resulting in a clinically variable phenotype. All of these variants cause reduced or abolished translation of the HBB protein.²²⁶ Approximately 98% of total adult hemoglobin is composed of hemoglobin A (HbA), which is formed by two β-globin subunits bound to two α-globin subunits.²²⁷ Reduced expression of the β-globin subunit results in a relative excess of the α-globin subunit, resulting in precipitation of the α-globin subunit in erythroblasts and erythrocytes. This ultimately leads to impaired erythropoiesis and hemolysis, and thus anemia.

Treatment of β-thalassemia depends on life-long supportive measures, of which blood transfusion is the main component. β-Thalassemia patients either have transfusion-dependent thalassemia (TDT) or non-TDT (NTDT).²²⁸ TDT patients require life-long blood transfusions for survival, starting at an average age of 2 years for every 2–5 weeks, while NTDT patients need blood transfusions only occasionally or for limited periods of time.²²⁸ Regular transfusions place patients at risk of blood-borne infections, iron overload, and transfusion reactions.²²⁹ In addition, 80% of TDT patients develop long-term complications.²³⁰ Although long-term complications due to iron overload result in decreased longevity, a life expectancy of over 50 years of age has been estimated.²³¹ Recurrent therapy, adverse events, and complications also negatively impact patients quality of life. Furthermore, treatment of β-thalassemia patients with iron chelation therapy is essential to reduce iron overload, but it results in considerable additional costs. In addition, through alloimmunization, it becomes increasingly challenging to find eligible blood products.²²⁹ The only curative therapy to date is allogeneic HSC transplantation, provided that a suitable donor is available. An HLA-matched sibling donor is available in about 30% of cases.²³² For the remaining patients an unrelated HLA-matched donor should be considered, which approaches success rates of sibling donors. However, for 20%–30% of patients needing an HSC transplantation (without considering the underlying disease), no optimal unrelated HLA-matched donor can be found even with the extensive donor registries that have been established in Europe.²³³ For 5% of patients, no donor could be identified at all. The alternative of cord blood transplantation from unrelated donors, for which HLA matching is less stringent, is less favorable due to higher rates of graft failure.²³⁴ Haploidentical, or half-matched (e.g., parents or children), stem cell transplantation seems inferior to HLA-matched unrelated transplantation due to delayed restoration of the immune system, although experience is limited.²³² Between 2000 and 2010, the European Society for Blood and Bone Marrow Transplantation Hemoglobinopathy Registry reported treatment outcomes for all HSC transplantations, showing a 2-year event-free survival rate of more than 80% in TDT patients. However, this study also revealed a 12% overall mortality within 2



(legend on next page)

years after allogeneic HSC transplantation and the required (myeloablative) conditioning. In addition, 10% of patients developed severe acute GVHD, and about 15% of patients developed chronic GVHD.²³⁵

In sickle cell disease (SCD), the β -globin subunit in HbA carries a point variant that results in the formation of an aberrant form termed hemoglobin S. The *HBB* p.Glu6Val variant in combination with the same or a second *HBB* disease-associated variant on the second allele leads to SCD, in which erythrocytes are malformed, resulting in chronic hemolytic anemia. The malformed erythrocytes can cause acute ischemia throughout the body due to obstruction of blood vessels, leading to (severe) pain, organ failure, and severe acute vaso-occlusive complications such as acute chest syndrome or stroke,²³⁶ which can be treated by exchange transfusion.²³⁷ With this therapy, the patients' blood is exchanged with donor blood to lower the percentage of sickle cells. Chronic transfusions are performed in patients with a history of stroke to prevent new cerebral ischemic events.²³⁷ Possible complications of frequent transfusions have been described previously. Frequently hospitalized patients, for example due to acute chest syndrome or the need for intravenous analgesics in the management of acute pain, are treated with hydroxyurea. These treatments, hospital admissions, acute complications, and many more chronic complications result in reduced life quality of patients.²³⁷ As in β -thalassemia, allogeneic HSC transplantation is the only cure for SCD. Although HSC transplantation with a product of a related HLA-matched donor seems successful in most cases, severe complications as described previously are also seen in SCD.²³⁸ Recent improvements in conditioning regimens have led to reduced intensity treatment without short-term GVHD, but serious adverse events still occurred.^{239,240} The experience with other HSC transplantation sources is scarce in SCD, but it seems inferior to related HLA-matched donors.^{238,241}

Curative options that are less toxic than allogeneic HSC transplantation are required for both β -thalassemia and SCD. As gene therapy allows the engineering of autologous stem cells, the need for a donor would be bypassed. Importantly, transfusion of autologous rather than allogeneic stem cells strongly reduces the HSC transplantation-related toxicity.²⁴² Reports of gene therapy using lentiviral vectors to add a healthy *HBB* copy to HSCs *in vitro* for reinfusion purposes have been published for β -thalassemia²⁴³ and SCD,²⁴⁴ and the first promising (interim) results of clinical trials have been reported.^{245,246} As the graft must replenish the hematopoietic system through rapid cell

division, an integrative vector, such as lentiviral vectors, is required. Although γ -retroviral vectors used in the past gave rise to leukemia through insertional mutagenesis,²⁴⁷ currently used third-generation self-inactivating lentiviruses have an improved safety profile and have been used without adverse events in several clinical trials up to 7 years follow-up.^{248–255} Because lentiviral transduction is highly efficient, it provides a strong competitor for gene editing approaches in strategies involving overexpression of transgenes.

The main strategy under current investigation for clinical application of gene editing is the induction of endogenous expression of fetal hemoglobin (HbF). This originated from the observation that co-inheritance of hereditary persistence of HbF (HPFH), a benign condition, reduces symptoms of SCD and β -thalassemia.²²⁷ The situation in SCD and β -thalassemia is depicted in Figure 3A. In HPFH, HbF protein production continues into adulthood, whereas under normal physiological conditions production shifts to adult hemoglobin after birth. HbF protein contains two subunits of α -globin and γ -globin each, the latter of which are translated from the *HBB* gene. Persistent HbF expression in HPFH compensates for the reduced production of HbA in β -thalassemia patients. There is a difference in HbF protein levels among β -thalassemia patients, and this has been linked to several genetic variants, with single nucleotide variants (SNVs) in the *BCL11A* gene correlating most strongly with HbF expression.²⁵⁶ Reduced *BCL11A* protein expression is correlated with increased HbF protein expression, likely because *BCL11A* suppresses HbF expression by binding directly to the *HBB* promoter.^{257,258} *BCL11A* null mice were shown to be unable to downregulate murine embryonic globin in erythrocytes, demonstrating the essential role of *BCL11A* in repression of HbF expression during development.²⁵⁹ However, *BCL11A* knockdown by gene disruption in HSCs results in impaired engraftment of HSC in mice, illustrating that knockout of *BCL11A* itself is not a feasible strategy to treat β -thalassemia.²⁶⁰ As *BCL11A* expression during erythropoiesis is specifically regulated by the intronic erythroid-specific enhancer,²⁶¹ disrupting this enhancer will result in *BCL11A* knockout during erythropoiesis, exclusively. This strategy was preclinically tested by using ZFN-mediated gene disruption of the GATAA element of the *BCL11A* erythroid-specific enhancer in HSCs (Figure 3B).^{260,262} These cells achieved robust long-term engraftment in mice and gave rise to erythroid cells with elevated HbF levels upon *ex vivo* culture of chimeric bone marrow.²⁶⁰ Multiple clinical trials are based on a

Figure 3. Gene Editing Strategies in β -Thalassemia and Sickle Cell Disease

(A) Situation in β -thalassemia and sickle cell disease. The locus control region (LCR) loops to the β -globin gene and β -globin is expressed; however, due to a disease-associated variant in the β -globin gene there is insufficient expression (β -thalassemia) or malformed (sickle cell disease) β -globin. The transcriptional repressor *BCL11A* recognizes the first TGACCA binding sequence, which leads to inhibition of expression of fetal-specific γ -globin. (B) In one strategy, CRISPR-Cas9 or ZFNs (not shown) are used for targeted disruption of the GATAA motif in the intronic erythroid-specific enhancer of *BCL11A*, which will result in disruption of *BCL11A* expression during erythropoiesis and consequently relief of inhibition of γ -globin expression. γ -Globin will substitute for the lack of β -globin to form functional hemoglobin: HbF. (C) In a related strategy, the TGACCA recognition site for *BCL11A* is disrupted using CRISPR-Cas9 or ZFNs (not shown). *BCL11A* remains expressed but cannot bind to the recognition site to inhibit the γ -globin expression, resulting in relief of inhibition of γ -globin expression. (D) In another scenario, the β -globin promoter sequence is disrupted using CRISPR-Cas9, leading to a loss of binding sites for proteins that repress expression of γ -globin and subsequent induction of γ -globin expression. (E) Finally, the disease-associated variant can be precisely corrected using CRISPR-Cas9. While strategies in (B), (C), and (D) will lead to the induction of fetal hemoglobin, the strategy in (E) will lead to production of adult hemoglobin. Red indicates the result of intervention.

strategy involving HSCs, of which the intronic erythroid-specific enhancer of BCL11A is disrupted *ex vivo* using CRISPR-Cas9 or ZFNs as a treatment for TDT^{263,264} or SCD^{265,266} patients. The long-term effects of infusing such cells are investigated in one clinical trial.²⁶⁷ Other strategies to increase HbF expression include disruption of the binding motif for BCL11A (and co-repressive proteins) within the HBG promoter sequence (Figure 3C)^{257,268,269} or the induction of a natural occurring variant termed the Sicilian HPFH disease-associated variant (Figure 3D).²⁷⁰ In the latter variant, the entire β -globin locus is deleted and the putative 3' β -globin enhancer is brought in closer proximity to the γ -globin locus. These strategies have been explored preclinically, but have not (yet) reached clinical application.

Besides induction of HbF, other applications of gene editing techniques to treat β -thalassemia and SCD have been tested mainly in preclinical studies. Cai et al.²⁷¹ showed an approach to correct various *HBB* disease-associated variants by inserting a cDNA sequence of exons 2 and 3 of the *HBB* gene downstream of *HBB* exon 1 *in vitro* using CRISPR-Cas9 in induced pluripotent stem cells (iPSCs). This strategy ensured expression of correct β -globin and prevented expression of the mutated variant in iPSC-derived erythrocytes. Other preclinical studies showed the (HDR-mediated) correction of a specific disease-associated variant in (iPSC-induced) HSCs to restore β -globin and thus HbA expression (Figure 3E).^{272–279} One clinical trial implies to investigate the infusion of autologous, iPSC-induced HSCs with a directly gene-corrected version of the *HBB* gene in β -thalassemia.²⁸⁰ However, very limited information is provided for this trial and the exact strategy is not elucidated. Another preclinical strategy involves the *in vitro* knockout of α -globin²⁸¹ to prevent its precipitation. No clinical trial has been reported that investigates this option.

Gene Editing in Hemophilia

Hemophilia A and B are congenital bleeding disorders caused by deficiencies in clotting factor VIII (FVIII) or IX (FIX), respectively. These diseases have a recessive X-linked inheritance pattern. Protein substitution therapy (PST) with recombinant clotting factor or protein derived from donor plasma is currently the main treatment for these patients.²⁸² Despite the steep increase in life expectancy and the improved prevention of arthropathies due to articular bleedings after introduction of PST, this treatment has its drawbacks.^{282,283}

As substituting a deficient protein is not curative, repeated administration is required and patients remain at risk of bleedings. In addition, costs related to PST are considerable.²⁸² Insertion of a functional copy of the deficient gene in patient cells could potentially provide a long-term cure for hemophilia. To this end, *in vivo* gene therapy by viral vectors has been applied in multiple phase I clinical trials,^{284–}

²⁸⁶ as well as by *ex vivo* electroporation of fibroblasts that provided a source of FVIII after engraftment.²⁸⁵ Initial results observed in these clinical trials were disappointing.²⁸² For lentiviral transduction, preclinical optimization of *ex vivo* HSC-mediated lentiviral gene therapy is paving the way for the first clinical studies.²⁸² In spite of subclinical effects of targeting muscle cells by AAV vectors in hemophilia, prom-

ising clinical results have been obtained by the use of AAV vectors targeting liver cells.²⁸² Transient liver toxicity and a temporary requirement for immunosuppressive therapy were drawbacks of this strategy.

Currently, gene editing strategies to target liver cells are also being explored for hemophilia. Sharma et al.²⁸⁸ achieved robust expression of human FVIII or FIX by integrating the cDNA of either gene into intron 1 of the albumin locus in primary hepatocytes *in vitro* and in hepatocytes of mice *in vivo* by using AAV-delivered ZFN-mediated gene editing. Despite the low *in vivo* genome editing efficiency, gene expression was achieved by placing the genes under the control of the highly active albumin promoter. This *in vivo* gene editing strategy is currently being investigated in hemophilia B patients in a clinical trial.²⁸⁹ A drawback for clinical implementation of such strategy is the long-term presence of active gene editing components in the liver of patients and the associated risk of damaging the genome by introducing double-stranded breaks at off-target loci. This is a serious concern, as the gene editing machinery delivered by AAV has an expected presence in the liver of several years, which significantly increases the chance for off-target effects to occur. This highlights the need for developing more transient ways to perform *in vivo* gene editing.

Other preclinical strategies that are under investigation include insertion of the transgene into the AAVS1 locus^{290,291} or in the native locus^{292–295} and correction of disease-associated variants^{296–298} or large chromosomal rearrangements.^{299–301}

For the clinical translation of gene editing in HSCs, a critical aspect is to maintain long-term engraftment capacity.^{201,302–305} Similar to most other cells, HDR-mediated gene editing is challenging in HSCs, as these cells prefer the NHEJ pathway. In addition, it has been found that genetic manipulation of HSCs with gene editing or viral vectors can reduce their engraftment capacity. This has been found to be caused by activation of the DNA damage response pathway, resulting in activation of p53. Transient inhibition of p53 has been found to improve long-term engraftment of HSCs after gene editing.³⁴⁴ In addition, technical optimizations related to cell culture, delivery, and use of reagents have resulted in enhanced long-term engraftment of HSCs after gene editing in xenograft experiments involving transplantation of human HSCs into immunodeficient mice. The clinical testing of long-term engraftment of gene-edited HSCs in human patients needs further testing.

Gene Editing in Metabolic Disorders

Mucopolysaccharidoses

Mucopolysaccharidoses (MPSs) are monogenic lysosomal storage diseases (LSDs) in which one of the enzymes involved in the lysosomal degradation of glycosaminoglycans (GAGs) is deficient. In MPS I and II, this concerns the α -L-iduronidase (IDUA) and iduronate-2-sulfatase (IDS) enzymes, respectively. Patients suffer from multisystemic symptoms and reduced life expectancy that can vary depending on the type of MPS and the severity of the

disease-associated variant.³⁰⁶ The currently available treatment for MPS I, MPS II, MPS IVA, MPS VI, and MPS VII is enzyme replacement therapy (ERT), in which recombinant enzyme is administered intravenously. Drawbacks of ERT include the non-curative nature of the treatment, the requirement of repeated intravenous infusions, high costs, and ineffectiveness in treating symptoms in bone, cartilage, heart valves, and the central nervous system.^{307,308} In addition, generation of antibodies against the recombinant enzyme can interfere with the efficacy of ERT.³⁰⁷ HSC transplantation is currently applied to treat MPS I.³⁰⁹ This relies on the principle that lysosomal enzymes are secreted and can be taken up by target cells via the cation-independent mannose 6-phosphate receptor (CI-M6PR). In HSC transplantation, HSCs and their progeny secrete the enzyme into the circulation and provide a continuous source of ERT. In the case of MPS I, the level of secretion and reuptake provides partial efficacy in target organs. However, HSC transplantation depends on the availability of HLA-matched donors and can have severe adverse events such as GVHD, infection, and even death, as described before.³⁰⁶ In addition, the therapeutic effect on the skeletal abnormalities and neurological symptoms is limited, and for many other LSDs, endogenous expression levels in HSCs are insufficient to treat target organs. Therefore, overexpression by *ex vivo* lentiviral transduction or gene editing provides (additional) therapeutic efficacy. For MPS I, liposome-mediated delivery of CRISPR-Cas9 has been successfully applied *in vivo* and resulted in increased IDUA expression in newborn MPS I mice.³¹⁰ Alternatively, direct gene addition using AAV vectors (without gene editing) has been shown feasible in preclinical studies for several MPS types.^{311–316} This strategy is being investigated in multiple clinical trials, and recent results using intracerebral delivery showed promising outcomes with respect to treating the neurological decline of MPS IIIB patients.³¹⁷

Another approach, similar to the approach in hemophilia, is the site-specific integration of a transgene in the liver by *in vivo* genome editing following intravenous administration using AAV as the delivery method.²⁸⁸ Most efforts have been made on integrating transgenes into the albumin locus. Sharma et al.²⁸⁸ achieved ZFN-mediated insertion of IDUA and IDS *in vivo* into the albumin locus of healthy mice, resulting in detectable protein levels in liver lysates. More recently, ZFN-mediated insertion of human IDS in the albumin locus in murine liver *in vivo* was accompanied by a dose-dependent rise in circulating enzyme levels.³¹⁸ This IDS insertion caused reduction of GAG levels in tissue and urine samples of MPS II mice. These results have led to clinical trials investigating the safety of ascending dose levels of AAV vectors containing components required for *in vivo* ZFN-mediated insertion of IDUA and IDS genes into the albumin locus of hepatocytes in the liver of MPS I patients³¹⁹ and MPS II patients,³²⁰ respectively. The same drawbacks as in the hemophilia trial with respect to safety due to the potential introduction of double-stranded breaks in the liver at off-target locations in the genome apply here due to the long-term exposure of the patient to the uncontrolled activity of ZNF-mediated double-stranded breaks.

Gene Editing in the Eye

Leber's Congenital Amaurosis

Leber's congenital amaurosis (LCA) is an inherited retinopathy in which severe visual impairment or blindness occurs within the first months of life.³²¹ It is a genetically heterogeneous disease that can be caused by any of more than 20 mutated genes. Based on the genes involved and the ocular phenotypes, LCA is divided into 13 subtypes.³²² Currently, there is no treatment for LCA. In clinical trials, it has already been shown that AAV-mediated gene transfer by subretinal injection resulted in improved visual parameters in patients with the LCA type LCA2, which is caused by variants in the *RPE6* gene.^{323–326} Retinal dystrophy in LCA was (at least partially) reversed by the therapy. AAV-mediated gene therapy has also been applied to other congenital retinopathies.³²⁷

In addition to AAV-mediated gene transfer, gene editing is in development for retinopathies. For subtype LCA10, which is caused by variants in the *CEP290* gene,^{321,328} a clinical trial is currently open³²⁹ with the strategy outline below. Gene transfer via viral vectors (especially AAV) is problematic for *CEP290* due to the large gene size. *CEP290* encodes a protein that is essential for cilia, which are microtubule-based, hair-like extensions of cell membranes.³³⁰ In photoreceptor cells, cilia are highly specialized into cone- or rod-shaped segments that act as light sensors and signal transducers.³²² In LCA10, *CEP290* disease-associated variants cause (peripheral) thickening of the retina by an unknown mechanism.³³⁰ The most common variant is the intronic variant IVS26, which results in the generation of a cryptic splice site that causes an abrogated protein product.³²⁸ Preclinical studies had shown that, using subretinal injections of AAV5 vectors containing the CRISPR-Cas9 gene editing machinery, it deletion of the cryptic splice site leads to restoration of canonical splicing and expression of wild-type protein.^{331,332} This concept is used in the ongoing clinical trial.³²⁹ Other preclinical studies are investigating gene editing strategies for other disease-associated variants in LCA and other retinopathies.^{333,334} However, long-term expression of CRISPR-Cas9 in the eye imposes safety risks, as discussed in approaches for *in vivo* gene editing in hemophilia and MPS I and II.

Conclusions and Future Prospects

Disease-Specific Challenges

The challenges and opportunities of applying gene editing for the treatment of human disease depend in part on disease-specific aspects. In cancer immunotherapy, a major challenge is to specifically target cancer cells while leaving healthy cells unharmed. Targeting immune checkpoints with gene editing has been shown to be a promising strategy, but the clinical feasibility relies in part on the inherent problem of specificity: by inhibiting a general checkpoint with the aim to inhibit negative immune regulation, there is a risk of auto-immune-related side effects. Considering the life-threatening nature of cancer, this disadvantage may be acceptable if survival rates can be improved and increased toxicity is manageable. Other challenges include the viability of T cells that have been gene edited *ex vivo* to knock out immune checkpoint regulators. These cells do not need

Molecular Therapy

Methods & Clinical Development

to be present life-long, but they should have sufficient viability in order to help eliminating cancer cells. If needed, repeated administration would be an option, but this will increase costs. The development of a universal ACT would be an elegant solution to the high costs of preparing autologous or HLA-matched allogeneic gene-edited T cells for each individual patient, although this approach has the risk of inducing GvHD.

Application of targeted knockout to viral infection such as HPV could provide a useful additional treatment option when it comes to treating the primary tumor. However, a high efficiency of gene knockout is required to effectively reduce the tumor, and treating metastases is not yet possible due to the difficulty to reach target tissues and to eliminate the HPV virus in a safe and efficient manner outside the primary tumor. It might be an advantage to target viral sequences rather than endogenous genomic locations to reduce the risks of undesired genomic alterations as the result of gene editing. This could also be a potential advantage for strategies that eradicate HIV provirus from the genome. In the case of HIV, disease-specific challenges include the targeting of the dormant HIV reservoir, and to target HIV strains that do not depend on CCR5 for infection.

Ex Vivo Gene Editing

In both genetic disease and viral infection, promising strategies using *ex vivo* gene editing lie ahead for disorders that can be cured via the blood, including hematological disorders, lysosomal storage disorders, and HIV infection. The main reason for this is the feasibility to target blood cells such as HSCs or T cells *ex vivo* and to engraft autologous gene-modified cells back into patients. This approach relies on the long-standing experience with successful engraftment of HSCs, which has nowadays become a standard procedure with a very good safety profile. In addition, engrafted HSCs can provide a life-long treatment because they can self-renew to sustain the stem cell population and to differentiate into the hematopoietic lineage. Because prolonged *ex vivo* culture reduces engraftment potential and stem cell properties of HSCs, fast and efficient methods are required to make *ex vivo* gene editing of HSCs feasible for clinical implementation. It remains to be seen whether *ex vivo* gene editing for overexpressing proteins will be able to successfully compete with *ex vivo* lentiviral transduction of HSCs when it comes to clinical implementation, because lentiviral transduction is highly efficient, has been used more than 7 years without adverse events in several clinical trials, and could be more cost-effective.^{248–255} Strategies that rely on NHEJ are inherently more efficient compared to the HDR-mediated insertion of transgenes, and these provide promising options for the treatment of HIV infection, by knocking out the CCR5 receptor, or some genetic disorders such as β -thalassemia and SCD, by knocking out regulatory elements required for BCL11A-mediated negative regulation of HbF expression.

Among the many other preclinical developments for using *ex vivo* gene-edited HSCs (not covered in this review), the primary immunodeficiency diseases (PIDs) represent a promising example.^{335,336} These patients usually benefit from allogeneic HSC transplantation

from HLA-matched donors, but these are not always available, and allogeneic HSCs can induce GvHD. Autologous HSC transplantation following *ex vivo* gene therapy employing third-generation lentiviruses is ongoing in a number of clinical trials for Wiskott-Aldrich syndrome, ADA severe combined immunodeficiency (ADA-SCID), X-linked SCID, and chronic granulomatous disease (CGD). However, many PIDs involve genes with a timed and restricted expression pattern during development and require endogenous expression levels via the natural promoter rather than overexpression. Gene editing would be advantageous above lentiviral transduction in these cases, as it enables precise correction of an endogenous allele to maintain endogenous expression levels. There are currently no clinical trials for PIDs reported using gene editing, but promising preclinical developments may change this in the near future.

Other promising preclinical developments include the *ex vivo* gene editing of primary hepatocytes for metabolic disease of the liver. As a proof of concept, AAV-mediated delivery of CRISPR-Cas9 to freshly isolated mouse hepatocytes was used, followed by engraftment into the liver of a mouse model. This concept was applied to treat hereditary tyrosinemia in a mouse model to correct a point variant in the fumarylacetoacetate hydrolase gene using HDR.³³⁷ A major challenge for this approach is the limited engraftment capacity of hepatocytes in human liver. In cystic fibrosis, investigators are pursuing gene editing of stem cells derived from the airways with the ultimate goal of developing a gene-edited autologous airway stem cell transplantation.³³⁸ Mitochondrial diseases that are caused by disease-associated variants in mitochondrial DNA form an attractive target for gene editing.^{339,340} However, gene editing of mitochondrial DNA is even more challenging than nuclear DNA, and improvements are required before clinical applications can be considered in the near future.

In all of these possible applications, the *ex vivo* mode of gene editing ensures that a quality control can be performed prior to decision-making of engrafting cells into patients. Reliable methods to assess undesired genomic alterations are essential, and a shift from methods that rely on predictions toward unbiased methods will be required. Quality control should also include functional analysis of cellular transformation, because rare events that result in formation of tumorigenic cells will be very difficult to detect in population-based assays.

In Vivo Gene Editing

In vivo gene editing trials have already started for a number of disorders including lysosomal storage disorders, hemophilia, precancerous cervical lesions and LCA. This is despite the uncertainties of gene editing technology with respect to possible off-target effects. This is particularly important when gene editing technology is introduced in patients without ways for spatiotemporal control (i.e., means to confine gene editing to a short time and specific target tissue, for example by using suicide genes in DNA combined with tissue-specific delivery, or local administration of gene editing tools as RNA/protein rather than DNA), such as is the case in trials so far. This means that gene editing may continue for years inside the patient, which will increase the risk of undesired events with several orders of magnitude

compared to *ex vivo* gene editing. For safe future clinical development, it will be important to develop ways that can control the activity of *in vivo* gene editing by including on and off switches to prevent the prolonged generation of DNA breaks or by providing the gene editing machinery in other ways than as DNA. In addition, targeting gene editing tools specifically to the cells of interest will further increase the safety by preventing unnecessary targeting events in irrelevant cell types.

These aspects will also guide ongoing preclinical efforts to develop treatments for human disease based on *in vivo* gene editing. Multiple preclinical developments in different fields are ongoing, and it is beyond the scope of this review to cover these. As examples we mention metabolic disorders that are amenable to correction via knockout of a gene in the metabolic pathway to enable redirecting of metabolism. For example, severe hereditary tyrosinemia type I was successfully redirected to a more benign tyrosinemia type III form by deletion of the upstream metabolic enzyme hydroxyphenylpyruvate dioxygenase in the liver. The method applied involved intravenous injection of DNA encoding Cas9 and sgRNAs in the mouse, which transected the liver more efficiently compared to other organs.³⁴¹ The same gene was also targeted in a study on *in utero* correction of hereditary tyrosinemia type I using injection of an adenovirus expressing a base editor and sgRNA into mouse fetuses via the vitelline vein. In the same study, *in utero* knockout of PCSK9 was achieved with the aim to lower cholesterol levels and the risk of coronary heart disease in wild-type mice.

Precise correction of a point variant *in vivo* has been demonstrated for example in a mouse model for phenylketonuria (PKU) using base editors that were delivered by intravenous injection and that were expressed via a liver-specific promoter.³⁴² Gene editing is even applied in preclinical research to increase the fitness of pig organs for future xenotransplantation into humans. By knockout of genes that activate an immune response and retroviral elements, the aims are to generate organs with reduced hazard of graft rejection and xenozoonosis (an infectious disease transmitted from animal to human), respectively.³⁴³ The ultimate goal of these efforts is to overcome the shortage of human organs such as kidneys, hearts, livers, and lungs for transplantation.

Keeping Up with Technological Developments

Finally, it will be important to educate the various stakeholders, including clinicians, patients, and regulatory institutions. The technology for gene editing is moving so fast that it is difficult to cope with all of the developments and their potential benefits and risks. Clinicians need to be educated in order to allow them to judge the feasibility of a clinical trial and whether they are willing to expose their patients to the novel treatment. Patients rely largely on the information that is provided by their treating physician. The prospect of a “cure” via gene repair may be tempting for a patient, and therefore providing balanced and fair information by the physician on the possible benefits and risks provides an essential ingredient for decision-making. The same arguments apply to regulatory institutions, as these will approve or decline clinical protocols and finally market authorization. While the scientific developments in the field of gene editing are

continuing with dazzling speed, it will be important to provide education in the field and to closely monitor and regulate clinical developments.

In this review, we compiled all current clinical applications of gene editing and explained the rationale for the underlying strategies. In addition, we summarized preclinical studies that preceded clinical trials and provided examples of preclinical work that might be translated in a clinical setting in the future. As most other reviews focus on specific areas involving gene editing applications, we envision that centralized information on gene therapies will increase awareness of clinicians and researchers in the field of gene therapy outside their specific field of interest, and that this might catalyze new developments. We propose that clinical applications of gene editing in general will be documented in an accessible and transparent manner. We hope that this review precedes the discussion of a central database that includes relevant information of the clinical studies applying gene editing, as well as the underlying considerations with respect to the mechanism of action, safety, and expected results. Ideally, this information should be contributed by investigators involved in these clinical trials, peer-reviewed by experts in the field, and made publicly available prior to the start of such trials. Preferably, an analysis of risks and benefits of gene editing for a specific disease in the context of current treatments should be included, contributing to discussions on technical and ethical aspects of the applications. Such efforts should contribute to increasing transparency and help to inform stakeholders that are involved in clinical trials involving gene editing.

AUTHOR CONTRIBUTIONS

M.P.T.E., P.H.-H., M.B., and W.W.W.P.P. conceptualized this review, performed literature studies, and wrote the manuscript. All authors interpreted the contents and approved the final manuscript.

CONFLICTS OF INTEREST

A.T.v.d.P. has provided consulting services for various industries in the field of Pompe disease under an agreement between these industries and Erasmus MC, Rotterdam, the Netherlands. The remaining authors declare no competing interests.

ACKNOWLEDGMENTS

We thank Anita Rijnveld for reading and commenting on the original manuscript. This work has received funding from Texnet, Zeldzame Ziekten Fonds/WE Foundation, Metakids (project number 2018-082), and Stofwisselkracht. The collaboration project is co-funded by the PPP Allowance made available by Health~Holland, Top Sector Life Sciences & Health, to the Prinses Beatrix Spierfonds to stimulate public-private partnerships (project numbers LSHM17075 and LSHM19015).

REFERENCES

1. Colella, P., Ronzitti, G., and Mingozzi, F. (2017). Emerging issues in AAV-mediated *in vivo* gene therapy. *Mol. Ther. Methods Clin. Dev.* 8, 87–104.
2. Naldini, L. (2015). Gene therapy returns to centre stage. *Nature* 526, 351–360.

Molecular Therapy Methods & Clinical Development

3. Shirley, J.L., de Jong, Y.P., Terhorst, C., and Herzog, R.W. (2020). Immune responses to viral gene therapy vectors. *Mol. Ther.* 28, 709–722.
4. Rainov, N.G., and Ren, H. (2003). Clinical trials with retrovirus mediated gene therapy—what have we learned? *J. Neurooncol.* 65, 227–236.
5. Ronzitti, G., Gross, D.A., and Mingozzi, F. (2020). Human immune responses to adeno-associated virus (AAV) vectors. *Front. Immunol.* 11, 670.
6. Gaj, T., Gersbach, C.A., and Barbas, C.F., 3rd (2013). ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31, 397–405.
7. Broeders, M., Herrero-Hernandez, P., Ernst, M.P.T., van der Ploeg, A.T., and Pijnappel, W.W.M.P. (2020). Sharpening the molecular scissors: advances in gene-editing technology. *iScience* 23, 100789.
8. Komor, A.C., Kim, Y.B., Packer, M.S., Zuris, J.A., and Liu, D.R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424.
9. Gaudelli, N.M., Komor, A.C., Rees, H.A., Packer, M.S., Badran, A.H., Bryson, D.I., and Liu, D.R. (2017). Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature* 551, 464–471.
10. Sadelain, M., Papapetrou, E.P., and Bushman, F.D. (2011). Safe harbours for the integration of new DNA in the human genome. *Nat. Rev. Cancer* 12, 51–58.
11. van der Wal, E., Herrero-Hernandez, P., Wan, R., Broeders, M., In 't Groen, S.L.M., van Gestel, T.J.M., van IJcken, W.F.J., Cheung, T.H., van der Ploeg, A.T., Schaaf, G.J., and Pijnappel, W.W.M.P. (2018). Large-scale expansion of human iPSC-derived skeletal muscle cells for disease modeling and cell-based therapeutic strategies. *Stem Cell Reports* 10, 1975–1990.
12. Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* 31, 827–832.
13. Kosicki, M., Tombberg, K., and Bradley, A. (2018). Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* 36, 765–771.
14. Cornu, T.I., Mussolino, C., and Cathomen, T. (2017). Refining strategies to translate genome editing to the clinic. *Nat. Med.* 23, 415–423.
15. Kim, D., Luk, K., Wolfe, S.A., and Kim, J.S. (2019). Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. *Annu. Rev. Biochem.* 88, 191–220.
16. Manghwir, H., Li, B., Ding, X., Hussain, A., Lindsey, K., Zhang, X., and Jin, S. (2020). CRISPR/Cas systems in genome editing: methodologies and tools for sgRNA design, off-target evaluation, and strategies to mitigate off-target effects. *Adv. Sci. (Weinh.)* 7, 1902312.
17. Pattanayak, V., Guilinger, J.P., and Liu, D.R. (2014). Determining the specificities of TALENs, Cas9, and other genome-editing enzymes. *Methods Enzymol.* 546, 47–78.
18. Yin, H., Kauffman, K.J., and Anderson, D.G. (2017). Delivery technologies for genome editing. *Nat. Rev. Drug Discov.* 16, 387–399.
19. Lino, C.A., Harper, J.C., Carney, J.P., and Timlin, J.A. (2018). Delivering CRISPR: a review of the challenges and approaches. *Drug Deliv.* 25, 1234–1257.
20. Tong, S., Moyo, B., Lee, C.M., Leong, K., and Bao, G. (2019). Engineered materials for in vivo delivery of genome-editing machinery. *Nat. Rev. Mater.* 4, 726–737.
21. Li, A., Tanner, M.R., Lee, C.M., Hurley, A.E., De Giorgi, M., Jarrett, K.E., Davis, T.H., Doerfler, A.M., Bao, G., Beeton, C., and Lagor, W.R. (2020). AAV-CRISPR gene editing is negated by pre-existing immunity to Cas9. *Mol. Ther.* 28, 1432–1441.
22. Charlesworth, C.T., Deshpande, P.S., Dever, D.P., Camarena, J., Lemgart, V.T., Cromer, M.K., Vakulskas, C.A., Collingwood, M.A., Zhang, L., Bode, N.M., et al. (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nat. Med.* 25, 249–254.
23. Crudele, J.M., and Chamberlain, J.S. (2018). Cas9 immunity creates challenges for CRISPR gene editing therapies. *Nat. Commun.* 9, 3497.
24. Pickar-Oliver, A., and Gersbach, C.A. (2019). The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* 20, 490–507.
25. Moon, S.B., Kim, D.Y., Ko, J.H., and Kim, Y.S. (2019). Recent advances in the CRISPR genome editing tool set. *Exp. Mol. Med.* 51, 1–11.
26. Carroll, D. (2017). Genome editing: past, present, and future. *Yale J. Biol. Med.* 90, 653–659.
27. Suzuki, K., and Izpisua Belmonte, J.C. (2018). In vivo genome editing via the HITI method as a tool for gene therapy. *J. Hum. Genet.* 63, 157–164.
28. Chang, H.H.Y., Pannunzio, N.R., Adachi, N., and Lieber, M.R. (2017). Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* 18, 495–506.
29. Molla, K.A., and Yang, Y. (2019). CRISPR/Cas-mediated base editing: technical considerations and practical applications. *Trends Biotechnol.* 37, 1121–1142.
30. Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., and Liu, D.R. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157.
31. Tsai, S.Q., Wyvekens, N., Khayter, C., Foden, J.A., Thapar, V., Reyon, D., Goodwin, M.J., Aryee, M.J., and Joung, J.K. (2014). Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* 32, 569–576.
32. Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., and Joung, J.K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* 529, 490–495.
33. Slaymaker, I.M., Gao, L., Zetsche, B., Scott, D.A., Yan, W.X., and Zhang, F. (2016). Rationally engineered Cas9 nucleases with improved specificity. *Science* 351, 84–88.
34. Casini, A., Olivieri, M., Petris, G., Montagna, C., Reginato, G., Maule, G., Lorenzin, F., Prandi, D., Romanel, A., Demicheli, F., et al. (2018). A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36, 265–271.
35. Chen, J.S., Dagdas, Y.S., Kleinstiver, B.P., Welch, M.M., Sousa, A.A., Harrington, L.B., Sternberg, S.H., Joung, J.K., Yildiz, A., and Doudna, J.A. (2017). Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature* 550, 407–410.
36. Kleinstiver, B.P., Prew, M.S., Tsai, S.Q., Topkar, V.V., Nguyen, N.T., Zheng, Z., Gonzales, A.P.W., Li, Z., Peterson, R.T., Yeh, J.R.J., et al. (2015). Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature* 523, 481–485.
37. Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32, 279–284.
38. Kocak, D.D., Josephs, E.A., Bhandarkar, V., Adkar, S.S., Kwon, J.B., and Gersbach, C.A. (2019). Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37, 657–666.
39. Yin, H., Song, C.Q., Suresh, S., Kwan, S.Y., Wu, Q., Walsh, S., Ding, J., Bogorad, R.L., Zhu, L.J., Wolfe, S.A., et al. (2018). Partial DNA-guided Cas9 enables genome editing with reduced off-target activity. *Nat. Chem. Biol.* 14, 311–316.
40. Lee, J., Bayarsaikhan, D., Bayarsaikhan, G., Kim, J.S., Schwarzbach, E., and Lee, B. (2020). Recent advances in genome editing of stem cells for drug discovery and therapeutic application. *Pharmacol. Ther.* 209, 107501.
41. You, L., Tong, R., Li, M., Liu, Y., Xue, J., and Lu, Y. (2019). Advancements and obstacles of CRISPR-Cas9 technology in translational research. *Mol. Ther. Methods Clin. Dev.* 13, 359–370.
42. Coller, B.S. (2019). Ethics of human genome editing. *Annu. Rev. Med.* 70, 289–305.
43. Lea, R.A., and Niakan, K.K. (2019). Human germline genome editing. *Nat. Cell Biol.* 21, 1479–1489.
44. Ormond, K.E., Mortlock, D.P., Scholes, D.T., Bombard, Y., Brody, L.C., Faucett, W.A., Garrison, N.A., Hercher, L., Isasi, R., Middleton, A., et al. (2017). Human germline genome editing. *Am. J. Hum. Genet.* 101, 167–176.
45. Sukari, A., Abdallah, N., and Nagasaka, M. (2019). Unleash the power of the mighty T cells-basis of adoptive cellular therapy. *Crit. Rev. Oncol. Hematol.* 136, 1–12.
46. Yee, C., Lizee, G., and Schueneman, A.J. (2015). Endogenous T-cell therapy: clinical experience. *Cancer J.* 21, 492–500.
47. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of cancer: the next generation. *Cell* 144, 646–674.
48. Pardoll, D.M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12, 252–264.

49. Hargadon, K.M., Johnson, C.E., and Williams, C.J. (2018). Immune checkpoint blockade therapy for cancer: an overview of FDA-approved immune checkpoint inhibitors. *Int. Immunopharmacol.* 62, 29–39.
50. Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., et al. (2012). Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* 366, 2443–2454.
51. Brahmer, J.R., Tykodi, S.S., Chow, L.Q., Hwu, W.J., Topalian, S.L., Hwu, P., Drake, C.G., Camacho, L.H., Kauh, J., Odunsi, K., et al. (2012). Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med.* 366, 2455–2465.
52. Hamid, O., Robert, C., Daud, A., Hodi, F.S., Hwu, W.J., Kefford, R., Wolchok, J.D., Hersey, P., Joseph, R.W., Weber, J.S., et al. (2013). Safety and tumor responses with lambrolizumab (anti-PD-1) in melanoma. *N. Engl. J. Med.* 369, 134–144.
53. Michot, J.M., Bigenwald, C., Champiat, S., Collins, M., Carbone, F., Postel-Vinay, S., Berdelou, A., Varga, A., Bahleda, R., Hollebecq, A., et al. (2016). Immune-related adverse events with immune checkpoint blockade: a comprehensive review. *Eur. J. Cancer* 54, 139–148.
54. Ho, W.Y., Nguyen, H.N., Wolff, M., Kuball, J., and Greenberg, P.D. (2006). In vitro methods for generating CD8⁺ T-cell clones for immunotherapy from the naive repertoire. *J. Immunol. Methods* 310, 40–52.
55. Beane, J.D., Lee, G., Zheng, Z., Mendel, M., Abate-Daga, D., Bharathan, M., Black, M., Gandhi, N., Yu, Z., Chandran, S., et al. (2015). Clinical scale zinc finger nuclease-mediated gene editing of PD-1 in tumor infiltrating lymphocytes for the treatment of metastatic melanoma. *Mol. Ther.* 23, 1380–1390.
56. Su, S., Hu, B., Shao, J., Shen, B., Du, J., Du, Y., Zhou, J., Yu, L., Zhang, L., Chen, F., et al. (2016). CRISPR-Cas9 mediated efficient PD-1 disruption on human primary T cells from cancer patients. *Sci. Rep.* 6, 20070.
57. Marotte, L., Simon, S., Vignard, V., Dupre, E., Gantier, M., Cruard, J., Alberge, J.B., Hussong, M., Deleine, C., Heslan, J.M., et al. (2020). Increased antitumor efficacy of PD-1-deficient melanoma-specific human lymphocytes. *J. Immunother. Cancer* 8, e000311.
58. Menger, L., Sledzinska, A., Bergerhoff, K., Vargas, F.A., Smith, J., Poirot, L., Pule, M., Herrero, J., Peggs, K.S., and Quezada, S.A. (2016). TALEN-mediated inactivation of PD-1 in tumor-reactive lymphocytes promotes intratumoral T-cell persistence and rejection of established tumors. *Cancer Res.* 76, 2087–2093.
59. Zhao, Z., Shi, L., Zhang, W., Han, J., Zhang, S., Fu, Z., and Cai, J. (2017). CRISPR knock-out of programmed cell death protein 1 enhances anti-tumor activity of cytotoxic T lymphocytes. *Oncotarget* 9, 5208–5215.
60. Lu, S., Yang, N., He, J., Gong, W., Lai, Z., Xie, L., Tao, L., Xu, C., Wang, H., Zhang, G., et al. (2019). Generation of cancer-specific cytotoxic PD-1⁻ T cells using liposome-encapsulated CRISPR/Cas system with dendritic/tumor fusion cells. *J. Biomed. Nanotechnol.* 15, 593–601.
61. Wu, S.; Hangzhou Cancer Hospital, Ltd.; Anhui Kedgene Biotechnology Co., Ltd. (2017). PD-1 knockout engineered T cells for advanced esophageal cancer. <https://clinicaltrials.gov/ct2/show/NCT03081715>.
62. Lu, Y.; Sichuan University; Chengdu MedGenCell Co., Ltd. (2016). PD-1 knockout engineered T cells for metastatic non-small cell lung cancer. <https://clinicaltrials.gov/ct2/show/NCT02793856>.
63. Chen, S.; Guangzhou Anjie Biomedical Technology Co., Ltd.; University of Technology, Sydney (2018). Therapeutic vaccine plus PD-1 knockout in prostate cancer treatment. <https://clinicaltrials.gov/ct2/show/NCT03525652>.
64. Yang, Y.; The Affiliated Nanjing Drum Tower Hospital of Nanjing University Medical School (2017). PD-1 knockout EBV-CTLs for advanced stage Epstein-Barr virus (EBV) associated malignancies. <https://clinicaltrials.gov/ct2/show/NCT03044743>.
65. Lu, Y., Xue, J., Deng, T., Zhou, X., Yu, K., Deng, L., Huang, M., Yi, X., Liang, M., Wang, Y., et al. (2020). Safety and feasibility of CRISPR-edited T cells in patients with refractory non-small-cell lung cancer. *Nat. Med.* 26, 732–740.
66. Rothlin, C.V., and Ghosh, S. (2020). Lifting the innate immune barriers to antitumor immunity. *J. Immunother. Cancer* 8, e000695.
67. Chiosone, L., Dumas, P.Y., Vienne, M., and Vivier, E. (2018). Natural killer cells and other innate lymphoid cells in cancer. *Nat. Rev. Immunol.* 18, 671–688.
68. Rautela, J., Surgenor, E., and Huntington, N.D. (2018). Efficient genome editing of human natural killer cells by CRISPR RNP. *bioRxiv*. <https://doi.org/10.1101/406934>.
69. Pomeroy, E.J., Hunzeker, J.T., Kluesner, M.G., Lahr, W.S., Smeester, B.A., Crosby, M.R., Lonetree, C.L., Yamamoto, K., Bendzick, L., Miller, J.S., et al. (2020). A genetically engineered primary human natural killer cell platform for cancer immunotherapy. *Mol. Ther.* 28, 52–63.
70. Ruella, M., and Kalos, M. (2014). Adoptive immunotherapy for cancer. *Immunol. Rev.* 257, 14–38.
71. Liu, X., and Zhao, Y. (2018). CRISPR/Cas9 genome editing: fueling the revolution in cancer immunotherapy. *Curr. Res. Transl. Med.* 66, 39–42.
72. Kershaw, M.H., Westwood, J.A., and Darcy, P.K. (2013). Gene-engineered T cells for cancer therapy. *Nat. Rev. Cancer* 13, 525–541.
73. Cartellieri, M., Bachmann, M., Feldmann, A., Bippes, C., Stamova, S., Wehner, R., Temme, A., and Schmitz, M. (2010). Chimeric antigen receptor-engineered T cells for immunotherapy of cancer. *J. Biomed. Biotechnol.* 2010, 956304.
74. Ouchi, Y., Patil, A., Tamura, Y., Nishimasu, H., Negishi, A., Paul, S.K., Takemura, N., Satoh, T., Kimura, Y., Kuwachi, M., et al. (2018). Generation of tumor antigen-specific murine CD8⁺ T cells with enhanced anti-tumor activity via highly efficient CRISPR/Cas9 genome editing. *Int. Immunol.* 30, 141–154.
75. Guo, X., Jiang, H., Shi, B., Zhou, M., Zhang, H., Shi, Z., Du, G., Luo, H., Wu, X., Wang, Y., et al. (2018). Disruption of PD-1 enhanced the anti-tumor activity of chimeric antigen receptor T cells against hepatocellular carcinoma. *Front. Pharmacol.* 9, 1118.
76. Hu, B., Zou, Y., Zhang, L., Tang, J., Niedermann, G., Firat, E., Huang, X., and Zhu, X. (2019). Nucleofection with plasmid DNA for CRISPR/Cas9-mediated inactivation of programmed cell death protein 1 in CD133-specific CAR T cells. *Hum. Gene Ther.* 30, 446–458.
77. Hu, W., Zi, Z., Jin, Y., Li, G., Shao, K., Cai, Q., Ma, X., and Wei, F. (2019). CRISPR/Cas9-mediated PD-1 disruption enhances human mesothelin-targeted CAR T cell effector functions. *Cancer Immunol. Immunother.* 68, 365–377.
78. Rupp, L.J., Schumann, K., Roybal, K.T., Gate, R.E., Ye, C.J., Lim, W.A., and Marson, A. (2017). CRISPR/Cas9-mediated PD-1 disruption enhances anti-tumor efficacy of human chimeric antigen receptor T cells. *Sci. Rep.* 7, 737.
79. Zhu, H., You, Y., Shen, Z., and Shi, L. (2020). EGFRvIII-CAR-T cells with PD-1 knockout have improved anti-glioma activity. *Pathol. Oncol. Res.*, Published online January 27, 2020. <https://doi.org/10.1007/s12253-019-00759-1>.
80. Maude, S.L., Hucks, G.E., Seif, A.E., Talekar, M.K., Teachey, D.T., Baniewicz, D., Callahan, C., Gonzalez, V., Nazimuddin, F., Gupta, M., et al. (2017). The effect of pembrolizumab in combination with CD19-targeted chimeric antigen receptor (CAR) T cells in relapsed acute lymphoblastic leukemia (ALL). *J. Clin. Oncol.* 35 (Suppl.), 103.
81. Chong, E.A., Melenhorst, J.J., Lacey, S.F., Ambrose, D.E., Gonzalez, V., Levine, B.L., June, C.H., and Schuster, S.J. (2017). PD-1 blockade modulates chimeric antigen receptor (CAR)-modified T cells: refueling the CAR. *Blood* 129, 1039–1041.
82. Shang, X.; Third Military Medical University (2017). CD19 CAR and PD-1 knockout engineered T cells for CD19 positive malignant B-cell derived leukemia and lymphoma. <https://clinicaltrials.gov/ct2/show/NCT03298828>.
83. Weidong, H.; Chinese PLA General Hospital (2018). Study of PD-1 gene-knockout out mesothelin-directed CAR-T cells with the conditioning of PC in mesothelin positive multiple solid tumors. <https://clinicaltrials.gov/ct2/show/NCT03747965>.
84. Chen, S.; Guangzhou Anjie Biomedical Technology Co., Ltd. (2018). CAR T and PD-1 knockout engineered T cells for esophageal cancer. <https://clinicaltrials.gov/ct2/show/NCT03706326>.
85. Chen, S.; Guangzhou Anjie Biomedical Technology Co., Ltd.; University of Technology, Sydney (2018). Anti-MUC1 CAR T cells and PD-1 knockout engineered T cells for NSCLC. <https://clinicaltrials.gov/ct2/show/NCT03525782>.
86. Guangxun, G.; Xi'an Yufan Biotechnology Co., Ltd. (2019). CRISPR (HPK1) edited CD19-specific CAR-T cells (XYF19 CAR-T cells) for CD19⁺ leukemia or lymphoma. <https://www.clinicaltrials.gov/ct2/show/NCT04037566>.

Molecular Therapy Methods & Clinical Development

87. Shui, J.W., Boomer, J.S., Han, J., Xu, J., Dement, G.A., Zhou, G., and Tan, T.H. (2007). Hematopoietic progenitor kinase 1 negatively regulates T cell receptor signaling and T cell-mediated immune responses. *Nat. Immunol.* 8, 84–91.
88. Alzabin, S., Pyarajan, S., Yee, H., Kiefer, F., Suzuki, A., Burakoff, S., and Sawasdikosol, S. (2010). Hematopoietic progenitor kinase 1 is a critical component of prostaglandin E2-mediated suppression of the anti-tumor immune response. *Cancer Immunol. Immunother.* 59, 419–429.
89. Liu, J., Curtin, J., You, D., Hillerman, S., Li-Wang, B., Eraslan, R., Xie, J., Swanson, J., Ho, C.P., Oppenheimer, S., et al. (2019). Critical role of kinase activity of hematopoietic progenitor kinase 1 in anti-tumor immune surveillance. *PLoS ONE* 14, e0212670.
90. Hernandez, S., Qing, J., Thibodeau, R.H., Du, X., Park, S., Lee, H.M., Xu, M., Oh, S., Navarro, A., Roose-Girma, M., et al. (2018). The kinase activity of hematopoietic progenitor kinase 1 is essential for the regulation of T cell function. *Cell Rep.* 25, 80–94.
91. Zhang, W., Shi, L., Zhao, Z., Du, P., Ye, X., Li, D., Cai, Z., Han, J., and Cai, J. (2019). Disruption of CTLA-4 expression on peripheral blood CD8 + T cell enhances anti-tumor efficacy in bladder cancer. *Cancer Chemother. Pharmacol.* 83, 911–920.
92. Shi, L., Meng, T., Zhao, Z., Han, J., Zhang, W., Gao, F., and Cai, J. (2017). CRISPR knock out CTLA-4 enhances the anti-tumor activity of cytotoxic T lymphocytes. *Gene* 636, 36–41.
93. Jung, I.Y., Kim, Y.Y., Yu, H.S., Lee, M., Kim, S., and Lee, J. (2018). CRISPR/Cas9-mediated knockout of DGK improves antitumor activities of human T cells. *Cancer Res.* 78, 4692–4703.
94. Zhong, X.-P., Hainey, E.A., Olenchok, B.A., Jordan, M.S., Maltzman, J.S., Nichols, K.E., Shen, H., and Koretzky, G.A. (2003). Enhanced T cell responses due to diacylglycerol kinase ζ deficiency. *Nat. Immunol.* 4, 882–890.
95. Zhang, Y., Zhang, X., Cheng, C., Mu, W., Liu, X., Li, N., Wei, X., Liu, X., Xia, C., and Wang, H. (2017). CRISPR-Cas9 mediated LAG-3 disruption in CAR-T cells. *Front. Med.* 11, 554–562.
96. Singh, N., Shi, J., June, C.H., and Ruella, M. (2017). Genome-editing technologies in adoptive T cell immunotherapy for cancer. *Curr. Hematol. Malig. Rep.* 12, 522–529.
97. Torikai, H., and Cooper, L.J. (2016). Translational implications for off-the-shelf immune cells expressing chimeric antigen receptors. *mol. ther.* 24, 1178–1186.
98. Yang, Y., Jacoby, E., and Fry, T.J. (2015). Challenges and opportunities of allogeneic donor-derived CAR T cells. *Curr. Opin. Hematol.* 22, 509–515.
99. Torikai, H., Reik, A., Liu, P.Q., Zhou, Y., Zhang, L., Maiti, S., Huls, H., Miller, J.C., Kebriaei, P., Rabinovich, B., et al. (2012). A foundation for universal T-cell based immunotherapy: T cells engineered to express a CD19-specific chimeric-antigen-receptor and eliminate expression of endogenous TCR. *Blood* 119, 5697–5705.
100. Poirat, L., Philip, B., Schiffer-Mannioui, C., Le Clerc, D., Chion-Sotinel, I., Derniame, S., Potrel, P., Bas, C., Lemaire, L., Galetto, R., et al. (2015). Multiplex genome-edited T-cell manufacturing platform for “off-the-shelf” adoptive T-cell immunotherapies. *Cancer Res.* 75, 3853–3864.
101. Georgiadis, C., Preece, R., Nickolay, L., Etuk, A., Petrova, A., Ladon, D., Danyi, A., Humphries-Kirilov, N., Ajetunmbi, A., Kim, D., et al. (2018). Long terminal repeat CRISPR-CAR-coupled “universal” T cells mediate potent anti-leukemic effects. *Mol. Ther.* 26, 1215–1227.
102. Qasim, W., Zhan, H., Samarasinghe, S., Adams, S., Amrolia, P., Stafford, S., Butler, K., Rivat, C., Wright, G., Somana, K., et al. (2017). Molecular remission of infant B-ALL after infusion of universal TALEN gene-edited CAR T cells. *Sci. Transl. Med.* 9, eaaj2013.
103. Servier (Institut de Recherches Internationales Servier); ADIR, a Servier Group company (2016). Study of UCART19 in pediatric patients with relapsed/refractory b acute lymphoblastic leukemia. <https://clinicaltrials.gov/ct2/show/NCT02808442>.
104. Servier (Institut de Recherches Internationales Servier); ADIR, a Servier Group company (2016). Dose escalation study of UCART19 in adult patients with relapsed refractory B-cell acute lymphoblastic leukaemia. <https://clinicaltrials.gov/ct2/show/NCT02746952>.
105. Allogene Therapeutics (2019). Safety and efficacy of ALLO-501 Anti-CD19 allogeneic CAR T cells in adults with relapsed/refractory large B cell or follicular lymphoma (ALPHA). <https://clinicaltrials.gov/ct2/show/NCT03939026>.
106. Allogene Therapeutics (2019). Safety and efficacy of ALLO-715 BCMA allogeneic CAR T cells in adults with relapsed or refractory multiple myeloma (UNIVERSAL). <https://clinicaltrials.gov/ct2/show/NCT04093596>.
107. Servier (Institut de Recherches Internationales Servier); ADIR, a Servier Group company (2016). A study to evaluate the long-term safety of patients with advanced lymphoid malignancies who have been previously administered with UCART19/ALLO-501. <https://clinicaltrials.gov/ct2/show/NCT02735083>.
108. Ren, J., Zhang, X., Liu, X., Fang, C., Jiang, S., June, C.H., and Zhao, Y. (2017). A versatile system for rapid multiplex genome-edited CAR T cell generation. *Oncotarget* 8, 17002–17011.
109. Ren, J., Liu, X., Fang, C., Jiang, S., June, C.H., and Zhao, Y. (2017). Multiplex genome editing to generate universal CAR T cells resistant to PD1 inhibition. *Clin. Cancer Res.* 23, 2255–2266.
110. Liu, X., Zhang, Y., Cheng, C., Cheng, A.W., Zhang, X., Li, N., Xia, C., Wei, X., Liu, X., and Wang, H. (2017). CRISPR-Cas9-mediated multiplex gene editing in CAR-T cells. *Cell Res.* 27, 154–157.
111. Choi, B.D., Yu, X., Castano, A.P., Darr, H., Henderson, D.B., Bouffard, A.A., Larson, R.C., Scarfo, L., Bailey, S.R., Gerhard, G.M., et al. (2019). CRISPR-Cas9 disruption of PD-1 enhances activity of universal EGFRvIII CAR T cells in a preclinical model of human glioblastoma. *J. Immunother.* 42, 304.
112. Weidong, H.; Chinese PLA General Hospital (2017). A study evaluating UCART019 in patients with relapsed or refractory CD19+ leukemia and lymphoma. <https://clinicaltrials.gov/ct2/show/NCT03166878>.
113. Eyquem, J., Mansilla-Soto, J., Giavridis, T., van der Stegen, S.J., Hamieh, M., Cunanan, K.M., Odak, A., Gonen, M., and Sadelain, M. (2017). Targeting a CAR to the TRAC locus with CRISPR/Cas9 enhances tumour rejection. *Nature* 543, 113–117.
114. Roth, T.L., Puig-Saus, C., Yu, R., Shifrut, E., Carnevale, J., Li, P.J., Hiatt, J., Saco, J., Krystofinski, P., Li, H., et al. (2018). Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature* 559, 405–409.
115. CRISPR Therapeutics AG (2019). A safety and efficacy study evaluating CTX110 in subjects with relapsed or refractory B-cell malignancies. <https://clinicaltrials.gov/ct2/show/NCT04035434>.
116. CRISPR Therapeutics AG (2020). A safety and efficacy study evaluating CTX120 in subjects with relapsed or refractory multiple myeloma. <https://clinicaltrials.gov/ct2/show/NCT04244656>.
117. Xu, K.L., and Zheng, J.N.; Nanjing Bioheng Biotech Co., Ltd. (2019). CTA101 UCAR-T cell injection for treatment of relapsed or refractory CD19+ B-cell acute lymphoblastic leukemia. <https://clinicaltrials.gov/ct2/show/NCT04154709>.
118. Collectis S.A. (2019). Phase I study of UCART22 in patients with relapsed or refractory CD22+ B-cell acute lymphoblastic leukemia (BALLI-01). <https://clinicaltrials.gov/ct2/show/NCT04150497>.
119. The First Affiliated Hospital with Nanjing Medical University; Nanjing Bioheng Biotech Co., Ltd. (2019). CTA101 in the treatment of relapsed or refractory diffuse large B-cell lymphoma. <https://clinicaltrials.gov/ct2/show/NCT04026100>.
120. Weidong, H.; Chinese PLA General Hospital (2018). A feasibility and safety study of universal dual specificity CD19 and CD20 or CD22 CAR-T cell immunotherapy for relapsed or refractory leukemia and lymphoma. <https://clinicaltrials.gov/ct2/show/NCT03398967>.
121. Collectis S.A. (2017). Study evaluating safety and efficacy of UCART123 in patients with relapsed/ refractory acute myeloid leukemia. <https://clinicaltrials.gov/ct2/show/NCT03190278>.
122. Collectis S.A. (2019). Study evaluating safety and efficacy of UCART targeting CS1 in patients with relapsed/refractory multiple myeloma (MELANI-01). <https://clinicaltrials.gov/ct2/show/NCT04142619>.
123. Zhang, X.; Gracell Biotechnologies (Shanghai) Co., Ltd.; 920th Hospital of Joint Logistics Support Force; The Second Affiliated Hospital of Chongqing Medical University; The Affiliated Hospital of Guizhou Medical University; Central South University; The First Affiliated Hospital of Kunming Medical College; The General Hospital of Western Theater Command; Second Affiliated Hospital of Xi'an Jiaotong University; Nanfang Hospital of Southern Medical University; Fujian Medical University Union Hospital; The First Affiliated Hospital of Anhui

- Medical University; Tang-Du Hospital (2020). Anti-CD19 U-CAR-T cell therapy for B cell hematologic malignancies. <https://clinicaltrials.gov/ct2/show/NCT04264039>.
124. Xiang, X.; Gracell Biotechnologies (Shanghai) Co., Ltd.; 920th Hospital of Joint Logistics Support Force; The Second Affiliated Hospital of Chongqing Medical University; The Affiliated Hospital of Guizhou Medical University; Central South University; The First Affiliated Hospital of Kunming Medical College; The General Hospital of Western Theater Command; Second Affiliated Hospital of Xi'an Jiaotong University; Nanfang Hospital of Southern Medical University; Fujian Medical University Union Hospital; The First Affiliated Hospital of Anhui Medical University; Tang-Du Hospital (2020). Anti-CD7 U-CAR-T cell therapy for T/NK cell hematologic malignancies. <https://clinicaltrials.gov/ct2/show/NCT04264078>.
 125. Shanghai Bioray Laboratory Inc.; Shanghai Tongji Hospital, Tongji University School of Medicine; Second Xiangya Hospital of Central South University (2018). Efficacy and safety evaluation of BCMA-UCART. <https://clinicaltrials.gov/ct2/show/NCT03752541>.
 126. Shanghai Bioray Laboratory Inc.; The First Affiliated Hospital of Zhengzhou University; Second Xiangya Hospital of Central South University; Shanghai 10th People's Hospital (2017). Safety and efficacy evaluation of CD19-UCART. <https://clinicaltrials.gov/ct2/show/NCT03229876>.
 127. Shanghai Longyao Biotechnology Inc. Ltd. (2019). The clinical study of CD19 UCART-T cells in patients with B-cell acute lymphoblastic leukemia (B-ALL). <https://clinicaltrials.gov/ct2/show/NCT04166838>.
 128. Ma, H., Padmanabhan Iyer, S., Parmar, S., and Gong, Y. (2019). Adoptive cell therapy for acute myeloid leukemia. *Leuk. Lymphoma* 60, 1370–1380.
 129. Kenderian, S.S., Ruella, M., Shestova, O., Klichinsky, M., Aikawa, V., Morrisette, J.J., Scholler, J., Song, D., Porter, D.L., Carroll, M., et al. (2015). CD33-specific chimeric antigen receptor T cells exhibit potent preclinical activity against human acute myeloid leukemia. *Leukemia* 29, 1637–1647.
 130. Gill, S., Tasian, S.K., Ruella, M., Shestova, O., Li, Y., Porter, D.L., Carroll, M., Danet-Desnoves, G., Scholler, J., Grupp, S.A., et al. (2014). Preclinical targeting of human acute myeloid leukemia and myeloablation using chimeric antigen receptor-modified T cells. *Blood* 123, 2343–2354.
 131. Kim, M.Y., Yu, K.R., Kenderian, S.S., Ruella, M., Chen, S., Shin, T.H., Aljanahi, A.A., Schreeder, D., Klichinsky, M., Shestova, O., et al. (2018). Genetic inactivation of CD33 in hematopoietic stem cells to enable CAR T cell immunotherapy for acute myeloid leukemia. *Cell* 173, 1439–1453.e19.
 132. Knipping, F., Osborn, M.J., Petri, K., Tolar, J., Glimm, H., von Kalle, C., Schmidt, M., and Gabriel, R. (2017). Genome-wide specificity of highly efficient TALENs and CRISPR/Cas9 for T cell receptor modification. *Mol. Ther. Methods Clin. Dev.* 4, 213–224.
 133. University of Pennsylvania (2018). NY-ESO-1-redirected CRISPR (TCRendo and PD1) edited T cells (NYCE T cells). <https://clinicaltrials.gov/ct2/show/NCT03399448>.
 134. Weidong, H.; Chinese PLA General Hospital (2018). Study of CRISPR-Cas9 mediated PD-1 and TCR gene-knockout on mesothelin-directed CAR-T cells in patients with mesothelin positive multiple solid tumors. <https://clinicaltrials.gov/ct2/show/NCT03545815>.
 135. Stadtmauer, E.A., Fraietta, J.A., Davis, M.M., Cohen, A.D., Weber, K.L., Lancaster, E., Mangan, P.A., Kulikovskaya, I., Gupta, M., Chen, F., et al. (2020). CRISPR-engineered T cells in patients with refractory cancer. *Science* 367, eaba7365.
 136. Rapoport, A.P., Stadtmauer, E.A., Binder-Scholl, G.K., Golubeva, O., Vogl, D.T., Lacey, S.F., Badros, A.Z., Garfall, A., Weiss, B., Finkelshtein, J., et al. (2015). NY-ESO-1-specific TCR-engineered T cells mediate sustained antigen-specific antitumor effects in myeloma. *Nat. Med.* 21, 914–921.
 137. Robbins, P.F., Kassim, S.H., Tran, T.L., Crystal, J.S., Morgan, R.A., Feldman, S.A., Yang, J.C., Dudley, M.E., Wunderlich, J.R., Sherry, R.M., et al. (2015). A pilot trial using lymphocytes genetically engineered with an NY-ESO-1-reactive T-cell receptor: long-term follow-up and correlates with response. *Clin. Cancer Res.* 21, 1019–1027.
 138. D'Angelo, S.P., Melchiori, L., Merchant, M.S., Bernstein, D., Glod, J., Kaplan, R., Grupp, S., Tap, W.D., Chagin, K., Binder, G.K., et al. (2018). Antitumor activity associated with prolonged persistence of adoptively transferred NY-ESO-1 ^{CD259}T cells in synovial sarcoma. *Cancer Discov.* 8, 944–957.
 139. Alcantara, M., Tesio, M., June, C.H., and Houot, R. (2018). CAR T-cells for T-cell malignancies: challenges in distinguishing between therapeutic, normal, and neoplastic T-cells. *Leukemia* 32, 2307–2315.
 140. Zhou, S., Zhu, X., Shen, N., Li, Q., Wang, N., You, Y., Zhong, Z., Cheng, F., Zou, P., and Zhu, X. (2019). T cells expressing CD26-specific chimeric antigen receptors exhibit extensive self-antigen-driven fratricide. *Immunopharmacol. Immunotoxicol.* 41, 490–496.
 141. Leisegang, M., Wilde, S., Spranger, S., Milosevic, S., Frankenberger, B., Uckert, W., and Schendel, D.J. (2010). MHC-restricted fratricide of human lymphocytes expressing survivin-specific transgenic T cell receptors. *J. Clin. Invest.* 120, 3869–3877.
 142. Cooper, M.L., Choi, J., Staser, K., Ritchey, J.K., Devenport, J.M., Eckardt, K., Rettig, M.P., Wang, B., Eissenberg, L.G., Ghebadi, A., et al. (2018). An "off-the-shelf" fratricide-resistant CAR-T for the treatment of T cell hematologic malignancies. *Leukemia* 32, 1970–1983.
 143. Gomes-Silva, D., Srinivasan, M., Sharma, S., Lee, C.M., Wagner, D.L., Davis, T.H., Rouce, R.H., Bao, G., Brenner, M.K., and Mamontkin, M. (2017). CD7-edited T cells expressing a CD7-specific CAR for the therapy of T-cell malignancies. *Blood* 130, 285–296.
 144. Rouce, R.; The Methodist Hospital System; Center for Cell and Gene Therapy, Baylor College of Medicine (2018). Cell therapy for high risk T-cell malignancies using CD7-specific CAR expressed on autologous T cells. <https://clinicaltrials.gov/ct2/show/NCT03690011>.
 145. Jemal, A., Bray, F., Center, M.M., Ferlay, J., Ward, E., and Forman, D. (2011). Global cancer statistics. *CA Cancer J. Clin.* 61, 69–90.
 146. Bruni, L., Diaz, M., Barrionuevo-Rosas, L., Herrero, R., Bray, F., Bosch, F.X., de Sanjosé, S., and Castellsagué, X. (2016). Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis. *Lancet Glob. Health* 4, e453–e463.
 147. Crosbie, E.J., Einstein, M.H., Franceschi, S., and Kitchener, H.C. (2013). Human papillomavirus and cervical cancer. *Lancet* 382, 889–899.
 148. Howlader, N., Noone, A.M., Krapcho, M., Miller, D., Bishop, K., Kosary, C.L., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D.R., Chen, H.S., Feuer, E.J., and Cronin, K.A., eds. (2018). SEER Cancer Statistics Review, 1975–2014 (National Cancer Institute), https://seer.cancer.gov/csr/1975_2014/.
 149. zur Hausen, H. (2002). Papillomaviruses and cancer: from basic studies to clinical application. *Nat. Rev. Cancer* 2, 342–350.
 150. Sima, N., Wang, W., Kong, D., Deng, D., Xu, Q., Zhou, J., Xu, G., Meng, L., Lu, Y., Wang, S., and Ma, D. (2008). RNA interference against HPV16 E7 oncogene leads to viral E6 and E7 suppression in cervical cancer cells and apoptosis via upregulation of Rb and p53. *Apoptosis* 13, 273–281.
 151. Wang, W., Sima, N., Kong, D., Luo, A., Gao, Q., Liao, S., Li, W., Han, L., Wang, J., Wang, S., et al. (2010). Selective targeting of HPV-16 E6/E7 in cervical cancer cells with a potent oncolytic adenovirus and its enhanced effect with radiotherapy in vitro and vivo. *Cancer Lett.* 291, 67–75.
 152. Almeida, A.M., Queiroz, J.A., Sousa, F., and Sousa, Á. (2019). Cervical cancer and HPV infection: ongoing therapeutic research to counteract the action of E6 and E7 oncoproteins. *Drug Discov. Today* 24, 2044–2057.
 153. Zheng, Z.M., Tang, S., and Tao, M. (2005). Development of resistance to RNAi in mammalian cells. *Ann. N Y Acad. Sci.* 1058, 105–118.
 154. Ding, W., Hu, Z., Zhu, D., Jiang, X., Yu, L., Wang, X., Zhang, C., Wang, L., Ji, T., Li, K., et al. (2014). Zinc finger nucleases targeting the human papillomavirus E7 oncogene induce E7 disruption and a transformed phenotype in HPV16/18-positive cervical cancer cells. *Clin. Cancer Res.* 20, 6495–6503.
 155. Hu, Z., Ding, W., Zhu, D., Yu, L., Jiang, X., Wang, X., Zhang, C., Wang, L., Ji, T., Liu, D., et al. (2015). TALEN-mediated targeting of HPV oncogenes ameliorates HPV-related cervical malignancy. *J. Clin. Invest.* 125, 425–436.
 156. Shankar, S., Prasad, D., Sanawar, R., Das, A.V., and Pillai, M.R. (2017). TALEN based HPV-E7 editing triggers necrotic cell death in cervical cancer cells. *Sci. Rep.* 7, 5500.

Molecular Therapy Methods & Clinical Development

157. Zhen, S., Hua, L., Takahashi, Y., Narita, S., Liu, Y.H., and Li, Y. (2014). In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9. *Biochem. Biophys. Res. Commun.* 450, 1422–1426.
158. Hu, Z., Yu, L., Zhu, D., Ding, W., Wang, X., Zhang, C., Wang, L., Jiang, X., Shen, H., He, D., et al. (2014). Disruption of HPV16-E7 by CRISPR/Cas system induces apoptosis and growth inhibition in HPV16 positive human cervical cancer cells. *BioMed Res. Int.* 2014, 612823.
159. Kennedy, E.M., Kornepati, A.V., Goldstein, M., Bogerd, H.P., Poling, B.C., Whisnant, A.W., Kastan, M.B., and Cullen, B.R. (2014). Inactivation of the human papillomavirus E6 or E7 gene in cervical carcinoma cells by using a bacterial CRISPR/Cas RNA-guided endonuclease. *J. Virol.* 88, 11965–11972.
160. Ma, D.; Huazhong University of Science and Technology (2017). Study of molecular-targeted therapy using zinc finger nuclease in cervical precancerous lesions. <https://clinicaltrials.gov/ct2/show/NCT02800369>.
161. Ma, D.; Huazhong University of Science and Technology (2017). Study of targeted therapy using transcription activator-like effector nucleases in cervical precancerous lesions. <https://clinicaltrials.gov/ct2/show/NCT03226470>.
162. Zheng, H.; First Affiliated Hospital, Sun Yat-Sen University; Jìngchū University of Technology (2017). A safety and efficacy study of TALEN and CRISPR/Cas9 in the treatment of HPV-related cervical intraepithelial neoplasia. <https://clinicaltrials.gov/ct2/show/NCT03057912>.
163. Zhen, S., Lu, J.J., Wang, L.J., Sun, X.M., Zhang, J.Q., Li, X., Luo, W.J., and Zhao, L. (2016). In vitro and in vivo synergistic therapeutic effect of cisplatin with human papillomavirus16 E6/E7 CRISPR/Cas9 on cervical cancer cell line. *Transl. Oncol.* 9, 498–504.
164. Münger, K. (2002). The role of human papillomaviruses in human cancers. *Front. Biosci.* 7, d641–d649.
165. Lin, S.R., Yang, H.C., Kuo, Y.T., Liu, C.J., Yang, T.Y., Sung, K.C., Lin, Y.Y., Wang, H.Y., Wang, C.C., Shen, Y.C., et al. (2014). The CRISPR/Cas9 system facilitates clearance of the intrahepatic HBV templates in vivo. *Mol. Ther. Nucleic Acids* 3, e186.
166. Seeger, C., and Sohn, J.A. (2014). Targeting hepatitis B virus with CRISPR/Cas9. *Mol. Ther. Nucleic Acids* 3, e216.
167. Kennedy, E.M., and Cullen, B.R. (2015). Bacterial CRISPR/Cas DNA endonucleases: a revolutionary technology that could dramatically impact viral research and treatment. *Virology* 479–480, 213–220.
168. Zhen, S., Hua, L., Liu, Y.H., Gao, L.C., Fu, J., Wan, D.Y., Dong, L.H., Song, H.F., and Gao, X. (2015). Harnessing the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated Cas9 system to disrupt the hepatitis B virus. *Gene Ther.* 22, 404–412.
169. Li, H., Sheng, C., Wang, S., Yang, L., Liang, Y., Huang, Y., Liu, H., Li, P., Yang, C., Yang, X., et al. (2017). Removal of integrated hepatitis B virus DNA using CRISPR-Cas9. *Front. Cell. Infect. Microbiol.* 7, 91.
170. Scott, T., Moyo, B., Nicholson, S., Maepa, M.B., Watashi, K., Ely, A., Weinberg, M.S., and Arbutnot, P. (2017). ssAAVs containing cassettes encoding SaCas9 and guides targeting hepatitis B virus inactivate replication of the virus in cultured cells. *Sci. Rep.* 7, 7401.
171. Liu, Y., Zhao, M., Gong, M., Xu, Y., Xie, C., Deng, H., Li, X., Wu, H., and Wang, Z. (2018). Inhibition of hepatitis B virus replication via HBV DNA cleavage by Cas9 from *Staphylococcus aureus*. *Antiviral Res.* 152, 58–67.
172. Jiang, C., Mei, M., Li, B., Zhu, X., Zu, W., Tian, Y., Wang, Q., Guo, Y., Dong, Y., and Tan, X. (2017). A non-viral CRISPR/Cas9 delivery system for therapeutically targeting HBV DNA and *pcsk9* in vivo. *Cell Res.* 27, 440–443.
173. Wang, J., and Quake, S.R. (2014). RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesviridae infection. *Proc. Natl. Acad. Sci. USA* 111, 13157–13162.
174. Yuen, K.S., Chan, C.P., Wong, N.M., Ho, C.H., Ho, T.H., Lei, T., Deng, W., Tsao, S.W., Chen, H., Kok, K.H., and Jin, D.Y. (2015). CRISPR/Cas9-mediated genome editing of Epstein-Barr virus in human cells. *J. Gen. Virol.* 96, 626–636.
175. van Diemen, F.R., Kruse, E.M., Hooykaas, M.J., Bruggeling, C.E., Schürch, A.C., van Ham, P.M., Imhof, S.M., Nijhuis, M., Wiertz, E.J., and Lebbink, R.J. (2016). CRISPR/Cas9-mediated genome editing of herpesviruses limits productive and latent infections. *PLoS Pathog.* 12, e1005701.
176. Yuen, K.S., Wang, Z.M., Wong, N.M., Zhang, Z.Q., Cheng, T.F., Lui, W.Y., Chan, C.P., and Jin, D.Y. (2018). Suppression of Epstein-Barr virus DNA load in latently infected nasopharyngeal carcinoma cells by CRISPR/Cas9. *Virus Res.* 244, 296–303.
177. de Buhr, H., and Lebbink, R.J. (2018). Harnessing CRISPR to combat human viral infections. *Curr. Opin. Immunol.* 54, 123–129.
178. Mbonye, U., and Karn, J. (2017). The molecular basis for human immunodeficiency virus latency. *Annu. Rev. Virol.* 4, 261–285.
179. Chun, T.W., Davey, R.T., Jr., Engel, D., Lane, H.C., and Fauci, A.S. (1999). Re-emergence of HIV after stopping therapy. *Nature* 401, 874–875.
180. Jacob, S.A., Jacob, D.G., and Jugulete, G. (2017). Improving the adherence to anti-retroviral therapy, a difficult but essential task for a successful HIV treatment-clinical points of view and practical considerations. *Front. Pharmacol.* 8, 831.
181. Allers, K., Hütter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E., and Schneider, T. (2011). Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood* 117, 2791–2799.
182. Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müsigg, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O., et al. (2009). Long-term control of HIV by CCR5 delta32/delta32 stem-cell transplantation. *N. Engl. J. Med.* 360, 692–698.
183. Wu, L., Gerard, N.P., Wyatt, R., Choe, H., Parolin, C., Ruffing, N., Borsetti, A., Cardoso, A.A., Desjardins, E., Newman, W.V., et al. (1996). CD4-induced interaction of primary HIV-1 gp120 glycoproteins with the chemokine receptor CCR-5. *Nature* 384, 179–183.
184. Gupta, R.K., Abdul-Jawad, S., McCoy, L.E., Mok, H.P., Peppas, D., Salgado, M., Martinez-Picado, J., Nijhuis, M., Wensing, A.M.J., Lee, H., et al. (2019). HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* 568, 244–248.
185. Ioannidis, J.P., Rosenberg, P.S., Goedert, J.J., Ashton, L.J., Benfield, T.L., Buchbinder, S.P., Coutinho, R.A., Eugen-Olsen, J., Gallart, T., Katzenstein, T.L., et al.; International Meta-Analysis of HIV Host Genetics (2001). Effects of CCR5-Δ32, CCR2-64I, and SDF-1 3′A alleles on HIV-1 disease progression: an international meta-analysis of individual-patient data. *Ann. Intern. Med.* 135, 782–795.
186. Ioannidis, J.P., Contopoulos-Ioannidis, D.G., Rosenberg, P.S., Goedert, J.J., De Rossi, A., Espanol, T., Frenkel, L., Mayaux, M.J., Newell, M.L., Pahwa, S.G., et al.; HIV Host Genetics International Meta-Analysis Group (2003). Effects of CCR5-delta32 and CCR2-64I alleles on disease progression of perinatally HIV-1-infected children: an international meta-analysis. *AIDS* 17, 1631–1638.
187. Mulherin, S.A., O'Brien, T.R., Ioannidis, J.P., Goedert, J.J., Buchbinder, S.P., Coutinho, R.A., Jamieson, B.D., Meyer, L., Michael, N.L., Pantaleo, G., et al.; International Meta-Analysis of HIV Host Genetics (2003). Effects of CCR5-delta32 and CCR2-64I alleles on HIV-1 disease progression: the protection varies with duration of infection. *AIDS* 17, 377–387.
188. Perez, E.E., Wang, J., Miller, J.C., Jouvenot, Y., Kim, K.A., Liu, O., Wang, N., Lee, G., Bartsevich, V.V., Lee, Y.L., et al. (2008). Establishment of HIV-1 resistance in CD4⁺ T cells by genome editing using zinc-finger nucleases. *Nat. Biotechnol.* 26, 808–816.
189. University of Pennsylvania; Sangamo Therapeutics (2009). Autologous T-cells genetically modified at the CCR5 gene by zinc finger nucleases SB-728 for HIV (zinc-finger). <https://clinicaltrials.gov/ct2/show/NCT00842634>.
190. Sangamo Therapeutics (2010). Phase I dose escalation study of autologous t-cells genetically modified at the CCR5 gene by zinc finger nucleases in HIV-infected patients. <https://clinicaltrials.gov/ct2/show/NCT01044654>.
191. Sangamo Therapeutics (2014). Repeat doses of SB-728mR-T after cyclophosphamide conditioning in HIV-infected subjects on HAART. <https://clinicaltrials.gov/ct2/show/NCT02225665>.
192. University of Pennsylvania; National Institute of Allergy and Infectious Diseases (NIAID) (2015). A phase I study of T-cells genetically modified at the CCR5 gene by zinc finger nucleases SB-728mR in HIV-infected patients. <https://clinicaltrials.gov/ct2/show/NCT02388594>.
193. Sangamo Therapeutics (2012). Dose escalation study of cyclophosphamide in HIV-infected subjects on HAART receiving SB-728-T. <https://clinicaltrials.gov/ct2/show/NCT01543152>.

194. Smith, C.; Case Western Reserve University; University of California, San Francisco; University of Cincinnati (2018). CCR5-modified CD4⁺ T cells for HIV infection (TRAILBLAZER). <https://clinicaltrials.gov/ct2/show/NCT03666871>.
195. Sangamo Therapeutics (2014). Study of autologous T-cells genetically modified at the CCR5 gene by zinc finger nucleases in HIV-infected subjects. <https://clinicaltrials.gov/ct2/show/NCT01252641>.
196. Tebas, P., Stein, D., Tang, W.W., Frank, I., Wang, S.Q., Lee, G., Spratt, S.K., Suroskey, R.T., Giedlin, M.A., Nichol, G., et al. (2014). Gene editing of CCR5 in autologous CD4 T cells of persons infected with HIV. *N. Engl. J. Med.* 370, 901–910.
197. Sangamo Therapeutics (2019). Long-term follow-up of HIV subjects exposed to SB-728-T or SB-728mR-T. <https://clinicaltrials.gov/ct2/show/NCT04201782>.
198. Wang, C.X., and Cannon, P.M. (2016). Clinical applications of genome editing to HIV cure. *AIDS Patient Care STDS* 30, 539–544.
199. Wang, C.X., and Cannon, P.M. (2016). The clinical applications of genome editing in HIV. *Blood* 127, 2546–2552.
200. Li, L., Krymskaya, L., Wang, J., Henley, J., Rao, A., Cao, L.F., Tran, C.A., Torres-Coronado, M., Gardner, A., Gonzalez, N., et al. (2013). Genomic editing of the HIV-1 coreceptor CCR5 in adult hematopoietic stem and progenitor cells using zinc finger nucleases. *Mol. Ther.* 21, 1259–1269.
201. Wang, J., Exline, C.M., DeClercq, J.J., Llewellyn, G.N., Hayward, S.B., Li, P.W., Shivak, D.A., Suroskey, R.T., Gregory, P.D., Holmes, M.C., and Cannon, P.M. (2015). Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat. Biotechnol.* 33, 1256–1263.
202. Holt, N., Wang, J., Kim, K., Friedman, G., Wang, X., Taupin, V., Crooks, G.M., Kohn, D.B., Gregory, P.D., Holmes, M.C., and Cannon, P.M. (2010). Human hematopoietic stem/progenitor cells modified by zinc-finger nucleases targeted to CCR5 control HIV-1 in vivo. *Nat. Biotechnol.* 28, 839–847.
203. Krishnan, A.Y.; City of Hope Medical Center; Sangamo Therapeutics/California Institute for Regenerative Medicine (CIRM) (2015). Safety study of zinc finger nuclease CCR5-modified hematopoietic stem/progenitor cells in HIV-1 infected patients. <https://clinicaltrials.gov/ct2/show/NCT02500849>.
204. Hu, C.; Affiliated Hospital to Academy of Military Medical Sciences/Peking University; Capital Medical University (2017). Safety of transplantation of CRISPR CCR5 modified CD34⁺ cells in HIV-infected subjects with hematological malignances. <https://clinicaltrials.gov/ct2/show/NCT03164135>.
205. Maldini, C.R., Ellis, G.L., and Riley, J.L. (2018). CAR T cells for infection, autoimmunity and allotransplantation. *Nat. Rev. Immunol.* 18, 605–616.
206. Tebas, P.; University of Pennsylvania (2018). CD4 CAR⁺ ZFN-modified T cells in HIV therapy. <https://clinicaltrials.gov/ct2/show/NCT03617198>.
207. Oh, D.Y., Jessen, H., Kücherer, C., Neumann, K., Oh, N., Poggensee, G., Bartmeyer, B., Jessen, A., Pruss, A., Schumann, R.R., and Hamouda, O. (2008). CCR5Δ32 genotypes in a German HIV-1 seroconverter cohort and report of HIV-1 infection in a CCR5Δ32 homozygous individual. *PLoS ONE* 3, e2747.
208. Kordelas, L., Verheyen, J., Beelen, D.W., Horn, P.A., Heinold, A., Kaiser, R., Trenschel, R., Schadendorf, D., Dittmer, U., and Esser, S.; Essen HIV AlloSCT Group (2014). Shift of HIV tropism in stem-cell transplantation with CCR5 Delta32 mutation. *N. Engl. J. Med.* 371, 880–882.
209. Henrich, T.J., Hanhauser, E., Hu, Z., Stellbrink, H.J., Noah, C., Martin, J.N., Deeks, S.G., Kuritzkes, D.R., and Pereyra, F. (2015). Viremic control and viral coreceptor usage in two HIV-1-infected persons homozygous for CCR5 Δ32. *AIDS* 29, 867–876.
210. Murray, A.J., Kwon, K.J., Farber, D.L., and Siliciano, R.F. (2016). The latent reservoir for HIV-1: how immunologic memory and clonal expansion contribute to HIV-1 persistence. *J. Immunol.* 197, 407–417.
211. Ebina, H., Misawa, N., Kanemura, Y., and Koyanagi, Y. (2013). Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci. Rep.* 3, 2510.
212. Hu, W., Kaminski, R., Yang, F., Zhang, Y., Cosentino, L., Li, F., Luo, B., Alvarez-Carbonell, D., Garcia-Mesa, Y., Karn, J., et al. (2014). RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc. Natl. Acad. Sci. USA* 111, 11461–11466.
213. Zhu, W., Lei, R., Le Duff, Y., Li, J., Guo, F., Wainberg, M.A., and Liang, C. (2015). The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology* 12, 22.
214. Kaminski, R., Chen, Y., Fischer, T., Tedaldi, E., Napoli, A., Zhang, Y., Karn, J., Hu, W., and Khalili, K. (2016). Elimination of HIV-1 genomes from human T-lymphoid cells by CRISPR/Cas9 gene editing. *Sci. Rep.* 6, 22555.
215. Wang, G., Zhao, N., Berkhout, B., and Das, A.T. (2016). A combinatorial CRISPR-Cas9 attack on HIV-1 DNA extinguishes all infectious provirus in infected T cell cultures. *Cell Rep.* 17, 2819–2826.
216. Liao, H.K., Gu, Y., Diaz, A., Marlett, J., Takahashi, Y., Li, M., Suzuki, K., Xu, R., Hishida, T., Chang, C.J., et al. (2015). Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat. Commun.* 6, 6413.
217. Yin, C., Zhang, T., Li, F., Yang, F., Putatunda, R., Young, W.B., Khalili, K., Hu, W., and Zhang, Y. (2016). Functional screening of guide RNAs targeting the regulatory and structural HIV-1 viral genome for a cure of AIDS. *AIDS* 30, 1163–1174.
218. Lebbink, R.J., de Jong, D.C., Wolters, F., Kruse, E.M., van Ham, P.M., Wiertz, E.J., and Nijhuis, M. (2017). A combinatorial CRISPR/Cas9 gene-editing approach can halt HIV replication and prevent viral escape. *Sci. Rep.* 7, 41968.
219. Yin, L., Hu, S., Mei, S., Sun, H., Xu, F., Li, J., Zhu, W., Liu, X., Zhao, F., Zhang, D., et al. (2018). CRISPR/Cas9 inhibits multiple steps of HIV-1 infection. *Hum. Gene Ther.* 29, 1264–1276.
220. Xiao, Q., Guo, D., and Chen, S. (2019). Application of CRISPR/Cas9-based gene editing in HIV-1/AIDS therapy. *Front. Cell. Infect. Microbiol.* 9, 69.
221. Tsukamoto, T. (2019). Gene therapy approaches to functional cure and protection of hematopoietic potential in HIV infection. *Pharmaceutics* 11, E114.
222. Bobbin, M.L., Burnett, J.C., and Rossi, J.J. (2015). RNA interference approaches for treatment of HIV-1 infection. *Genome Med.* 7, 50.
223. Hartweger, H., McGuire, A.T., Horning, M., Taylor, J.J., Dosenovic, P., Yost, D., Gazumyan, A., Seaman, M.S., Stamatatos, L., Jankovic, M., and Nussenzweig, M.C. (2019). HIV-specific humoral immune responses by CRISPR/Cas9-edited B cells. *J. Exp. Med.* 216, 1301–1310.
224. Voss, J.E., Gonzalez-Martin, A., Andrabi, R., Fuller, R.P., Murrell, B., McCoy, L.E., Porter, K., Huang, D., Li, W., Sok, D., et al. (2019). Reprogramming the antigen specificity of B cells using genome-editing technologies. *eLife* 8, e42995.
225. Moffett, H.F., Harms, C.K., Fitzpatrick, K.S., Tooley, M.R., Boonyaratankornkit, J., and Taylor, J.J. (2019). B cells engineered to express pathogen-specific antibodies protect against infection. *Sci. Immunol.* 4, eaax0644.
226. Weatherall, D.J. (2001). Phenotype-genotype relationships in monogenic disease: lessons from the thalassaemias. *Nat. Rev. Genet.* 2, 245–255.
227. Bank, A. (2006). Regulation of human fetal hemoglobin: new players, new complexities. *Blood* 107, 435–443.
228. Cappellini, M.D., Porter, J.B., Viprakasit, V., and Taher, A.T. (2018). A paradigm shift on beta-thalassaemia treatment: how will we manage this old disease with new therapies? *Blood Rev.* 32, 300–311.
229. Shah, F.T., Sayani, F., Trompeter, S., Drasar, E., and Piga, A. (2019). Challenges of blood transfusions in β-thalassaemia. *Blood Rev.* 37, 100588.
230. Bonifazi, F., Conte, R., Baiardi, P., Bonifazi, D., Felisi, M., Giordano, P., Giannuzzi, V., Iacono, A., Padula, R., Pepe, A., et al.; HTA-THAL Multiregional Registry (2017). Pattern of complications and burden of disease in patients affected by beta thalassaemia major. *Curr. Med. Res. Opin.* 33, 1525–1533.
231. Telfer, P. (2009). Update on survival in thalassaemia major. *Hemoglobin* 33 (Suppl 1), S76–S80.
232. Locatelli, F., Merli, P., and Stocchio, L. (2016). Transplantation for thalassaemia major: alternative donors. *Curr. Opin. Hematol.* 23, 515–523.
233. Tiercy, J.M., and Claas, F. (2013). Impact of HLA diversity on donor selection in organ and stem cell transplantation. *Hum. Hered.* 76, 178–186.
234. Stocchio, L., Romano, M., Cefalo, M.G., Vinti, L., Gaspari, S., and Locatelli, F. (2015). Cord blood transplantation in children with hemoglobinopathies. *Expert Opin. Orphan Drugs* 3, 1125–1136.
235. Baronciani, D., Angelucci, E., Potschger, U., Gaziev, J., Yesilipek, A., Zecca, M., Orofino, M.G., Giardini, C., Al-Ahmari, A., Marktel, S., et al. (2016). Hemopoietic stem cell transplantation in thalassaemia: a report from the European Society for Blood and Bone Marrow Transplantation Hemoglobinopathy Registry, 2000–2010. *Bone Marrow Transplant.* 51, 536–541.

Molecular Therapy Methods & Clinical Development

236. Piel, F.B., Steinberg, M.H., and Rees, D.C. (2017). Sickle cell disease. *N. Engl. J. Med.* 376, 1561–1573.
237. Kato, G.J., Piel, F.B., Reid, C.D., Gaston, M.H., Ohene-Frempong, K., Krishnamurti, L., Smith, W.R., Panepinto, J.A., Weatherall, D.J., Costa, F.F., and Vichinsky, E.P. (2018). Sickle cell disease. *Nat. Rev. Dis. Primers* 4, 18010.
238. Leonard, A., and Tisdale, J.F. (2018). Stem cell transplantation in sickle cell disease: therapeutic potential and challenges faced. *Expert Rev. Hematol.* 11, 547–565.
239. Hsieh, M.M., Fitzhugh, C.D., Weitzel, R.P., Link, M.E., Coles, W.A., Zhao, X., Rodgers, G.P., Powell, J.D., and Tisdale, J.F. (2014). Nonmyeloablative HLA-matched sibling allogeneic hematopoietic stem cell transplantation for severe sickle cell phenotype. *JAMA* 312, 48–56.
240. Hsieh, M.M., Kang, E.M., Fitzhugh, C.D., Link, M.B., Bolan, C.D., Kurlander, R., Childs, R.W., Rodgers, G.P., Powell, J.D., and Tisdale, J.F. (2009). Allogeneic hematopoietic stem-cell transplantation for sickle cell disease. *N. Engl. J. Med.* 361, 2309–2317.
241. Fitzhugh, C.D., Abraham, A.A., Tisdale, J.F., and Hsieh, M.M. (2014). Hematopoietic stem cell transplantation for patients with sickle cell disease: progress and future directions. *Hematol. Oncol. Clin. North Am.* 28, 1171–1185.
242. Goodman, M.A., and Malik, P. (2016). The potential of gene therapy approaches for the treatment of hemoglobinopathies: achievements and challenges. *Ther. Adv. Hematol.* 7, 302–315.
243. Cavazzana-Calvo, M., Payen, E., Negre, O., Wang, G., Hehir, K., Fusil, F., Down, J., Denaro, M., Brady, T., Westerman, K., et al. (2010). Transfusion independence and *HMG2* activation after gene therapy of human β -thalassaemia. *Nature* 467, 318–322.
244. Ribeil, J.A., Hacein-Bey-Abina, S., Payen, E., Magnani, A., Semeraro, M., Magrin, E., Caccavelli, L., Neven, B., Bourget, P., El Nemer, W., et al. (2017). Gene therapy in a patient with sickle cell disease. *N. Engl. J. Med.* 376, 848–855.
245. Thompson, A.A., Walters, M.C., Kwiatkowski, J., Rasko, J.E.J., Ribeil, J.A., Hongeng, S., Magrin, E., Schiller, G.J., Payen, E., Semeraro, M., et al. (2018). Gene therapy in patients with transfusion-dependent β -thalassaemia. *N. Engl. J. Med.* 378, 1479–1493.
246. Marktel, S., Scaramuzza, S., Cicalese, M.P., Giglio, F., Galimberti, S., Lidonnici, M.R., Calbi, V., Assanelli, A., Bernardo, M.E., Rossi, C., et al. (2019). Intrabone hematopoietic stem cell gene therapy for adult and pediatric patients affected by transfusion-dependent β -thalassaemia. *Nat. Med.* 25, 234–241.
247. Davé, U.P., Jenkins, N.A., and Copeland, N.G. (2004). Gene therapy insertional mutagenesis insights. *Science* 303, 333.
248. Hacein-Bey Abina, S., Gaspar, H.B., Blondeau, J., Caccavelli, L., Charrier, S., Buckland, K., Picard, C., Six, E., Himoudi, N., Gilmour, K., et al. (2015). Outcomes following gene therapy in patients with severe Wiskott-Aldrich syndrome. *JAMA* 313, 1550–1563.
249. De Ravin, S.S., Wu, X., Moir, S., Anaya-O'Brien, S., Kwatema, N., Littell, P., Theobald, N., Choi, U., Su, L., Marquesen, M., et al. (2016). Lentiviral hematopoietic stem cell gene therapy for X-linked severe combined immunodeficiency. *Sci. Transl. Med.* 8, 335ra57.
250. Biffi, A., Montini, E., Liorio, L., Cesani, M., Fumagalli, F., Plati, T., Baldoli, C., Martino, S., Calabria, A., Canale, S., et al. (2013). Lentiviral hematopoietic stem cell gene therapy benefits metachromatic leukodystrophy. *Science* 341, 1233158.
251. Sessa, M., Liorio, L., Fumagalli, F., Acquati, S., Redaelli, D., Baldoli, C., Canale, S., Lopez, I.D., Morena, F., Calabria, A., et al. (2016). Lentiviral haemopoietic stem-cell gene therapy in early-onset metachromatic leukodystrophy: an ad-hoc analysis of a non-randomised, open-label, phase 1/2 trial. *Lancet* 388, 476–487.
252. Eichler, F., Duncan, C., Musolino, P.L., Orchard, P.J., De Oliveira, S., Thrasher, A.J., Armant, M., Dansereau, C., Lund, T.C., Miller, W.P., et al. (2017). Hematopoietic stem-cell gene therapy for cerebral adrenoleukodystrophy. *N. Engl. J. Med.* 377, 1630–1638.
253. Ferrua, F., Cicalese, M.P., Galimberti, S., Giannelli, S., Dionisio, F., Barzaghi, F., Migliauacca, M., Bernardo, M.E., Calbi, V., Assanelli, A.A., et al. (2019). Lentiviral haemopoietic stem/progenitor cell gene therapy for treatment of Wiskott-Aldrich syndrome: interim results of a non-randomised, open-label, phase 1/2 clinical study. *Lancet Haematol.* 6, e239–e253.
254. Mamcarz, E., Zhou, S., Lockey, T., Abdelsamed, H., Cross, S.J., Kang, G., Ma, Z., Condori, J., Dowdy, J., Triplett, B., et al. (2019). Lentiviral gene therapy combined with low-dose busulfan in infants with SCID-X1. *N. Engl. J. Med.* 380, 1525–1534.
255. Cicalese, M.P., Ferrua, F., Castagnaro, L., Rolfe, K., De Boever, E., Reinhardt, R.R., Appleby, J., Roncarolo, M.G., and Aiuti, A. (2018). Gene therapy for adenosine deaminase deficiency: a comprehensive evaluation of short- and medium-term safety. *Mol. Ther.* 26, 917–931.
256. Thein, S.L., Menzel, S., Lathrop, M., and Garner, C. (2009). Control of fetal hemoglobin: new insights emerging from genomics and clinical implications. *Hum. Mol. Genet.* 18 (R2), R216–R223.
257. Liu, N., Hargreaves, V.V., Zhu, Q., Kurland, J.V., Hong, J., Kim, W., Sher, F., Macias-Trevino, C., Rogers, J.M., Kurita, R., et al. (2018). Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* 173, 430–442.e17.
258. Chen, Z., Luo, H.Y., Steinberg, M.H., and Chui, D.H. (2009). BCL11A represses *HBB* transcription in K562 cells. *Blood Cells Mol. Dis.* 42, 144–149.
259. Sankaran, V.G., Xu, J., Ragoczy, T., Ippolito, G.C., Walkley, C.R., Maika, S.D., Fujiwara, Y., Ito, M., Groudine, M., Bender, M.A., et al. (2009). Developmental and species-divergent globin switching are driven by BCL11A. *Nature* 460, 1093–1097.
260. Chang, K.H., Smith, S.E., Sullivan, T., Chen, K., Zhou, Q., West, J.A., Liu, M., Liu, Y., Vieira, B.F., Sun, C., et al. (2017). Long-term engraftment and fetal globin induction upon *BCL11A* gene editing in bone-marrow-derived CD34⁺ hematopoietic stem and progenitor cells. *Mol. Ther. Methods Clin. Dev.* 4, 137–148.
261. Smith, E.C., Luc, S., Croney, D.M., Woodworth, M.B., Greig, L.C., Fujiwara, Y., Nguyen, M., Sher, F., Macklis, J.D., Bauer, D.E., and Orkin, S.H. (2016). Strict in vivo specificity of the *Bcl11a* erythroid enhancer. *Blood* 128, 2338–2342.
262. Psatha, N., Reik, A., Phelps, S., Zhou, Y., Dalas, D., Yannaki, E., Levasseur, D.N., Urvov, F.D., Holmes, M.C., and Papayannopoulou, T. (2018). Disruption of the BCL11A erythroid enhancer reactivates fetal hemoglobin in erythroid cells of patients with β -thalassaemia major. *Mol. Ther. Methods Clin. Dev.* 10, 313–326.
263. Chapin, J.; Vertex Pharmaceuticals Incorporated; CRISPR Therapeutics (2018) A safety and efficacy study evaluating CTX001 in subjects with transfusion-dependent β -thalassaemia. <https://clinicaltrials.gov/ct2/show/NCT03655678>.
264. Schiller, G., Walters, M., Williams, D., and Smith, A.; Sangamo Therapeutics; Sanofi. (2018). A study to assess the safety, tolerability, and efficacy of ST-400 for treatment of transfusion-dependent beta-thalassemia (TDT). <https://clinicaltrials.gov/ct2/show/NCT03432364>.
265. Chapin, J.; Vertex Pharmaceuticals Incorporated; CRISPR Therapeutics (2018). A safety and efficacy study evaluating CTX001 in subjects with severe sickle cell disease. <https://clinicaltrials.gov/ct2/show/NCT03745287>.
266. Sanofi; Bioerativ Therapeutics Inc. (2018). A study to assess the safety, tolerability, and efficacy of BIVV003 for autologous hematopoietic stem cell transplantation in patients with severe sickle cell disease (BIVV003). <https://clinicaltrials.gov/ct2/show/NCT03653247>.
267. Vertex Pharmaceuticals Incorporated; CRISPR Therapeutics (2019). A long-term follow-up study in subjects who received CTX001. <https://clinicaltrials.gov/ct2/show/NCT04208529>.
268. Traxler, E.A., Yao, Y., Wang, Y.D., Woodard, K.J., Kurita, R., Nakamura, Y., Hughes, J.R., Hardison, R.C., Blobel, G.A., Li, C., and Weiss, M.J. (2016). A genome-editing strategy to treat β -globinopathies that recapitulates a mutation associated with a benign genetic condition. *Nat. Med.* 22, 987–990.
269. Lux, C.T., Pattabhi, S., Berger, M., Nourigat, C., Flowers, D.A., Negre, O., Humbert, O., Yang, J.G., Lee, C., Jacoby, K., et al. (2018). TALEN-mediated gene editing of *HBB* in human hematopoietic stem cells leads to therapeutic fetal hemoglobin induction. *Mol. Ther. Methods Clin. Dev.* 12, 175–183.
270. Ye, L., Wang, J., Tan, Y., Beyer, A.I., Xie, F., Muench, M.O., and Kan, Y.W. (2016). Genome editing using CRISPR-Cas9 to create the HPFH genotype in HSPCs: an approach for treating sickle cell disease and β -thalassaemia. *Proc. Natl. Acad. Sci. USA* 113, 10661–10665.
271. Cai, L., Bai, H., Mahairaki, V., Gao, Y., He, C., Wen, Y., Jin, Y.C., Wang, Y., Pan, R.L., Qasba, A., et al. (2018). A universal approach to correct various *HBB* gene mutations in human stem cells for gene therapy of beta-thalassemia and sickle cell disease. *Stem Cells Transl. Med.* 7, 87–97.

272. Wattanapitch, M., Damkham, N., Potirat, P., Trakarnsanga, K., Janan, M., U-Pratya, Y., Kheolamai, P., Klincumhom, N., and Issaragrisil, S. (2018). One-step genetic correction of hemoglobin E/beta-thalassemia patient-derived iPSCs by the CRISPR/Cas9 system. *Stem Cell Res. Ther.* 9, 46.
273. Martin, R.M., Ikeda, K., Cromer, M.K., Uchida, N., Nishimura, T., Romano, R., Tong, A.J., Lemgart, V.T., Camarena, J., Pavel-Dinu, M., et al. (2019). Highly efficient and marker-free genome editing of human pluripotent stem cells by CRISPR-Cas9 RNP and AAV6 donor-mediated homologous recombination. *Cell Stem Cell* 24, 821–828.e5.
274. Park, S.H., Lee, C.M., Dever, D.P., Davis, T.H., Camarena, J., Srifa, W., Zhang, Y., Paikari, A., Chang, A.K., Porteus, M.H., et al. (2019). Highly efficient editing of the β -globin gene in patient-derived hematopoietic stem and progenitor cells to treat sickle cell disease. *Nucleic Acids Res.* 47, 7955–7972.
275. Xie, F., Ye, L., Chang, J.C., Beyer, A.L., Wang, J., Muench, M.O., and Kan, Y.W. (2014). Seamless gene correction of β -thalassaemia mutations in patient-specific iPSCs using CRISPR/Cas9 and piggyBac. *Genome Res.* 24, 1526–1533.
276. Song, B., Fan, Y., He, W., Zhu, D., Niu, X., Wang, D., Ou, Z., Luo, M., and Sun, X. (2015). Improved hematopoietic differentiation efficiency of gene-corrected beta-thalassemia induced pluripotent stem cells by CRISPR/Cas9 system. *Stem Cells Dev.* 24, 1053–1065.
277. Xu, P., Tong, Y., Liu, X.Z., Wang, T.T., Cheng, L., Wang, B.Y., Lv, X., Huang, Y., and Liu, D.P. (2015). Both TALENs and CRISPR/Cas9 directly target the HBB IVS2-654 (C > T) mutation in β -thalassaemia-derived iPSCs. *Sci. Rep.* 5, 12065.
278. Niu, X., He, W., Song, B., Ou, Z., Fan, D., Chen, Y., Fan, Y., and Sun, X. (2016). Combining single strand oligodeoxynucleotides and CRISPR/Cas9 to correct gene mutations in β -thalassaemia-induced pluripotent stem cells. *J. Biol. Chem.* 291, 16576–16585.
279. Liu, Y., Yang, Y., Kang, X., Lin, B., Yu, Q., Song, B., Gao, G., Chen, Y., Sun, X., Li, X., et al. (2017). One-step allelic and scarless correction of a β -thalassaemia mutation in patient-specific iPSCs without Drug Selection. *Mol. Ther. Nucleic Acids* 6, 57–67.
280. Alllife Medical Science and Technology Co., Ltd. (2018). iHSCs with the gene correction of HBB intervent subjects with β -thalassaemia mutations. <https://clinicaltrials.gov/ct2/show/NCT03728322>.
281. Mettananda, S., Fisher, C.A., Hay, D., Badat, M., Quek, L., Clark, K., Hublitz, P., Downes, D., Kerry, J., Gosden, M., et al. (2017). Editing an α -globin enhancer in primary human hematopoietic stem cells as a treatment for β -thalassaemia. *Nat. Commun.* 8, 424.
282. VandenDriessche, T., and Chuah, M.K. (2017). Hemophilia gene therapy: ready for prime time? *Hum. Gene Ther.* 28, 1013–1023.
283. Franchini, M., and Mannucci, P.M. (2012). Past, present and future of hemophilia: a narrative review. *Orphanet J. Rare Dis.* 7, 24.
284. Powell, J.S., Ragni, M.V., White, G.C., 2nd, Lusher, J.M., Hillman-Wiseman, C., Moon, T.E., Cole, V., Ramanathan-Girish, S., Roehl, H., Sajjadi, N., et al. (2003). Phase I trial of FVIII gene transfer for severe hemophilia A using a retroviral construct administered by peripheral intravenous infusion. *Blood* 102, 2038–2045.
285. Kay, M.A., Manno, C.S., Ragni, M.V., Larson, P.J., Couto, L.B., McClelland, A., Glader, B., Chew, A.J., Tai, S.J., Herzog, R.W., et al. (2000). Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* 24, 257–261.
286. Manno, C.S., Pierce, G.F., Arruda, V.R., Glader, B., Ragni, M., Rasko, J.J., Ozelo, M.C., Hoots, K., Blatt, P., Konkle, B., et al. (2006). Successful transduction of liver in hemophilia by AAV-factor IX and limitations imposed by the host immune response. *Nat. Med.* 12, 342–347.
287. Roth, D.A., Tawa, N.E., Jr., O'Brien, J.M., Treco, D.A., and Selden, R.F.; Factor VIII Transkaryotic Therapy Study Group (2001). Nonviral transfer of the gene encoding coagulation factor VIII in patients with severe hemophilia A. *N. Engl. J. Med.* 344, 1735–1742.
288. Sharma, R., Anguela, X.M., Doyon, Y., Wechsler, T., DeKolver, R.C., Sproul, S., Paschon, D.E., Miller, J.C., Davidson, R.J., Shivak, D., et al. (2015). In vivo genome editing of the albumin locus as a platform for protein replacement therapy. *Blood* 126, 1777–1784.
289. Quon, D., Kuriakose, P.; Sangamo Therapeutics (2016). Ascending dose study of genome editing by zinc finger nuclease therapeutic SB-FIX in subjects with severe hemophilia B. <https://clinicaltrials.gov/ct2/show/NCT02695160>.
290. Lyu, C., Shen, J., Wang, R., Gu, H., Zhang, J., Xue, F., Liu, X., Liu, W., Fu, R., Zhang, L., et al. (2018). Targeted genome engineering in human induced pluripotent stem cells from patients with hemophilia B using the CRISPR-Cas9 system. *Stem Cell Res. Ther.* 9, 92.
291. Sivalingam, J., Kenanov, D., Han, H., Nirmal, A.J., Ng, W.H., Lee, S.S., Masilamani, J., Phan, T.T., Maurer-Stroh, S., and Kon, O.L. (2016). Multidimensional genome-wide analyses show accurate FVIII integration by ZFN in primary human cells. *Mol. Ther.* 24, 607–619.
292. Li, H., Haurigot, V., Doyon, Y., Li, T., Wong, S.Y., Bhagwat, A.S., Malani, N., Anguela, X.M., Sharma, R., Ivanciu, L., et al. (2011). In vivo genome editing restores haemostasis in a mouse model of haemophilia. *Nature* 475, 217–221.
293. Anguela, X.M., Sharma, R., Doyon, Y., Miller, J.C., Li, H., Haurigot, V., Rohde, M.E., Wong, S.Y., Davidson, R.J., Zhou, S., et al. (2013). Robust ZFN-mediated genome editing in adult hemophilic mice. *Blood* 122, 3283–3287.
294. Bergmann, T., Ehrke-Schulz, E., Gao, J., Schiwoon, M., Schildgen, V., David, S., Schildgen, O., and Ehrhardt, A. (2018). Designer nuclease-mediated gene correction via homology-directed repair in an in vitro model of canine hemophilia B. *J. Gene Med.* 20, e3020.
295. Ohmori, T., Nagao, Y., Mizukami, H., Sakata, A., Muramatsu, S.I., Ozawa, K., Tominaga, S.I., Hanazono, Y., Nishimura, S., Nureki, O., and Sakata, Y. (2017). CRISPR/Cas9-mediated genome editing via postnatal administration of AAV vector cures haemophilia B mice. *Sci. Rep.* 7, 4159.
296. Guan, Y., Ma, Y., Li, Q., Sun, Z., Ma, L., Wu, L., Wang, L., Zeng, L., Shao, Y., Chen, Y., et al. (2016). CRISPR/Cas9-mediated somatic correction of a novel coagulator factor IX gene mutation ameliorates hemophilia in mouse. *EMBO Mol. Med.* 8, 477–488.
297. He, Q., Wang, H.H., Cheng, T., Yuan, W.P., Ma, Y.P., Jiang, Y.P., and Ren, Z.H. (2017). Genetic correction and hepatic differentiation of hemophilia B-specific human induced pluripotent stem cells. *Chin. Med. Sci. J.* 32, 135–144.
298. Huai, C., Jia, C., Sun, R., Xu, P., Min, T., Wang, Q., Zheng, C., Chen, H., and Lu, D. (2017). CRISPR/Cas9-mediated somatic and germline gene correction to restore hemostasis in hemophilia B mice. *Hum. Genet.* 136, 875–883.
299. Park, C.Y., Kim, D.H., Son, J.S., Sung, J.J., Lee, J., Bae, S., Kim, J.H., Kim, D.W., and Kim, J.S. (2015). Functional correction of large factor VIII gene chromosomal inversions in hemophilia A patient-derived iPSCs using CRISPR-Cas9. *Cell Stem Cell* 17, 213–220.
300. Park, C.Y., Kim, J., Kweon, J., Son, J.S., Lee, J.S., Yoo, J.E., Cho, S.R., Kim, J.H., Kim, J.S., and Kim, D.W. (2014). Targeted inversion and reversion of the blood coagulation factor 8 gene in human iPSC cells using TALENs. *Proc. Natl. Acad. Sci. USA* 111, 9253–9258.
301. Park, C.Y., Sung, J.J., Choi, S.H., Lee, D.R., Park, I.H., and Kim, D.W. (2016). Modeling and correction of structural variations in patient-derived iPSCs using CRISPR/Cas9. *Nat. Protoc.* 11, 2154–2169.
302. Wagenblast, E., Azkanaz, M., Smith, S.A., Shakib, L., McLeod, J.L., Krivdova, G., Araújo, J., Shultz, L.D., Gan, O.L., Dick, J.E., and Lechman, E.R. (2019). Functional profiling of single CRISPR/Cas9-edited human long-term hematopoietic stem cells. *Nat. Commun.* 10, 4730.
303. Genovese, P., Schirolli, G., Escobar, G., Tomaso, T.D., Firrito, C., Calabria, A., Moi, D., Mazzieri, R., Bonini, C., Holmes, M.C., et al. (2014). Targeted genome editing in human repopulating hematopoietic stem cells. *Nature* 510, 235–240.
304. De Ravin, S.S., Reik, A., Liu, P.Q., Li, L., Wu, X., Su, L., Raley, C., Theobald, N., Choi, U., Song, A.H., et al. (2016). Targeted gene addition in human CD34⁺ hematopoietic cells for correction of X-linked chronic granulomatous disease. *Nat. Biotechnol.* 34, 424–429.
305. Schirolli, G., Ferrari, S., Conway, A., Jacob, A., Capo, V., Albano, L., Plati, T., Castiello, M.C., Sanvito, F., Gennery, A.R., et al. (2017). Preclinical modeling highlights the therapeutic potential of hematopoietic stem cell gene editing for correction of SCID-X1. *Sci. Transl. Med.* 9, ean0820.
306. Sawamoto, K., Chen, H.H., Alméciga-Díaz, C.J., Mason, R.W., and Tomatsu, S. (2018). Gene therapy for mucopolysaccharidoses. *Mol. Genet. Metab.* 123, 59–68.

Molecular Therapy Methods & Clinical Development

307. Concolino, D., Deodato, F., and Parini, R. (2018). Enzyme replacement therapy: efficacy and limitations. *Ital. J. Pediatr.* *44* (Suppl 2), 120.
308. Tomatsu, S., Almcéiga-Díaz, C.J., Montaño, A.M., Yabe, H., Tanaka, A., Dung, V.C., Giugliani, R., Kubaski, F., Mason, R.W., Yasuda, E., et al. (2015). Therapies for the bone in mucopolysaccharidoses. *Mol. Genet. Metab.* *114*, 94–109.
309. Taylor, M., Khan, S., Stapleton, M., Wang, J., Chen, J., Wynn, R., Yabe, H., Chinen, Y., Boelens, J.J., Mason, R.W., et al. (2019). Hematopoietic stem cell transplantation for mucopolysaccharidoses: past, present, and future. *Biol. Blood Marrow Transplant.* *25*, e226–e246.
310. Schuh, R.S., Poletto, E., Pasqualim, G., Tavares, A.M.V., Meyer, F.S., Gonzalez, E.A., Giugliani, R., Matte, U., Teixeira, H.F., and Baldo, G. (2018). In vivo genome editing of mucopolysaccharidosis I mice using the CRISPR/Cas9 system. *J. Control. Release* *288*, 23–33.
311. Ellinwood, N.M., Ausseil, J., Desmaris, N., Bigou, S., Liu, S., Jens, J.K., Snella, E.M., Mohammed, E.E., Thomson, C.B., Raoul, S., et al. (2011). Safe, efficient, and reproducible gene therapy of the brain in the dog models of Sanfilippo and Hurler syndromes. *Mol. Ther.* *19*, 251–259.
312. Motas, S., Haurigot, V., Garcia, M., Marcó, S., Ribera, A., Roca, C., Sánchez, X., Sánchez, V., Molas, M., Bertolin, J., et al. (2016). CNS-directed gene therapy for the treatment of neurologic and somatic mucopolysaccharidosis type II (Hunter syndrome). *JCI Insight* *1*, e86696.
313. Fu, H., Dirosario, J., Killeddar, S., Zaraspe, K., and McCarty, D.M. (2011). Correction of neurological disease of mucopolysaccharidosis IIIB in adult mice by rAAV9 trans-blood-brain barrier gene delivery. *Mol. Ther.* *19*, 1025–1033.
314. Sorrentino, N.C., D'Orsi, L., Sambri, L., Nusco, E., Monaco, C., Spannato, C., Polishchuk, E., Saccone, P., De Leonibus, E., Ballabio, A., and Fraldi, A. (2013). A highly secreted sulphamidase engineered to cross the blood-brain barrier corrects brain lesions of mice with mucopolysaccharidoses type IIIB. *EMBO Mol. Med.* *5*, 675–690.
315. Tessitore, A., Faella, A., O'Malley, T., Cotugno, G., Doria, M., Kunieda, T., Matarese, G., Haskins, M., and Auricchio, A. (2008). Biochemical, pathological, and skeletal improvement of mucopolysaccharidosis VI after gene transfer to liver but not to muscle. *Mol. Ther.* *16*, 30–37.
316. Gurda, B.L., De Guilhem De Lataillade, A., Bell, P., Zhu, Y., Yu, H., Wang, P., Bagel, J., Vite, C.H., Sikora, T., Hinderer, C., et al. (2016). Evaluation of AAV-mediated gene therapy for central nervous system disease in canine mucopolysaccharidosis VII. *Mol. Ther.* *24*, 206–216.
317. Tardieu, M., Zérab, M., Gougeon, M.L., Ausseil, J., de Bournoville, S., Husson, B., Zaferiou, D., Parenti, G., Bourget, P., Poirier, B., et al. (2017). Intracerebral gene therapy in children with mucopolysaccharidosis type IIIB syndrome: an uncontrolled phase 1/2 clinical trial. *Lancet Neurol.* *16*, 712–720.
318. Laoharawee, K., DeKelver, R.C., Podetz-Pedersen, K.M., Rohde, M., Sproul, S., Nguyen, H.O., Nguyen, T., St Martin, S.J., Ou, L., Tom, S., et al. (2018). Dose-dependent prevention of metabolic and neurologic disease in murine MPS II by ZFN-mediated in vivo genome editing. *Mol. Ther.* *26*, 1127–1136.
319. Harmatz, P., Heldermon, C., Wilcox, W., Whitley, C., Lau, H., and Leslie, N.; Sangamo Therapeutics (2018). Ascending dose study of genome editing by the zinc finger nuclease (ZFN) therapeutic SB-318 in subjects with MPS I. <https://clinicaltrials.gov/ct2/show/NCT02702115>.
320. Burton, B., Whitley, C., Lau, H., Muenzer, J., Prada, C., and Ficcioglu, C.; Sangamo Therapeutics (2017). Ascending dose study of genome editing by the zinc finger nuclease (ZFN) therapeutic SB-913 in subjects with MPS II. <https://clinicaltrials.gov/ct2/show/NCT03041324>.
321. Kumaran, N., Moore, A.T., Weleber, R.G., and Michaelides, M. (2017). Leber congenital amaurosis/early-onset severe retinal dystrophy: clinical features, molecular genetics and therapeutic interventions. *Br. J. Ophthalmol.* *101*, 1147–1154.
322. May-Simera, H., Nagel-Wolfrum, K., and Wolfrum, U. (2017). Cilia—the sensory antennae in the eye. *Prog. Retin. Eye Res.* *60*, 144–180.
323. Jacobson, S.G., Cideciyan, A.V., Ratnakaram, R., Heon, E., Schwartz, S.B., Roman, A.J., Peden, M.C., Aleman, T.S., Boye, S.L., Sumaroka, A., et al. (2012). Gene therapy for leber congenital amaurosis caused by RPE65 mutations: safety and efficacy in 15 children and adults followed up to 3 years. *Arch. Ophthalmol.* *130*, 9–24.
324. Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balagun, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., et al. (2008). Effect of gene therapy on visual function in Leber's congenital amaurosis. *N. Engl. J. Med.* *358*, 2231–2239.
325. Russell, S., Bennett, J., Wellman, J.A., Chung, D.C., Yu, Z.F., Tillman, A., Wittes, J., Pappas, J., Elci, O., McCague, S., et al. (2017). Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet* *390*, 849–860.
326. Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh, E.N., Jr., Mingozzi, F., Bencicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., et al. (2008). Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N. Engl. J. Med.* *358*, 2240–2248.
327. Sharif, W., and Sharif, Z. (2017). Leber's congenital amaurosis and the role of gene therapy in congenital retinal disorders. *Int. J. Ophthalmol.* *10*, 480–484.
328. Xu, C.L., Cho, G.Y., Sengillo, J.D., Park, K.S., Mahajan, V.B., and Tsang, S.H. (2018). Translation of CRISPR genome surgery to the bedside for retinal diseases. *Front. Cell Dev. Biol.* *6*, 46.
329. Allergan; Editas Medicine, Inc. (2019). Single ascending dose study in participants with LCA10. <https://clinicaltrials.gov/ct2/show/NCT03872479>.
330. Coppieters, F., Lefever, S., Leroy, B.P., and De Baere, E. (2010). CEP290, a gene with many faces: mutation overview and presentation of CEP290base. *Hum. Mutat.* *31*, 1097–1108.
331. Ruan, G.X., Barry, E., Yu, D., Lukason, M., Cheng, S.H., and Scaria, A. (2017). CRISPR/Cas9-mediated genome editing as a therapeutic approach for Leber congenital amaurosis 10. *Mol. Ther.* *25*, 331–341.
332. Maeder, M.L., Stefanidakis, M., Wilson, C.J., Baral, R., Barrera, L.A., Bounoutas, G.S., Bumcrot, D., Chao, H., Ciulla, D.M., DaSilva, J.A., et al. (2019). Development of a gene-editing approach to restore vision loss in Leber congenital amaurosis type 10. *Nat. Med.* *25*, 229–233.
333. Sanjurjo-Soriano, C., and Kalatzis, V. (2018). Guiding lights in genome editing for inherited retinal disorders: implications for gene and cell therapy. *Neural Plast.* *2018*, 5056279.
334. Xu, C.L., Park, K.S., and Tsang, S.H. (2018). CRISPR/Cas9 genome surgery for retinal diseases. *Drug Discov. Today. Technol.* *28*, 23–32.
335. Zhang, Z.Y., Thrasher, A.J., and Zhang, F. (2019). Gene therapy and genome editing for primary immunodeficiency diseases. *Genes Dis.* *7*, 38–51.
336. Booth, C., Romano, R., Roncarolo, M.G., and Thrasher, A.J. (2019). Gene therapy for primary immunodeficiency. *Hum. Mol. Genet.* *28* (R1), R15–R23.
337. VanLith, C.J., Guthman, R.M., Nicolas, C.T., Allen, K.L., Liu, Y., Chilton, J.A., Tritz, Z.P., Nyberg, S.L., Kaiser, R.A., Lilligard, J.B., and Hickey, R.D. (2019). Ex vivo hepatocyte reprogramming promotes homology-directed DNA repair to correct metabolic disease in mice after transplantation. *Hepatol. Commun.* *3*, 558–573.
338. Vaidyanathan, S., Salahudeen, A.A., Sellers, Z.M., Bravo, D.T., Choi, S.S., Batish, A., Le, W., Baik, R., de la O, S., Kaushik, M.P., et al. (2020). High-efficiency, selection-free gene repair in airway stem cells from cystic fibrosis patients rescues CFTR function in differentiated epithelia. *Cell Stem Cell* *26*, 161–171.e4.
339. Nissanka, N., and Moraes, C.T. (2020). Mitochondrial DNA heteroplasmy in disease and targeted nuclease-based therapeutic approaches. *EMBO Rep.* *21*, e49612.
340. Zekonyte, U., Bacman, R., and Moraes, C.T. (2020). DNA-editing enzymes as potential treatments for heteroplasmic mtDNA diseases. *J. Intern. Med.* *287*, 685–697.
341. Pankowicz, F.P., Barzi, M., Legras, X., Hubert, L., Mi, T., Tomolonis, J.A., Ravishanker, M., Sun, Q., Yang, D., Borowiak, M., et al. (2016). Reprogramming metabolic pathways in vivo with CRISPR/Cas9 genome editing to treat hereditary tyrosinaemia. *Nat. Commun.* *7*, 12642.
342. Villerger, L., Grisch-Can, H.M., Lindsay, H., Ringnald, F., Pogliano, C.B., Allegri, G., Fingerhut, R., Häberle, J., Matos, J., Robinson, M.D., et al. (2018). Treatment of a metabolic liver disease by in vivo genome base editing in adult mice. *Nat. Med.* *24*, 1519–1525.
343. Lu, T., Yang, B., Wang, R., and Qin, C. (2020). Xenotransplantation: current status in preclinical research. *Front. Immunol.* *10*, 3060.
344. Ferrari, S., Jacob, A., Beretta, S., Unali, G., Albano, L., Vavassori, V., et al. (2020). Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-020-0551-y>.



CHAPTER 6

LARGE-SCALE EXPANSION OF HUMAN IPSC-DERIVED SKELETAL MUSCLE CELLS FOR DISEASE MODELING AND CELL-BASED THERAPEUTIC STRATEGIES

Erik van der Wal, Pablo Herrero-Hernandez, Raymond Wan, **Mike Broeders**,
Stijn L.M. in 't Groen, Tom J.M. van Gestel, Wilfred F.J. van IJcken, Tom H. Cheung,
Ans T. van der Ploeg, Gerben J. Schaaf, W.W.M. Pim Pijnappel

keywords: skeletal muscle differentiation; pluripotent stem cells; disease modeling
muscle regeneration; satellite cells; gene editing; CRISPR/Cas9; Pompe disease;
lysosomal storage disease; metabolic disease

Large-Scale Expansion of Human iPSC-Derived Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies

Erik van der Wal,^{1,2,3} Pablo Herrero-Hernandez,^{1,2,3} Raymond Wan,⁴ Mike Broeders,^{1,2,3} Stijn L.M. in 't Groen,^{1,2,3} Tom J.M. van Gestel,^{1,2,3} Wilfred F.J. van IJcken,⁵ Tom H. Cheung,⁴ Ans T. van der Ploeg,^{2,3} Gerben J. Schaaf,^{1,2,3} and W.W.M. Pim Pijnappel^{1,2,3,*}

¹Department of Clinical Genetics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

²Department of Pediatrics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

³Center for Lysosomal and Metabolic Diseases, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

⁴Division of Life Science, Center for Stem Cell Research, Center of Systems Biology and Human Health, State Key Laboratory in Molecular Neuroscience, Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong 999077, China

⁵Erasmus Center for Biomics, Erasmus University Medical Center, 3000 CA Rotterdam, Netherlands

*Correspondence: w.pijnappel@erasmusmc.nl

<https://doi.org/10.1016/j.stemcr.2018.04.002>

SUMMARY

Although skeletal muscle cells can be generated from human induced pluripotent stem cells (iPSCs), transgene-free protocols include only limited options for their purification and expansion. In this study, we found that fluorescence-activated cell sorting-purified myogenic progenitors generated from healthy controls and Pompe disease iPSCs can be robustly expanded as much as 5×10^{11} -fold. At all steps during expansion, cells could be cryopreserved or differentiated into myotubes with a high fusion index. *In vitro*, cells were amenable to maturation into striated and contractile myofibers. Insertion of *acid α -glucosidase* cDNA into the *AAVS1* locus in iPSCs using CRISPR/Cas9 prevented glycogen accumulation in myotubes generated from a patient with classic infantile Pompe disease. *In vivo*, the expression of human-specific nuclear and sarcolemmal antigens indicated that myogenic progenitors engraft into murine muscle to form human myofibers. This protocol is useful for modeling of skeletal muscle disorders and for using patient-derived, gene-corrected cells to develop cell-based strategies.

INTRODUCTION

Although over 700 human genetic disorders are known that affect skeletal muscle (Kaplan and Hamroun, 2015), very few therapies are available. Skeletal muscle nonetheless has a high capacity for regeneration after injury (Baghdadi and Tajbakhsh, 2017; Bursac et al., 2015; Dumont et al., 2015). Muscle regeneration is mediated by satellite cells (SCs) (Lepper et al., 2011; Murphy et al., 2011; Sambasivan et al., 2011); i.e., adult stem cells located between the sarcolemma and the plasma membrane (Mauro, 1961) that are quiescent in healthy, uninjured muscle. Upon injury, SCs expand to contribute to fiber formation and to self-renew the SC pool.

SCs are considered useful for *in vitro* disease modeling to investigate molecular mechanisms of disease, test drugs, or develop cell-based therapies. To decipher molecular mechanisms of disease, it is important to generate isogenic controls, given the high variability of gene expression and functional parameters between individuals (Hockemeyer and Jaenisch, 2016; Soldner et al., 2011). To develop cell-based therapy, the ultimate goal is to engraft gene-corrected, autologous cells. However, it has not proved easy to date to establish robust *in vitro* disease models for skeletal muscle disorders, to efficiently restore gene function in skeletal muscle cells, and to develop cell-based therapeutic strategies based on muscle regeneration.

Pluripotent stem cells (PSCs) offer a potential source of skeletal muscle cells. PSCs, including induced PSCs (iPSCs), are easily expanded and maintain their full stem cell potential (Takahashi and Yamanaka, 2016). Differentiation of PSCs to SC-like cells was difficult until the recent development of two major strategies, the first involving the inducible overexpression of PAX7, the master transcription factor for SCs (Darabi et al., 2012). After generation from human embryonic stem cells and iPSCs, purified SC-like cells showed capacity for *in vitro* expansion and differentiation, and also for *in vivo* engraftment and contribution to muscle-fiber formation in immunodeficient mice (Darabi et al., 2012; Magli et al., 2017). The second strategy involved the use of small molecules to develop transgene-free differentiation. After using GSK3 β inhibition to activate the Wnt pathway, the basic procedure consists of treatment with fibroblast growth factor 2 (FGF2) and culturing in a minimal medium (see Table S1) (Borchin et al., 2013; Caron et al., 2016; Shelton et al., 2014, 2016; van der Wal et al., 2017b; Xu et al., 2013). In some cases, differentiation into the myogenic lineage has been promoted by including BMP4 inhibition (Chal et al., 2015, 2016; Swartz et al., 2016). In others, FGF2 has been replaced by the Notch signaling inhibitor DAPT (Choi et al., 2016).

Transgene-free protocols can be divided into those that use fluorescence-activated cell sorting (FACS) purification (Borchin et al., 2013; Choi et al., 2016; van der Wal et al.,

2017b) and those that use unpurified cell mixtures or partial purification through preplating (Caron et al., 2016; Chal et al., 2015; Shelton et al., 2014; Swartz et al., 2016; Xu et al., 2013) (Table S1). Upon terminal differentiation *in vitro*, unpurified/partially purified myogenic progenitors showed matured myotubes and even myofibers (Chal et al., 2015, 2016; Swartz et al., 2016). Three reports showed engraftment of myogenic cells from unpurified cultures into immunodeficient mice (Choi et al., 2016; Kim et al., 2017; Xu et al., 2013). Choi et al. (2016) reported that purification of myogenic progenitors by FACS resulted in myogenic progenitors that could be expanded 10⁵-fold. Upon *in vitro* differentiation to myotubes, these cells also showed a low (10%–15%) fusion index (Table S1).

In vivo engraftment of purified myogenic progenitors using a transgene-free procedure has not been reported so far. Similarly, it has not been possible yet to expand transgene-free, purified myogenic progenitors and differentiate and mature these cells to myotubes with high fusion index. Recently, we have modified a protocol by Borchin et al. (2013) for the transgene-free differentiation of human iPSC into SC-like cells, and used a simplified FACS purification procedure that selects C-MET-expressing cells that are HNK negative (Borchin et al., 2013; van der Wal et al., 2017b). The purified cells could be expanded at least 5×10^7 -fold and cryopreserved. At any point during the expansion, cells could be differentiated into myotubes with a high (60%–80%) fusion index. We have applied this protocol to model Pompe disease, which is a progressive inheritable metabolic myopathy caused by deficiency of acid α -glucosidase (GAA), resulting in lysosomal glycogen accumulation (van der Ploeg and Reuser, 2008). This protocol allowed the quantitative analysis of the effects of antisense oligonucleotides designed to restore canonical pre-mRNA splicing of GAA in skeletal muscle cells from Pompe patients (van der Wal et al., 2017a).

Here, we further explored the expansion capacity and the *in vitro* and *in vivo* potential of myogenic progenitors, generated from iPSCs in a transgene-free manner and FACS purified, for the future development of therapies for skeletal muscle disorders.

RESULTS

Optimization of the Generation of Myogenic Progenitors from iPSCs

As a starting point, we took the protocol published by Borchin et al. (2013), which we had modified recently (van der Wal et al., 2017b). This protocol consists of treating human iPSCs first with the GSK3 β inhibitor CHIR99021, then with FGF2, followed by prolonged culturing in minimal medium. The treatment with

CHIR99021 is a critical step, as too-low concentrations fail to yield myogenic progenitors, while too-high concentrations can be toxic. The optimal concentration most likely depends on the cell culture conditions used. We assume, for example, that the outcome can be affected by culturing iPSCs with or without feeders.

In our experiments, we cultured iPSCs on γ -irradiated mouse embryonic fibroblasts. To determine the optimal treatment with CHIR99021, we varied the concentration and duration of treatment and scored for confluency and PAX7 expression (Table S2). The results in two independent iPSC lines showed that the highest number of PAX7⁺ cells was induced after 4–5 days at a concentration of 4 μ M CHIR99021 in the absence of toxicity. To avoid any risk of toxicity in subsequent experiments, we chose 5-day incubation at a concentration of 3.5 μ M CHIR99021.

Robustness of the Myogenic Differentiation Protocol

As outlined in Figure 1A, we used primary fibroblast-derived iPSCs from 15 different donors, applying the myogenic differentiation procedure in over 50 individual differentiation experiments. Eight of these iPSC lines were derived from healthy individuals, while seven were from patients with Pompe disease. Figure 1B shows robust generation of PAX7⁺ areas in six examples of healthy control iPSCs after 35 days of differentiation as described previously (van der Wal et al., 2017b). During the differentiation procedure, phase-contrast microscopy showed small colonies with a confluency of between 20% and 40% at day 1 (Figure S1A). After 5 days of culture, iPSC colonies had reached a medium size. At this stage CHIR99021 treatment was started. After 5 days of incubation, we observed increased cell detachment, which was attenuated after a further 3–4 days in FGF2-containing medium. From day 17 onwards, the cells started to proliferate rapidly, and cultures reached complete confluency after 24 days. Multinucleated myotube-like cells were observed between 30 and 40 days. During this differentiation procedure, we observed similar morphological changes in all iPSC lines (Figure S1A and data not shown).

Differentiation of 59 cultures from a total of 15 donors yielded an average of $4.26\% \pm 3.96\%$ of C-MET⁺/Hoechst⁺/HNK-1[−] cells (Figure 1C). There were no significant differences in the number of C-MET⁺/Hoechst⁺ cells between iPSCs from healthy controls and from Pompe patients. Sorting differentiation cultures with low levels of C-MET⁺/Hoechst⁺/HNK-1[−] cells ($\sim 0.2\%$ of cells) resulted in expandable myogenic progenitors whose differentiation capacity was similar to that of cultures with a high recovery ($>2\%$) (data not shown). C-MET[−]/HNK-1⁺ cells were unable to form myosin heavy chain (MHC)-positive cells after 4 days of differentiation (data not shown). After 24 hr of plating, sorted myogenic progenitors revealed a rather

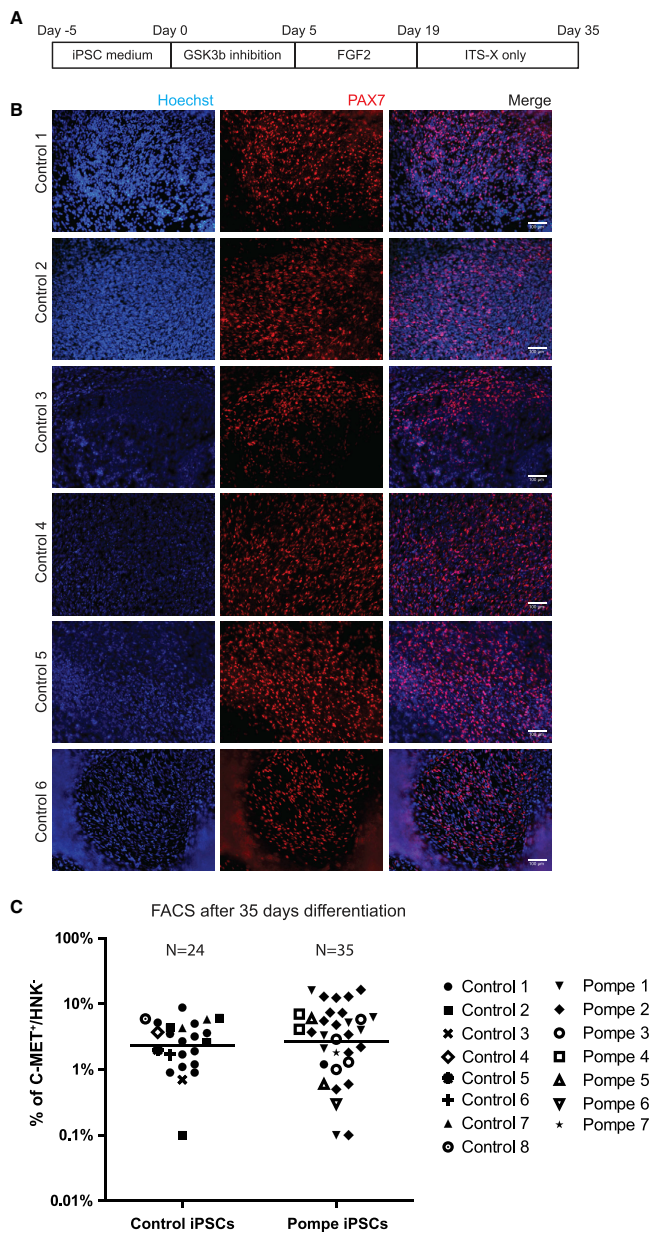
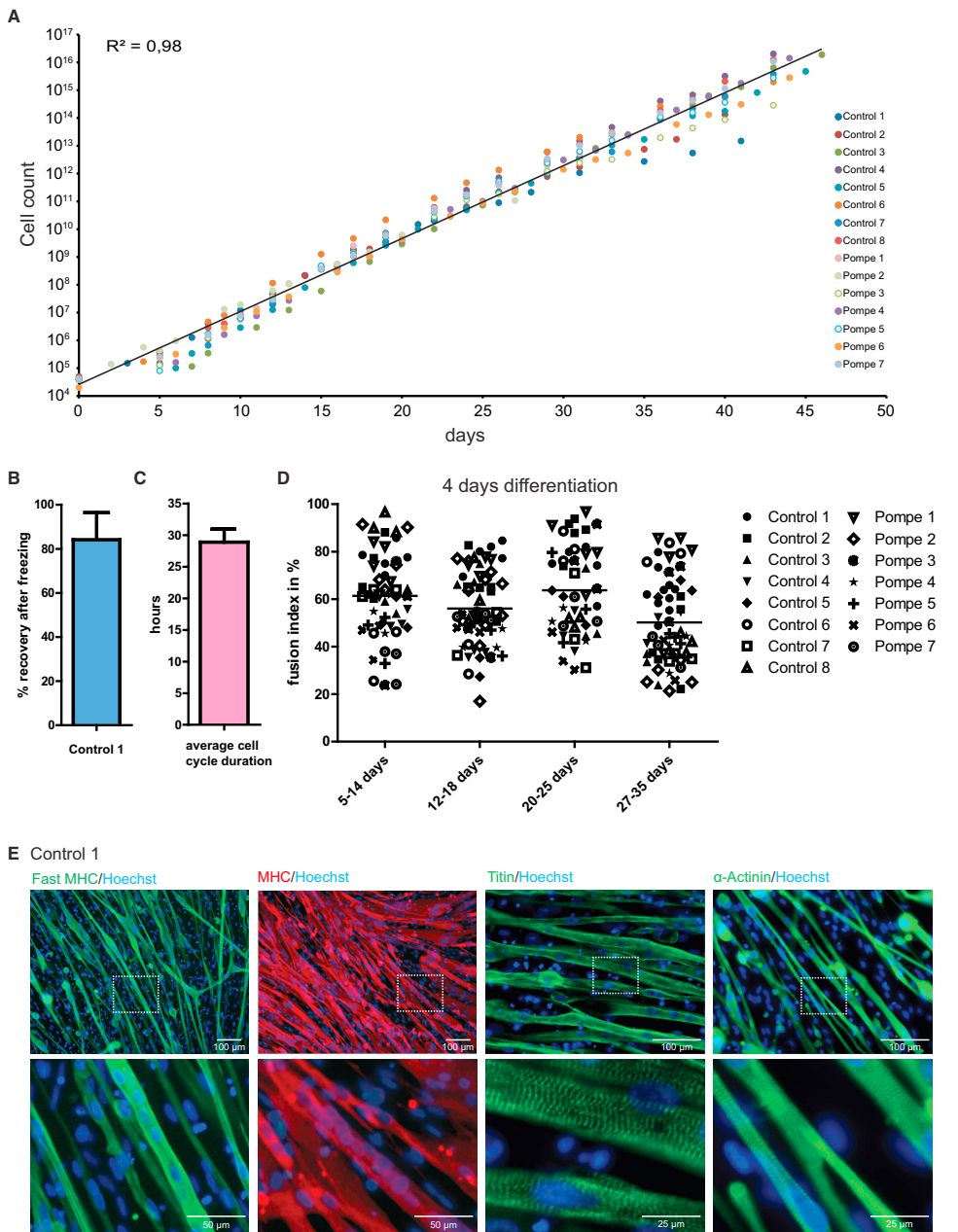


Figure 1. Robustness of Generation and Purification of Myogenic Progenitors from iPSCs

(A) Scheme for myogenic differentiation of iPSCs.

(B) Representative examples of PAX7⁺ cells obtained in the original culture dishes from six different control iPSC lines that were differentiated using a 35-day protocol consisting of consecutive treatment with CHIR, FGF2, and minimal medium (Borchin et al., 2013; van der Wal et al., 2017b). The other two control iPSC lines showed similar patches of PAX7⁺ cells (data not shown). Red: PAX7⁺ nuclei using immunofluorescent staining. Blue: nuclei stained with Hoechst. For images of the plates during this 35-day protocol see Figure S1A.

(C) Differentiations described in (B) were purified using a one-step FACS purification based on selection for C-MET⁺ myogenic cells and counter selection of HNK1⁺ neural crest cells (Borchin et al., 2013). Results are shown for 59 differentiations performed on iPSCs derived from eight healthy controls and seven Pompe patients. Each symbol represents an individual differentiation experiment. Means are indicated by horizontal lines. A total number of 24 differentiations of control iPSCs and 35 differentiations of Pompe iPSCs were performed.



(legend on next page)

uniform morphology (Figure S1B). These results demonstrated that the differentiation protocol robustly generated C-MET⁺/Hoechst⁺/HNK⁻ myogenic cells.

***In Vitro* Expansion, Differentiation, and Maturation of Purified Myogenic Progenitors**

During 31 days of culture we had previously determined the proliferation rate of purified myogenic progenitors derived from two healthy controls and two Pompe patients (van der Wal et al., 2017b). To further determine expansion capacity, we determined the expansion capacities of myogenic progenitors generated from iPSCs from six additional healthy controls and five additional patients with Pompe disease. Proliferation rates were observed for all myogenic progenitor lines that reached 2×10^{16} cells during 43 days of expansion (Figure 2A). Cells could be cryopreserved with 84% recovery (Figure 2B) without affecting differentiation capacity (data not shown), and showed an average cell cycle of 28.9 hr (Figure 2C). After 43 days of expansion, the proliferation rate diminished, the morphology of cells changed, and differentiation capacity decreased (data not shown). This showed that the myogenic progenitors generated with this protocol could be expanded by a maximum of 5×10^{11} -fold.

Previously we had used a 4-day differentiation protocol to demonstrate that the differentiation capacity remained intact during the expansion phase of myogenic progenitors, based on similar fusion indexes (van der Wal et al., 2017b). Here we extended this analysis to demonstrate that all myogenic progenitors derived from eight healthy control and seven Pompe iPSCs retain their capacity to differentiate into multinucleated myotubes during expansion (Figure 2D). The average fusion index ranged between 20% and 97% and showed no expansion-induced differences (Figure S2A). Next, we tested whether maturation to contractile skeletal muscle cells is possible from purified myogenic progenitors. However, extending culture of myogenic progenitor-derived myotubes in conventional differentiation medium (1% ITS-X [insulin-transferrin-selenium-ethanolamine] in DMEM/F12) beyond day 4 of differentiation increased cell detachment and death (data

not shown). Supplementation of the myogenic progenitors' differentiation medium with 0.5%–2% fetal bovine serum increased the overall survival of the culture but also increased the proliferation rate of mononucleated cells, resulting in overgrowth of the cell culture (data not shown). In contrast, supplementation with 1% knockout serum replacement supported further differentiation of myogenic progenitors into skeletal muscle cells for up to 12 days. Longer differentiation resulted in fibers that expressed fast MHC, MHC, titin, and α -actinin; that showed patterns of striation (Figures 2E and S2B); and that contracted spontaneously (Videos S1 and S2). This demonstrated that functional sarcomeres, the strongest evidence of terminal differentiation, were formed.

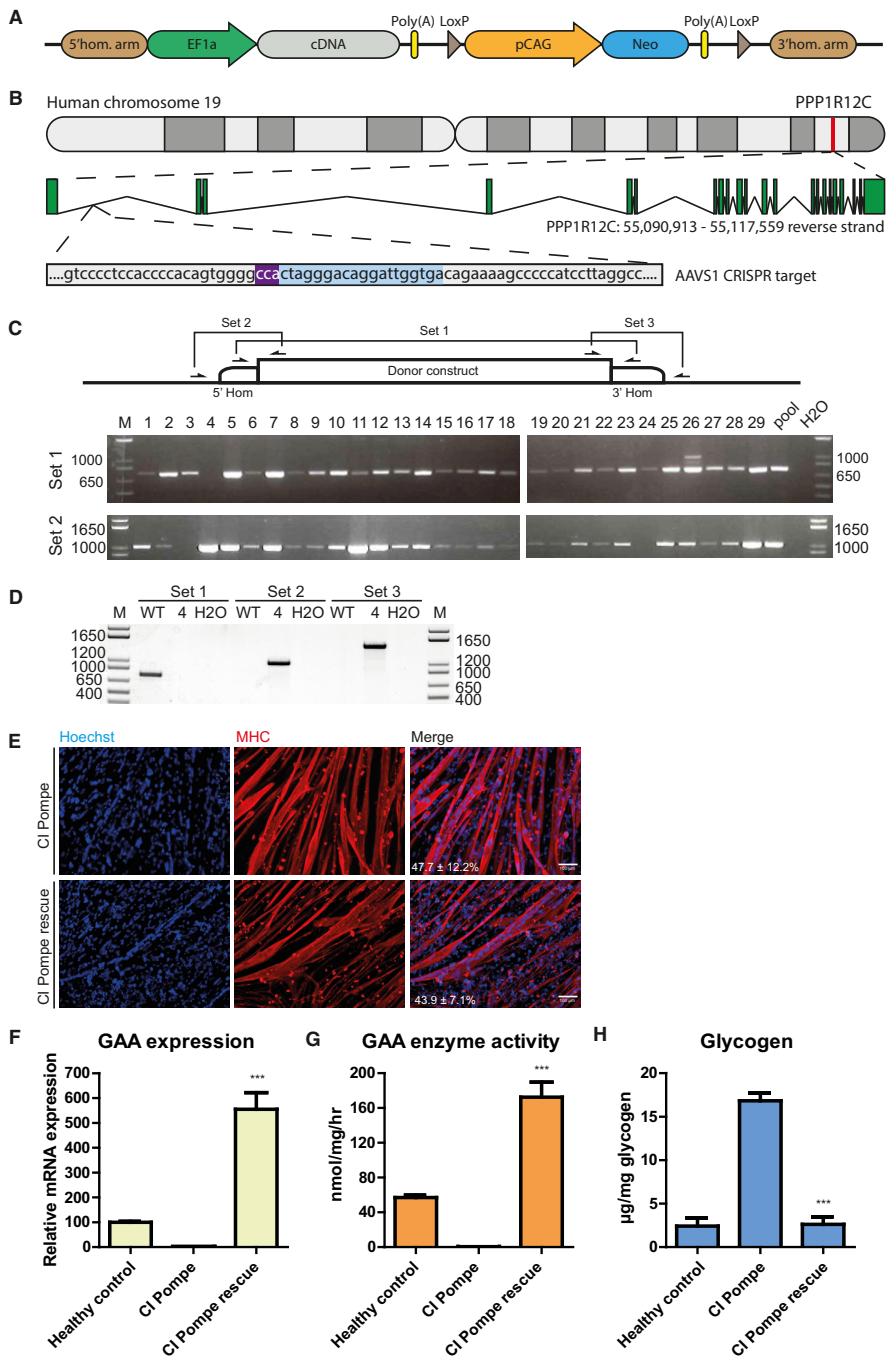
Generation of Gene-Corrected Myogenic Progenitors Using CRISPR/Cas9-Mediated Insertion of a cDNA into a Safe Harbor

Using gene editing, it is possible to perform genetic correction of human disease *in vitro* by placing an extra copy of the wild-type gene into a so-called safe harbor; i.e., a safe location of the genome (Hockemeyer and Jaenisch, 2016). As such a strategy relies on homology-directed DNA repair, which is inefficient, we generated a targeting construct that allows the selection and subsequent removal of the selection marker. The generic donor vector is shown in Figure 3A. As a proof of concept, we chose the PPP1R12C gene in the AAVS1 locus (Figure 3B) (Lombardo et al., 2011). As well as unique restriction sites that enable cloning of the 5' and 3' homology arms, the donor vector contains a ubiquitous EF1 α promoter in front of the cDNA of interest (flanked by unique restriction sites); a poly(A) site; and a *neomycin* selection marker driven by the CAG promoter flanked by loxP sites, which provide the option of removing the selection marker by transient expression of CRE recombinase (Figure 3A).

As proof of principle, we aimed to correct the glycogen accumulation caused by deficiency of lysosomal acid alpha glucosidase (GAA) in skeletal muscle cells of Pompe patients *in vitro*. To this end, we cloned the native GAA cDNA in the donor construct. iPSCs were generated from

Figure 2. *In Vitro* Expansion, Differentiation, and Maturation of Purified Myogenic Progenitor Cells

- (A) Proliferation curves of myogenic progenitors derived from 15 iPSC lines derived from healthy controls or Pompe patients, cultured in proliferation medium. An exponential trend line was plotted and an R^2 was calculated from all data points, which showed similar proliferation rates for all cell lines.
- (B) Recovery of control 1 myogenic progenitors from freezing. Data are means \pm SD from three independent cultures.
- (C) Average cell cycle duration of all cell lines shown in (A). Data are means \pm SD of all cell lines shown in (A).
- (D) After expansion for the number of days indicated on the X axis, skeletal muscle differentiation was induced for 4 days by switching to differentiation medium. The fusion index was quantified after staining for MHC and Hoechst. Individual values of random fields per cell line ($n = 3$ –5 fields per cell line) are plotted as symbols. Mean values of all cell lines per expansion period are indicated as horizontal lines.
- (E) At day 8 of differentiation, myotubes further matured as indicated by staining for fast MHC, MHC, titin, and α -actinin, a striated pattern, and spontaneous contractions (see Videos S1 and S2). Blue, nuclei as stained with Hoechst.



(legend on next page)

a patient with classic infantile (CI) Pompe disease (the most severe phenotype, which is characterized by complete deficiency of GAA enzyme activity), and co-transfected the donor vector containing the *GAA* cDNA with vectors that expressed a guide RNA targeting the AAVS1 locus and a human codon-optimized Cas9 nuclease. After selection with G418, an average of 200 colonies were obtained per 2×10^6 cells, suggesting a targeting frequency of $1 \times 10^{-4}\%$. Twenty-nine colonies were picked and genotyped using two PCR strategies (Figure 3C). With PCR primer set 1, the untargeted allele yields a product of 749 bp, while the targeted allele yields a product that is too large to be amplified under the conditions employed. With primer set 2, insertion of the *GAA* cDNA at the correct location is detected. The results with primer set 2 showed that 27/29 colonies had inserted the *GAA* cDNA at the desired location. With primer set 1, 28/29 colonies showed that the second allele had not been targeted. One colony (clone 4) contained two targeted alleles. iPSCs from clone 4 were expanded, and the correct integration site was further validated at the 3' site using primer set 3 (Figure 3D). iPSCs from clone 4 were expanded, and myogenic progenitors were generated and compared with myogenic progenitors from the original iPSC line before gene editing. Myogenic progenitors from these lines were purified, expanded, and subjected to myotube differentiation. Similar differentiation capacities and fusion indexes were observed before and after gene editing (Figure 3E). RT-qPCR analysis showed the absence of *GAA* mRNA expression in the untargeted Pompe myotubes; this was caused by mRNA decay following a frameshift in both alleles (*GAA* genotype c.525del/c.525del). In the gene-edited myotubes, *GAA* mRNA expression had been restored 5.5-fold over levels in healthy control myotubes (Figure 3F). *GAA* enzyme activity measurements showed complete restoration of *GAA* activity in the gene-edited myotubes to levels

that were ~3-fold higher than those of healthy control myotubes (Figure 3G). Myotubes from the CI Pompe patient showed accumulation of glycogen that was restored in the gene-edited myotubes to the levels of healthy control myotubes (Figure 3H). Altogether, these results demonstrate the feasibility of combining gene editing in iPSCs with the myogenic differentiation protocol to generate gene-corrected skeletal muscle cells.

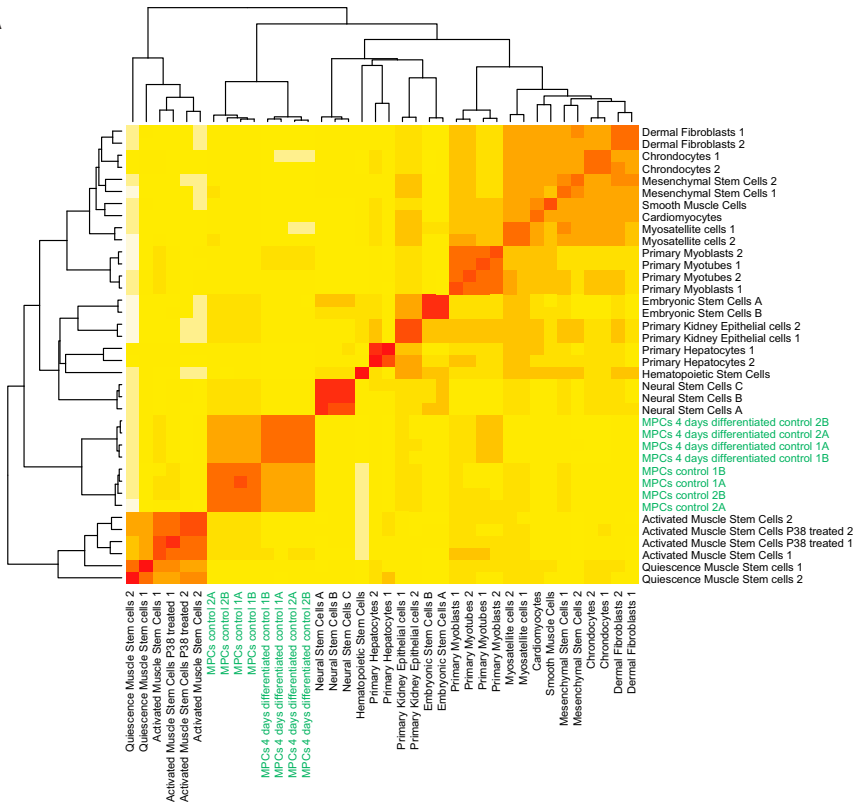
Expression Profiling of iPSC-Derived Myogenic Progenitors

To characterize myogenic progenitors, we used RNA sequencing (RNA-seq) to perform genome-wide mRNA expression analysis. Profiles from purified, expanded (~15 days) iPSC-derived myogenic progenitors from healthy controls were compared with publicly available datasets (see Table S3) on cell types of different lineages, including adult SCs (FACS purified), myoblasts/myosatellite cells (prepared using preplating), neuronal cells, chondrocytes, cardiomyocytes, hepatocytes, embryonic stem cells, smooth-muscle cells, mesenchymal stem cells, and fibroblasts (Figure 4A). The “new Tuxedo” pipeline (Pertea et al., 2016) was used. Spearman correlation analysis showed that profiles of two independent biological replicates of myogenic progenitors from independent individuals clustered together, indicating that these cells contained similar and defined gene expression profiles (Figure 4, myogenic progenitors from the present study are indicated in green). The profiles of myogenic progenitors clustered away from all other cell types, while the profiles of the adult quiescent and activated muscle stem cells showed an early split from all other profiles. A total of 1,852 out of 13,193 genes were differentially expressed between activated muscle stem cells and myogenic progenitors (false discovery rate < 0.01; Table S5). The dissimilarity between quiescent and

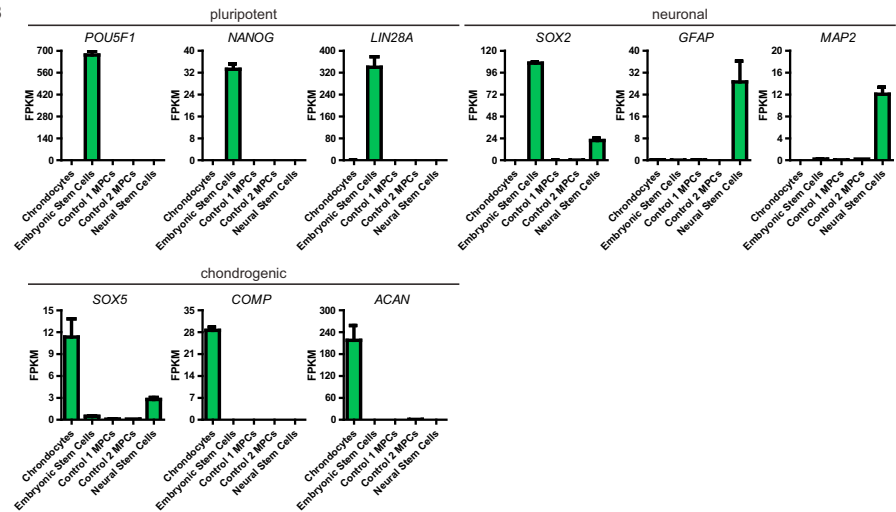
Figure 3. Gene Editing in iPSCs Restores the Pompe Disease Phenotype in Skeletal Muscle Cells *In Vitro*

(A) Generic construct for insertion of a cDNA in a safe harbor following CRISPR/Cas9-mediated targeting.
(B) The construct shown in (A) was tailored to express *GAA* in the AAVS1 locus. After transfection into iPSCs from a classic infantile (CI) Pompe patient, G418 selection was used, and single colonies were picked.
(C) Genotyping was performed using PCR. Primer sets 2 and 3 amplified a product that is only present in correctly targeted clones, while primer set 1 spanned the insertion site to give a product only in the absence of targeting. With primer set 1, 28/29 clones were positive, indicating that most clones also contained an untargeted allele; with primer set 2, 27/29 clones were positive, indicating that most clones showed efficient targeting of at least one allele.
(D and E) (D) One clone (#4) showed targeting of both alleles, which was validated using primer set 3, and was differentiated into myogenic progenitors for further analysis. Myogenic progenitors were generated from healthy controls, a CI Pompe patient (CI Pompe), and the isogenic, gene-corrected, CI Pompe patient (CI Pompe rescue). Myogenic progenitors were purified, expanded, differentiated for 6 days into myotubes, and the fusion index was determined (E).
(F–H) Myogenic progenitors were analyzed for *GAA* mRNA expression at day 4 (F); *GAA* enzyme activity at day 4 (G); and glycogen accumulation at day 6 (H). *GAA* mRNA expression was measured by RT-qPCR using primers spanning exon 1–2. *GAA* enzyme activity was measured using the 4-methylumbelliferone assay. Glycogen accumulation was measured biochemically. To deplete cytoplasmic glycogen, cells were cultured in glucose-free medium for the last 24 hr, as described in Bergsma et al. (2015). For (F, G, and H), data are means \pm SD of two independent (healthy control) or three independent (CI Pompe disease and rescue) cultures. Two-tailed Student's *t* test: ****p* < 0.001.

A



B



(legend on next page)

activated muscle stem cells from myosatellite cells and primary myoblasts can be explained by the fact that the former cells were FACS purified, while the latter cells were obtained using preplating and probably contained contaminating cell types. KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis of genes that were differentially expressed in myogenic progenitors relative to activated muscle stem cells showed enrichment of the AMPK, MAPK, and ErbB signaling pathways in myogenic progenitors (Figure S2C). These pathways have been involved in cell cycle regulation, muscle regeneration, and/or satellite cell function (Charville et al., 2015; Golding et al., 2007; Theret et al., 2017). Overall, this suggests that the myogenic progenitors were dissimilar from the other cell types tested and contained a defined mRNA expression profile.

To assess the purity of the myogenic progenitors, we used the datasets shown in Figure 4A to examine the expression of markers for pluripotent cells (*POU5F1*, *NANOG*, and *LIN28A*), neuronal cells (*SOX2*, *GFAP*, and *MAP2*), and chondrogenic cells (*SOX5*, *COMP*, and *ACAN*). None of these markers were expressed in the purified iPSC-derived myogenic progenitor cultures, suggesting that contaminating cells from the lineages tested were absent (Figure 4B).

In earlier work we showed that, upon expansion, purified iPSC-derived myogenic progenitors express several myogenic markers, including the MyoD protein (van der Wal et al., 2017b). To examine PAX7 protein expression during *in vitro* expansion and differentiation, we used a PAX7 antibody to perform immunofluorescent analysis. Under proliferating conditions, expanded myogenic progenitors (~25 days) from two independent iPSCs expressed PAX7 in a subset of cells (Figure 5A). Although myogenic progenitor cultures contained a stable ~3% of PAX7⁺ cells during the majority of the expansion period, the percentage of Pax7⁺ cells started to decline at day 39 (control 1) or day 28 (control 2) (Figure 5B). After differentiation to myotubes, PAX7⁺ cells remained present in the culture (Figure 5C and data not shown). These results indicate that, during expansion, a subset of iPSC-derived myogenic progenitors continue to express markers of SCs during both proliferation and differentiation.

***In Vivo* Myogenic Potential of Myogenic Progenitors**

To test the capacity of purified and expanded myogenic progenitors to engraft and contribute to muscle regeneration *in vivo*, we performed cell transplantations in tibialis anterior (TA) muscles of NSG immunodeficient recipient mice that had been pre-injured with BaCl₂. Analysis of engraftment was performed 4 weeks after transplantation. Using human-specific epitopes (Lamin A/C, Spectrin, and Dystrophin; for controls, see Figure S3A), we observed that myogenic progenitors that had been expanded for 3 days were able to engraft and participate in the formation of new myofibers (Figure 6A). In addition, myogenic progenitors were engrafted after longer periods of expansion (6 and 11 days), and at different cell concentrations (2.5×10^5 to 1×10^6 , healthy control 1 line) ($n = 6$ mice) (data not shown). Quantification of the number of Spectrin⁺ fibers showed that cell engraftment efficiency was 35–58 fibers/section, with 87–127 Lamin A/C⁺ nuclei/section (Figure 6B, using two independent cell lines: control 1 and control 5). Lamin A/C⁺ nuclei were found within myofibers and in the interstitium. A subset of Lamin A/C⁺ nuclei was found at a satellite cell position (Figure S3B top); however, very few of those were Pax7⁺ (Figure S3B bottom). The location of Lamin A/C⁺ nuclei was as follows: ~45% was found within human Spectrin⁺ myofibers, suggesting that these contributed to myofiber formation (Figure 6C); 25%–36% was found in the interstitium (Figure S3C); the remaining 23%–40% was found within Spectrin[−] myofibers, which may indicate that in those (multinucleated) fibers mouse nuclei were dominant. These results demonstrate the engraftment potential and regenerative capacities of expanded myogenic progenitors and their participation in muscle regeneration *in vivo*.

DISCUSSION

In this study, we have characterized FACS-purified myogenic progenitors for their applicability *in vitro* and *in vivo*, and provide a detailed protocol to generate these cells. The principle of the procedure and its possible applications are shown in Figure 7. We showed that it is possible to reproducibly generate myogenic progenitors from 15

Figure 4. Molecular Profiling and Purity of Myogenic Progenitors

(A) Purified myogenic progenitors have a myogenic gene expression signature. Heatmap showing a comparison of genome-wide mRNA expression (as measured by RNA-seq) from myogenic progenitors and publicly available datasets. Purified myogenic progenitors from two healthy control iPSCs were included: cells were either expanded for ~15 days in proliferation medium or differentiated for 4 days. Published datasets are listed in Table S3. Datasets were analyzed using the “new Tuxedo” pipeline as described in Pertea et al. (2016). Spearman correlations are shown. Datasets generated in this study are indicated in green.

(B) Purified myogenic progenitors do not express pluripotency markers (*POU5F1*, *LIN28A*, and *NANOG*), neuronal markers (*SOX2*, *GFAP*, and *MAP2*), or chondrogenic markers (*SOX5*, *COMP*, and *ACAN*). Data were extracted from (A). Data are means \pm SD of two independent cultures per cell line.

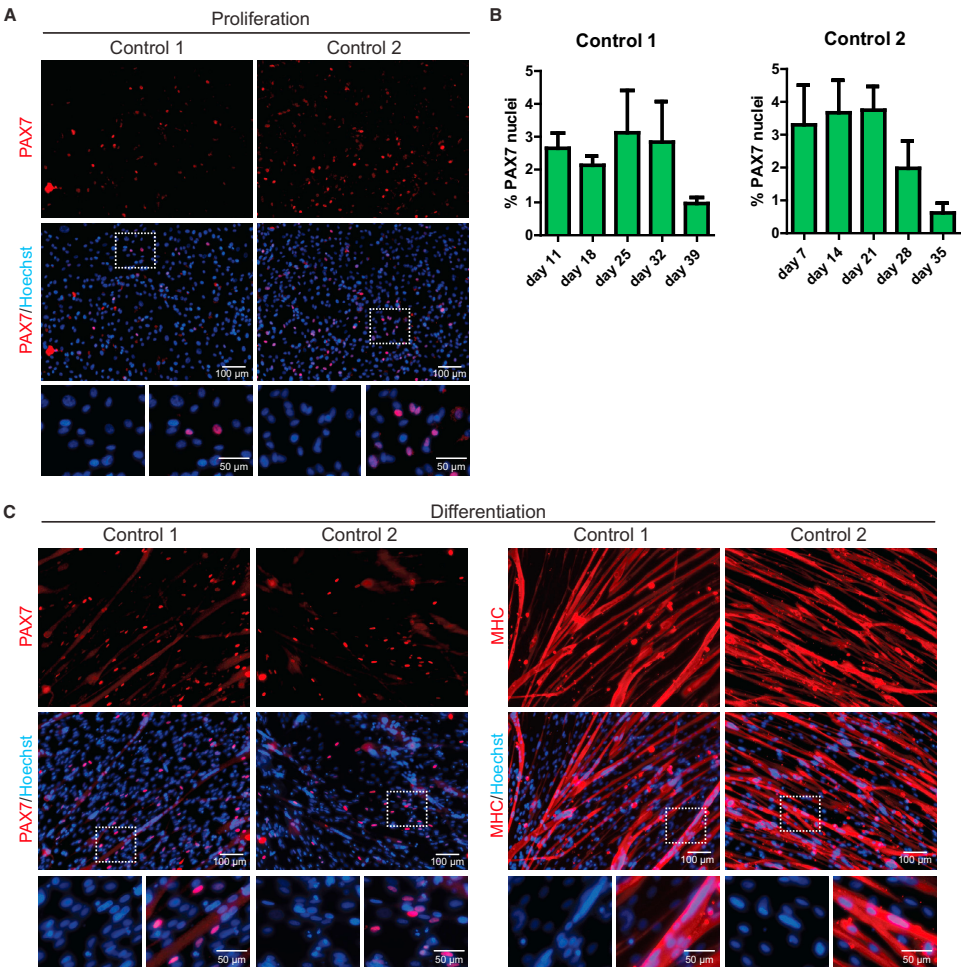


Figure 5. PAX7 Expression during *In Vitro* Proliferation and Differentiation of Purified Myogenic Progenitors

(A) Purified myogenic progenitors from two healthy control iPSCs were expanded for ~25 days in proliferation medium and stained with a PAX7 antibody and Hoechst to stain nuclei.

(B) Quantification of PAX7⁺ cells during expansion of myogenic progenitors from the two healthy control iPSCs shown in (A). Data are means \pm SD of $n = 5$ fields per point.

(C) Myogenic progenitors were differentiated for 6 days to myotubes. Immunofluorescent analysis was performed using a PAX7 antibody (in red) or an MHC antibody (in red) to monitor myotube formation, as indicated. Nuclei were stained with Hoechst (blue).

iPSC lines that were derived from different donors. As we have shown previously, 4×10^4 sorted myogenic progenitors could be expanded to as much as 1×10^{12} cells within

31 days without losing differentiation capacity (van der Wal et al., 2017b). Our current data show that the period during which myogenic progenitors can be expanded can

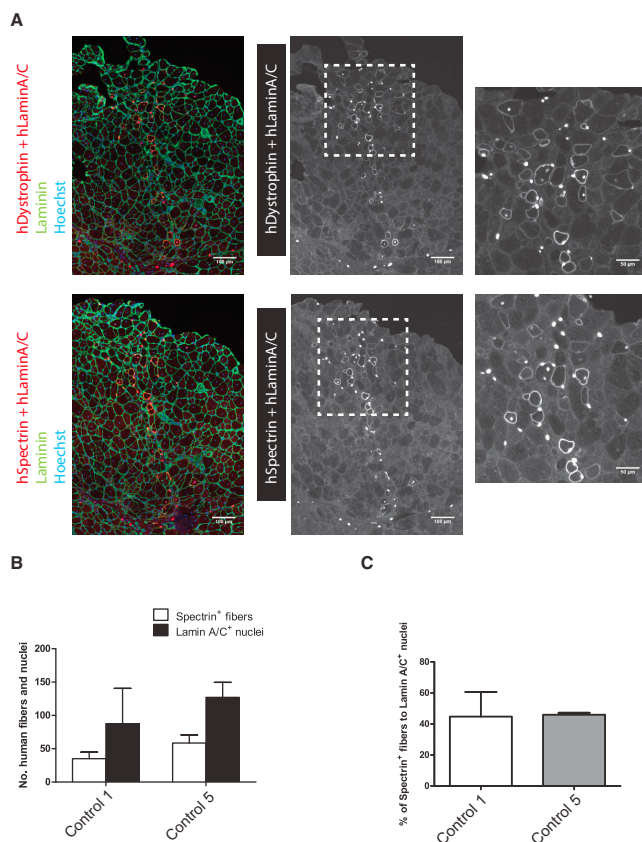


Figure 6. *In Vivo* Myogenic Potential of Purified Myogenic Progenitors Following Engraftment in Immunodeficient Mice

(A) Twenty-four hours before transplantation, the TA of NSG mice was injured using BaCl₂. Myogenic progenitors were administered using intramuscular injection of 5×10^5 cells. Four weeks after transplantation, engraftment was determined by immunohistochemistry of human-specific Lamin A/C and Dystrophin or Spectrin (white or red) and multi-species Laminin (green) on consecutive cross sections.

(B and C) (B) Quantification of Spectrin⁺ muscle fibers and Lamin A/C⁺ nuclei and (C) the percentage of Spectrin⁺ fibers relative to the total number of Lamin A/C⁺ nuclei per section of each biological replicate. Data in (B) and (C) are means \pm SD ($n = 2$ TAs transplanted per line used. Each replicate was transplanted in different mice). All sections were counterstained with Hoechst (blue). Scale bars represent 100 μ m, and 50 μ m on insets.

be extended to up to 43 days. After ~ 50 days of expansion, changes in morphology and proliferation rate suggested the initiation of a senescent phenotype. It is therefore likely that, during the expansion, myogenic progenitors slowly progress to a myoblast-like phenotype, a cell type that is known to undergo replicative senescence during passaging (Bigot et al., 2008). After 43 days of culture, myogenic progenitors had expanded as much as 5×10^{11} -fold (the maximum value obtained), allowing the generation of at least 2×10^{16} cells, which should be sufficient for subsequent analyses, including high-throughput screenings and engraftment studies.

We generated a generic donor construct that can be used for precise and highly efficient gene correction. The selection of positive clones is facilitated by its inclusion of a se-

lection marker. The option of removing the selection marker using transient CRE recombinase expression may be useful in future *in vitro* and *in vivo* applications. A prerequisite for the strategy of inserting a wild-type copy of a cDNA of interest is that overexpression of the transgene product should not be harmful; if it is, the choice of promoter that drives the transgene should be optimized. Overexpression of a transgene is expected to benefit the development of cell-based therapeutic strategies. In the case of skeletal muscle, which consists of multinucleated cells, it can be envisioned that overexpression in a subset of engrafted myonuclei that become part of the syncytium of the affected myofibers would cross-correct part of the myofiber. If the transgene product is secreted, as GAA is secreted in Pompe disease, overexpression potentially

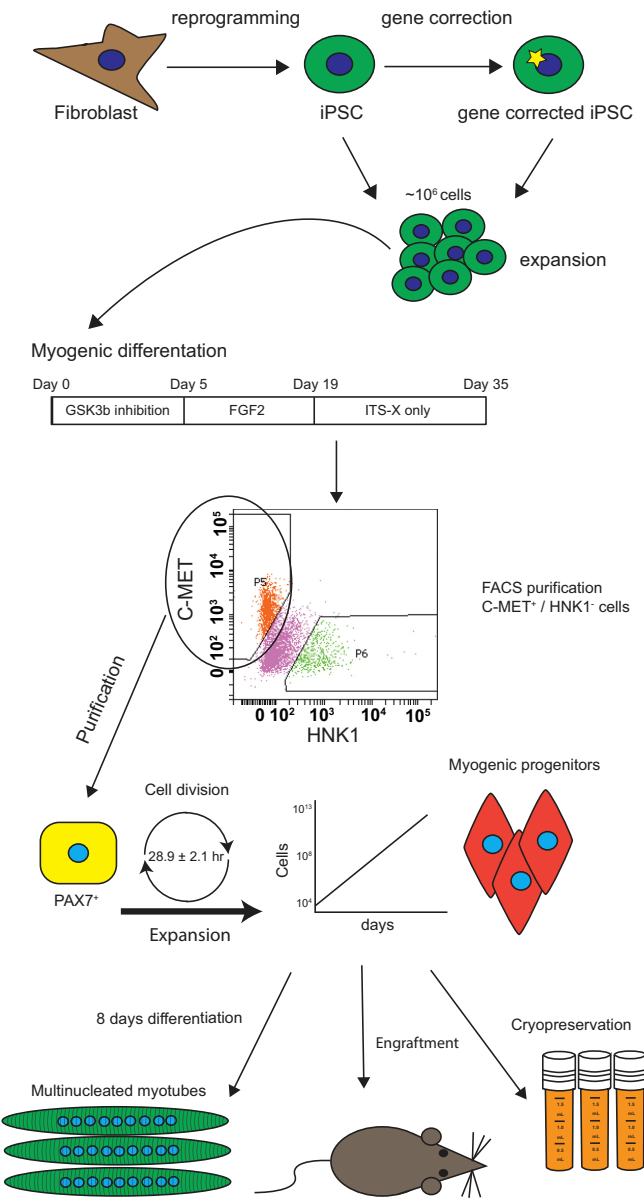


Figure 7. Cartoon Highlighting the Applications of Myogenic Progenitors Described Here

Human iPSCs derived from healthy controls or patients are used as starting cells. Gene correction is applied to iPSCs using CRISPR/Cas9-mediated insertion of a cDNA into a safe harbor. Original or gene-corrected iPSCs are differentiated into the myogenic lineage using a 35-day transgene-free protocol. Myogenic progenitors are purified using a 1-step FACS procedure, and are then expanded (up to 5×10^{11} -fold) and cryopreserved. During expansion, purified myogenic progenitors are differentiated *in vitro* into myotubes with high fusion index, and show striation and spontaneous contraction upon *in vitro* maturation. Upon engraftment in immunodeficient mice, purified and expanded myogenic progenitors form human mononuclear cells and contribute to myofiber formation *in vivo*.

results in cross-correction of neighboring myofibers (Zaretzky et al., 1997).

Both iPSC-derived myogenic progenitors and SCs express PAX7 during *in vitro* proliferation and differentiation, and contribute to myofiber formation after intramuscular engraftment in immunodeficient mice *in vivo*. We showed that, upon optimization of the differentiation process using defined medium conditions, the cells not only expressed fast MHC, α -actinin, and titin but also formed functional sarcomeres, thereby allowing spontaneous contractions. These data revealed enhanced maturation compared with that found in our previous report (van der Wal et al., 2017b) and showed that it is possible to generate mature myofibers from purified iPSC-derived cultures. The formation of myogenic progenitors from Pompe patient-derived iPSCs was not hampered by the underlying disorder, and we expect that this approach can be used to model other disorders that affect muscle cells. It remains to be determined whether, as Chal et al. (2015) report, purified myogenic progenitors have the capacity to differentiate into millimeter-long skeletal muscle cells with PAX7⁺ cells embedded between the sarcolemma and the basal lamina.

It is essential for the development of stem cell-based therapies that the transplantable cell preparations are highly pure and well characterized before transplantation in human patients can be considered. The development of techniques for the expansion and manipulation of pure myogenic progenitor populations *ex vivo* is therefore critical to the further development of this field. In this paper we have provided evidence for the successful engraftment of myogenic progenitors in pre-injured muscles of mice over a period of 4 weeks post-transplantation. The efficiencies of engraftment of mononuclear cells and their contribution to myofibers were comparable with those recently obtained using inducible PAX7 overexpression (Magli et al., 2017). Transplanted myogenic progenitors demonstrated their ability to regenerate injured muscle, as was shown by the detection of centrally located lamin A/C⁺ human nuclei, a characteristic perceived only in fusion-competent myoblasts.

Future studies should identify the stem cell properties of transplanted human myogenic progenitors that allow transplanted donor cells to make a long-term contribution to muscle regeneration. This would provide researchers with novel tools that would help them make progress in the development of muscle stem cell therapies for treating muscle-wasting diseases.

EXPERIMENTAL PROCEDURES

Ethics Approval and Consent to Participate

The Institutional Review Board approved the study protocol, and all patients provided written informed consent. All animal experiments were approved by the animal experiments committee DEC-Consult.

Culture of Myogenic Progenitors

Myogenic progenitors were expanded in myogenic progenitor proliferation medium consisting of DMEM high glucose (Gibco, Waltham, MA) supplemented with 10% fetal bovine serum (HyClone, Thermo Scientific, Waltham, MA), 1% penicillin-streptomycin-glutamine (P/S/G) (Gibco, Waltham, MA), and 100 ng/mL FGF2 (Peprotech, Rocky Hill, NJ) on extracellular matrix-coated dishes (1:200 diluted, Sigma-Aldrich, E6909). For splitting, myogenic progenitors were detached with TrypLe reagent (Gibco, Waltham, MA) diluted 2× with PBS (Gibco, Waltham, MA). For cryopreservation, myogenic progenitors were detached as described above, and after centrifugation the cell pellet was resuspended in myogenic progenitor proliferation medium supplemented with 10% DMSO. Standard cell culture techniques were used for the freeze and thaw procedure.

RNA Isolation and RNA-Seq

Myogenic progenitors were expanded for ~15 days and harvested either in proliferation conditions or after 4 days of differentiation as described previously (van der Wal et al., 2017b). RNA was extracted using the RNeasy minikit with DNase treatment (QIAGEN, Germantown, MD). Sequencing libraries were prepared using TruSeq Stranded mRNA Library Prep Kit (Illumina, San Diego, CA) according to the manufacturer's instructions. Libraries were sequenced on a HiSeq2500 sequencer (Illumina, San Diego, CA) in rapid-run mode according to the manufacturer's instructions. Reads 50 bp in length were generated. The RNA-seq datasets listed in Table S3 were downloaded and aligned with the datasets generated in this study using the new Tuxedo pipeline as described by Pertea et al. (2016). Shortly, RNA-seq data were aligned using Hisat2 (version 2.1.0) to hg38 from University of California, Santa Cruz. The alignments were converted to BAM format using Samtools (version 1.3.1). Then, StringTie was used to quantify transcript expression levels according to the reference transcripts. For KEGG analysis, gene expression was quantified using StringTie with the -e option.

Maturation of Myogenic Progenitors into Skeletal Muscle Cells

When myogenic progenitors reached 90% confluence, cells were switched to myogenic progenitor differentiation medium containing DMEM high glucose supplemented with 1% P/S/G, 1× ITS-X, and 1% knockout serum replacement (all Gibco). Medium was not refreshed during differentiation and cells were harvested at 6 days, 8 days, or 12 days.

Construction of Donor Vector

To generate the generic donor vector for the overexpression of the gene of interest via CRISPR/Cas9-mediated knockin, we used the pCAGEN and pEF-GFP vectors (available on addgene: #11160 and #11154) as starting points. The neomycin selection cassette was introduced via PCR amplification into the pCAGEN vector, destroying the EcoRI and NotI sites. KpnI and ClaI sites were then added to the SalI site, and loxP and SfiI sites were added to the HindIII site. In the pEF-GFP vector, a KpnI site was added to the SalI site and loxP and ClaI sites were added to the HindIII site. Vectors were

combined using the KpnI and ClaI sites. The GAA cDNA was introduced via PCR amplification with EcoRI and NotI fragments. The 5' homology arm of 700 bp was added via PCR amplification with KpnI fragments, and the 3' homology arm of 972 bp was added via PCR amplification with HindIII fragments. All constructs were validated by sequencing. Cloning details are available on request.

Glycogen Assay

Myogenic progenitors were differentiated for 6 days in myogenic progenitor differentiation medium. On day 5 of differentiation, skeletal muscle cells were starved with differentiation medium without glucose (DMEM no glucose, Gibco). On day 6, skeletal muscle cells were detached with a scraper and the pellet was lysed with ice-cold protein lysis buffer (see [Supplemental Experimental Procedures](#)). Glycogen was measured as described in [Bergsma et al. \(2015\)](#).

Gene Editing of iPSCs

To select optimal target sites for the AAVS1 locus, single guide RNA (sgRNA) sequences were designed using the CRISPRscan program ([Moreno-Mateos et al., 2015](#)). The sgRNA CCACTAGGGA-CAGGATTGGTGA was expressed from a TOPO vector containing the U6 promoter (addgene: 41824). Confluent iPSCs on feeders were pretreated 4 hr before nucleofection with 10 μ M Rock inhibitor (Y-27632 dihydrochloride, Ascent Scientific, Asc-129). Single cells were generated from iPSC colonies by incubating with Accutase (Thermo Scientific, Waltham, MA), and 2×10^5 cells were nucleofected with 4 μ g of pCAG-hCAS9-GFP (addgene: 44719), 3 μ g of TOPO-sgRNA, and 2 μ g of donor vector using Amaxa Human Stem Cell Nucleofector Kit2 (VPH-5022, Lonza, Walkersville, MD) with program B-016. After nucleofection, cells were recovered in iPSC-conditioned medium (iPSC medium incubated for 24 hr on feeder cells) supplemented with 20 ng/mL FGF2 (Peprotech, Rocky Hill, NJ) and 10 μ M rock inhibitor. iPSCs were selected after 48 hr of nucleofection with 100 μ g/mL G-418 (Invivogen, San Diego, CA). Approximately 14 days after selection of the iPSCs, single colonies were picked and genotyped using primers from [Table S4](#).

Transplantation into NSG Mice

NSG (Jackson Laboratories) mice aged 2–6 months were used for transplantation studies. Mice (independently of gender) were anesthetized with isoflurane in oxygen from a vaporizer. Regeneration of skeletal muscle was induced by chemical injury. The endogenous skeletal muscle fibers of the mice were injured by injection with 50 μ L of 1.2% barium chloride (BaCl₂) into the TA muscle. Twenty-four hours later, 20 μ L of 5×10^5 dissociated cells were injected into the TA muscle in duplicates (one female and one male). Transplanted cells in this study were expanded for 3 days. PBS-injected TAs were used as negative control for cell transplantations. Mice were sacrificed 4 weeks after cell transplantation, and their TA muscles harvested. TA muscles were frozen in isopentane cooled in liquid nitrogen and stored at -80°C until analysis; 10 μ m cryosections were obtained at intervals throughout the entire muscle and were either stored at -80°C for further immunostaining or were used immediately for PAX7 staining.

Immunofluorescent Stainings

Muscle cryosections were fixed in ice-cold acetone for 5 min, followed by a permeabilization step with 0.3% Triton X-100 in PBS for 20 min. Samples were incubated with a blocking solution of 20% goat serum (DAKO, Santa Clara, CA) and 2% BSA (Sigma-Aldrich, Irvine, UK) in 0.1% Tween in PBS for 1 hr. Sections were incubated with primary antibodies mouse anti-human Lamin A/C (1:100, VP-L550, Vector Laboratories, Burlingame, CA) plus mouse anti-human Spectrin (1:100, SPEC1-CE, Leica, Wetzlar, Germany) or mouse anti-human Dystrophin (1:150, MABT827, Millipore) co-stained with rabbit anti-Laminin (1:100, L9393, Sigma-Aldrich, Irvine, UK) overnight at 4°C . Tissue sections were stained with secondary antibodies goat anti-rabbit (Alexa Fluor 488, 1:500, A-21141, Life Technologies, Carlsbad, CA) and horse anti-mouse biotin (1:250, BA-2000, Vector Laboratories, Burlingame, CA) for 1 hr at room temperature, followed by incubation with Streptavidin 594 (1:500, S-32356, Invitrogen, Carlsbad, CA) for 30 min. Freshly cut tissue was used for PAX7 stainings. Sections were fixed in 4% paraformaldehyde for 5 min and blocked with 20% goat serum and 2% BSA in 0.5% Triton X-100 in PBS for 1 hr, then incubated with mouse anti-PAX7 (1/20, DSHB), Lamin A/C, and Laminin in blocking solution for 2 hr at room temperature. Goat anti-mouse IgG1 Cy3 (1:500, 115-165-205, Jackson ImmunoResearch), goat anti-mouse IgG2b Alexa Fluor 488 (1:500, A-21141, Thermo Fisher), and goat anti-rabbit Alexa Fluor 647 (1:500, A21245, Invitrogen, Carlsbad, CA) were used in 0.1% PBST for 1 hr. All sections were incubated with Hoechst nuclear staining (1:15,000 Invitrogen, Carlsbad, CA) for 10 min and mounted with Mowiol medium (Sigma-Aldrich, Irvine, UK). Images were obtained using confocal microscopy (Zeiss LSM 700).

Statistical Analysis

Data represent mean \pm SD, and p values refer to two-sided t tests. Multiple groups were tested with one-way ANOVA followed by individual two-sided t tests. A p value of <0.05 was considered to be significant. Data showed normal variance and no samples were excluded from the analysis. Images for quantification were randomly selected.

ACCESSION NUMBERS

RNA-seq fastq files and data are accessible at GEO under GEO: GSE111163.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, five tables, and two videos and can be found with this article online at <https://doi.org/10.1016/j.stemcr.2018.04.002>.

AUTHOR CONTRIBUTIONS

Myogenic protocol, E.v.d.W., S.i.G., T.J.M.v.G., and W.W.M.P.P.; Engraftment, P.H.-H., T.J.M.v.G., G.J.S., and W.W.M.P.P.; Expression analysis, R.W., T.H.C., W.F.J.v.L., E.v.d.W., and W.W.M.P.P.; Gene edit and pathology, M.B., E.v.d.W., S.i.G., and W.W.M.P.P.; Funding, W.W.M.P.P., A.T.v.d.P., and G.J.S.; Data interpretation,

all authors; Writing, E.v.d.W., P.H.-H., and W.W.M.P.P.; Supervision, G.J.S., T.H.C., and W.W.M.P.P.

ACKNOWLEDGMENTS

We thank Dr. Tiziano Barberi for discussion; Dr. Schambach for the OSKM-lentiviral vector; and Dr. Christian Freund, Prof. Dr. Christine Mummery, Dr. Mehrnaz Ghazvini, and Prof. Dr. Joost Gribnau for providing healthy control fibroblasts and healthy control iPSC lines. This work was funded by the Prinses Beatrix Spierfonds/Stichting Spieren voor Spieren (grant WOR13-21), Tex Net, and the Croucher Innovation Award (to T.H.C.). A.T.v.d.P. has provided consulting services for various industries in the field of Pompe disease under an agreement between these industries and Erasmus MC, Rotterdam, the Netherlands. All the other authors declare no conflict of interest.

Received: September 24, 2017

Revised: March 31, 2018

Accepted: April 3, 2018

Published: May 3, 2018

REFERENCES

- Baghdadi, M.B., and Tajbakhsh, S. (2017). Regulation and phylogeny of skeletal muscle regeneration. *Dev. Biol.* **433**, 200–209.
- Bergsma, A.J., Kroos, M., Hoogveen-Westerveld, M., Halley, D., van der Ploeg, A.T., and Pijnappel, W.W. (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in Pompe disease using a generic approach. *Hum. Mutat.* **36**, 57–68.
- Bigot, A., Jacquemin, V., Debacq-Chainiaux, F., Butler-Browne, G.S., Toussaint, O., Furling, D., and Mouly, V. (2008). Replicative aging down-regulates the myogenic regulatory factors in human myoblasts. *Biol. Cell* **100**, 189–199.
- Borchin, B., Chen, J., and Barberi, T. (2013). Derivation and FACS-mediated purification of PAX3+/PAX7+ skeletal muscle precursors from human pluripotent stem cells. *Stem Cell Reports* **1**, 620–631.
- Bursac, N., Juhas, M., and Rando, T.A. (2015). Synergizing engineering and biology to treat and model skeletal muscle injury and disease. *Annu. Rev. Biomed. Eng.* **17**, 217–242.
- Caron, L., Kher, D., Lee, K.L., McKernan, R., Dumevska, B., Hidalgo, A., Li, J., Yang, H., Main, H., Ferri, G., et al. (2016). A human pluripotent stem cell model of facioscapulohumeral muscular dystrophy-affected skeletal muscles. *Stem Cells Transl. Med.* **5**, 1145–1161.
- Chal, J., Al Tanoury, Z., Hestin, M., Gobert, B., Aivio, S., Hick, A., Cherrier, T., Nesmith, A.P., Parker, K.K., and Pourquie, O. (2016). Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat. Protoc.* **11**, 1833–1850.
- Chal, J., Oginuma, M., Al Tanoury, Z., Gobert, B., Sumara, O., Hick, A., Bousson, F., Zidouni, Y., Mursch, C., Moncuquet, P., et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat. Biotechnol.* **33**, 962–969.
- Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex vivo expansion and in vivo self-renewal of human muscle stem cells. *Stem Cell Reports* **5**, 621–632.
- Choi, I.Y., Lim, H., Estrellas, K., Mula, J., Cohen, T.V., Zhang, Y., Donnelly, C.J., Richard, J.P., Kim, Y.J., Kim, H., et al. (2016). Concordant but varied phenotypes among Duchenne muscular dystrophy patient-specific myoblasts derived using a human iPSC-based model. *Cell Rep.* **15**, 2301–2312.
- Darabi, R., Arpke, R.W., Irion, S., Dimos, J.T., Grskovic, M., Kyba, M., and Perlingeiro, R.C. (2012). Human ES- and iPSC-derived myogenic progenitors restore DYSTROPHIN and improve contractility upon transplantation in dystrophic mice. *Cell Stem Cell* **10**, 610–619.
- Dumont, N.A., Bentzinger, C.F., Sincennes, M.C., and Rudnicki, M.A. (2015). Satellite cells and skeletal muscle regeneration. *Compr. Physiol.* **5**, 1027–1059.
- Golding, J.P., Calderbank, E., Partridge, T.A., and Beauchamp, J.R. (2007). Skeletal muscle stem cells express anti-apoptotic ErbB receptors during activation from quiescence. *Exp. Cell Res.* **313**, 341–356.
- Hockemeyer, D., and Jaenisch, R. (2016). Induced pluripotent stem cells meet genome editing. *Cell Stem Cell* **18**, 573–586.
- Kaplan, J.C., and Hamroun, D. (2015). The 2016 version of the gene table of monogenic neuromuscular disorders (nuclear genome). *Neuromuscul. Disord.* **25**, 991–1020.
- Kim, J., Magli, A., Chan, S.S.K., Oliveira, V.K.P., Wu, J., Darabi, R., Kyba, M., and Perlingeiro, R.C.R. (2017). Expansion and purification are critical for the therapeutic application of pluripotent stem cell-derived myogenic progenitors. *Stem Cell Reports* **9**, 12–22.
- Lepper, C., Partridge, T.A., and Fan, C.M. (2011). An absolute requirement for Pax7-positive satellite cells in acute injury-induced skeletal muscle regeneration. *Development* **138**, 3639–3646.
- Lombardo, A., Cesana, D., Genovese, P., Di Stefano, B., Provati, E., Colombo, D.F., Neri, M., Magnani, Z., Cantore, A., Lo Riso, P., et al. (2011). Site-specific integration and tailoring of cassette design for sustainable gene transfer. *Nat. Methods* **8**, 861–869.
- Magli, A., Incitti, T., Kiley, J., Swanson, S.A., Darabi, R., Rinaldi, F., Selvaraj, S., Yamamoto, A., Tolar, J., Yuan, C., et al. (2017). PAX7 targets, CD54, integrin alpha9beta1, and SDC2, allow isolation of human ESC/iPSC-derived myogenic progenitors. *Cell Rep.* **19**, 2867–2877.
- Mauro, A. (1961). Satellite cell of skeletal muscle fibers. *J. Biophys. Biochem. Cytol.* **9**, 493–495.
- Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods* **12**, 982–988.
- Murphy, M.M., Lawson, J.A., Mathew, S.J., Hutcheson, D.A., and Kardon, G. (2011). Satellite cells, connective tissue fibroblasts and their interactions are crucial for muscle regeneration. *Development* **138**, 3625–3637.
- Pertea, M., Kim, D., Pertea, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **11**, 1650–1667.

Stem Cell Reports

- Sambasivan, R., Yao, R., Kissenpfennig, A., Van Wittenberghe, L., Paldi, A., Gayraud-Morel, B., Guenou, H., Malissen, B., Tajbakhsh, S., and Galy, A. (2011). Pax7-expressing satellite cells are indispensable for adult skeletal muscle regeneration. *Development* 138, 3647–3656.
- Shelton, M., Kocharyan, A., Liu, J., Skerjanc, I.S., and Stanford, W.L. (2016). Robust generation and expansion of skeletal muscle progenitors and myocytes from human pluripotent stem cells. *Methods* 101, 73–84.
- Shelton, M., Metz, J., Liu, J., Carpenedo, R.L., Demers, S.P., Stanford, W.L., and Skerjanc, I.S. (2014). Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells. *Stem Cell Reports* 3, 516–529.
- Soldner, F., Laganieri, J., Cheng, A.W., Hockemeyer, D., Gao, Q., Alagappan, R., Khurana, V., Golbe, L.L., Myers, R.H., Lindquist, S., et al. (2011). Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell* 146, 318–331.
- Swartz, E.W., Baek, J., Pribadi, M., Wojta, K.J., Almeida, S., Karydas, A., Gao, F.B., Miller, B.L., and Coppola, G. (2016). A novel protocol for directed differentiation of C9orf72-associated human induced pluripotent stem cells into contractile skeletal myotubes. *Stem Cells Transl. Med.* 5, 1461–1472.
- Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. *Nat. Rev. Mol. Cell Biol.* 17, 183–193.
- Theret, M., Gsaier, L., Schaffer, B., Juban, G., Ben Larbi, S., Weiss-Gayet, M., Bultot, L., Collodet, C., Foretz, M., Desplanches, D., et al. (2017). AMPKalpha1-LDH pathway regulates muscle stem cell self-renewal by controlling metabolic homeostasis. *EMBO J.* 36, 1946–1962.
- van der Ploeg, A.T., and Reuser, A.J. (2008). Pompe's disease. *Lancet* 372, 1342–1353.
- van der Wal, E., Bergsma, A.J., Pijnenburg, J.M., van der Ploeg, A.T., and Pijnappel, W. (2017a). Antisense oligonucleotides promote exon inclusion and correct the common c.-32-13T>G GAA splicing variant in Pompe disease. *Mol. Ther. Nucleic Acids* 7, 90–100.
- van der Wal, E., Bergsma, A.J., van Gestel, T.J.M., In 't Groen, S.L.M., Zaehres, H., Arauzo-Bravo, M.J., Scholer, H.R., van der Ploeg, A.T., and Pijnappel, W. (2017b). GAA deficiency in Pompe disease is alleviated by exon inclusion in iPSC-derived skeletal muscle cells. *Mol. Ther. Nucleic Acids* 7, 101–115.
- Xu, C., Tabebordbar, M., Iovino, S., Ciarlo, C., Liu, J., Castiglioni, A., Price, E., Liu, M., Barton, E.R., Kahn, C.R., et al. (2013). A zebrafish embryo culture system defines factors that promote vertebrate myogenesis across species. *Cell* 155, 909–921.
- Zaretzky, J.Z., Candotti, F., Boerkoel, C., Adams, E.M., Yewdell, J.W., Blaese, R.M., and Plotz, P.H. (1997). Retroviral transfer of acid alpha-glucosidase cDNA to enzyme-deficient myoblasts results in phenotypic spread of the genotypic correction by both secretion and fusion. *Hum. Gene Ther.* 8, 1555–1563.

Supplemental Information

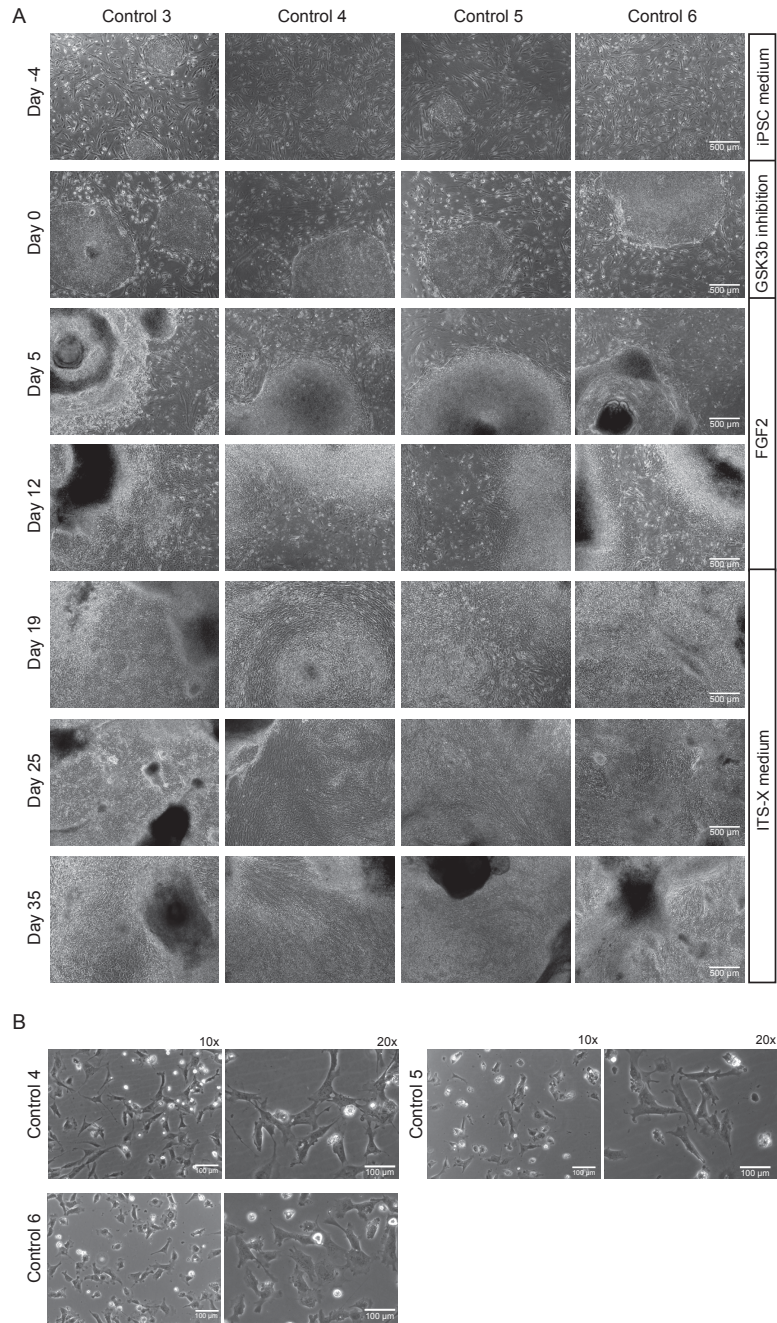


Figure S1

Supplemental Information

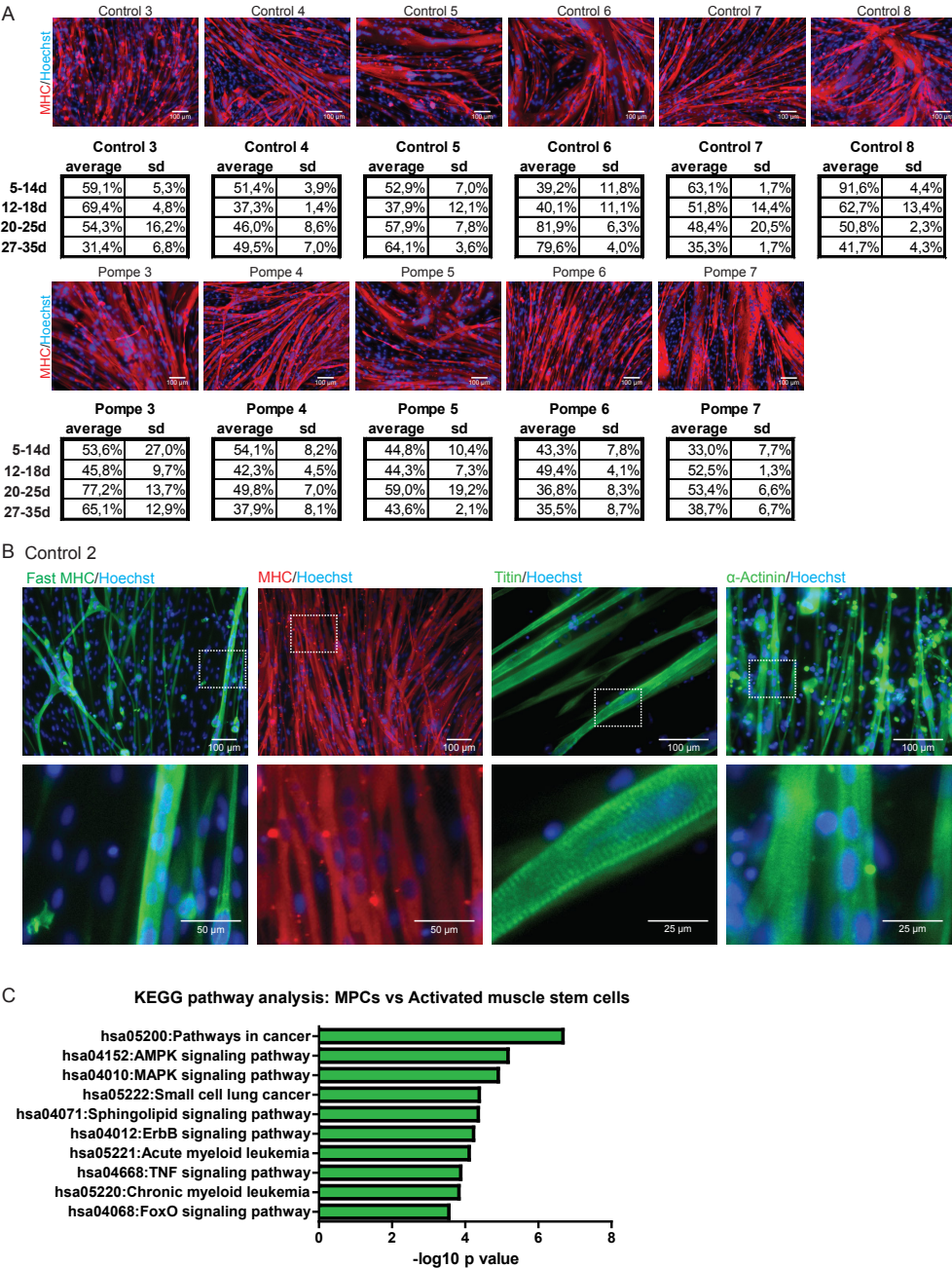
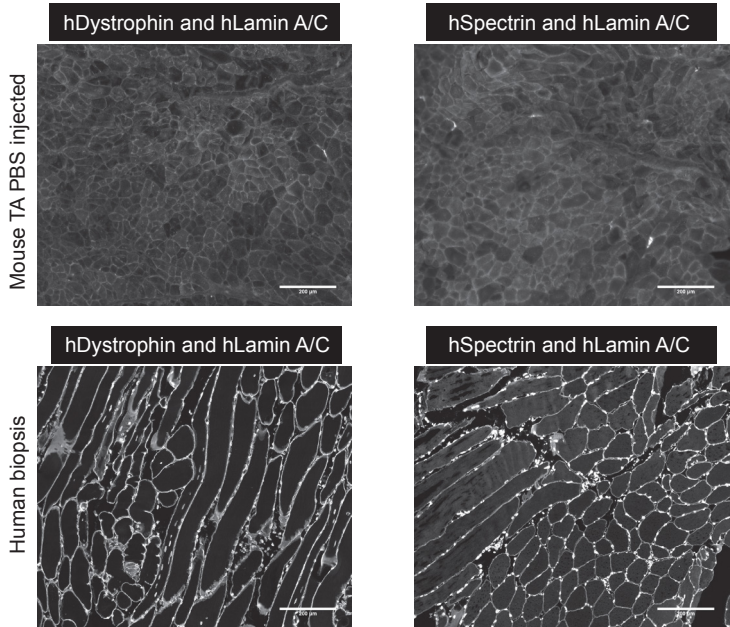


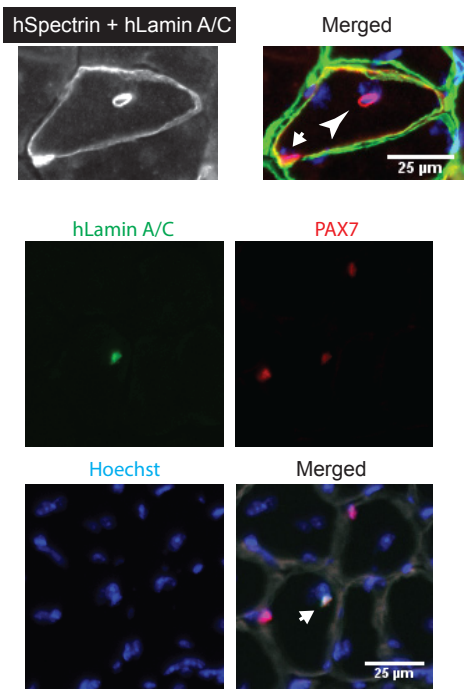
Figure S2

Supplemental Information

A



B



C

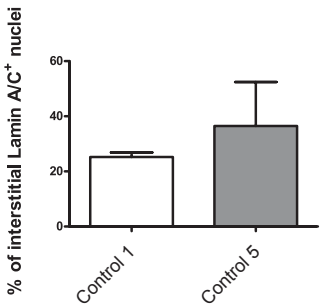


Figure S3

Supplemental Information

SUPPLEMENTARY FIGURE LEGENDS

Figure S1 (related to Figure 1): Cell morphologies during differentiation of iPSCs into the myogenic lineage and after purification of myogenic progenitors. (A) Healthy-control iPSCs 3, 4, 5 and 6 were differentiated using a 35-day protocol consisting of GSK3 β inhibition, FGF2 treatment, and a minimal medium (see Figure 7). Light microscope (4x magnification) images were taken before and during differentiation at the days indicated. Representative images are shown. (B) One day after FACS purification (described in Figure 7), light microscope images were taken from myogenic progenitors generated from healthy controls 4, 5 and 6 at a magnification of 10x and 20x. Representative images are shown.

Figure S2 (related to Figures 2 and 4): Differentiation, maturation of purified myogenic progenitors, and KEGG pathway analysis. (A) Myosin heavy chain (MHC) staining on 4 days differentiated myogenic progenitors from healthy controls 3- 8 and Pompe 3 – 7. For images of healthy controls 1 and 2, and Pompe 1 and 2, see van der Wal et al., 2017. Nuclei were stained with Hoechst. Images are representative for each differentiation. Fusion index during expansion was quantified and data are mean \pm SD of 3 fields per point. (B) Staining of matured fibers from myogenic progenitors of healthy control 2. After 6-8 days of differentiation, cells were stained with Fast MHC, MHC, Titin and α -Actinin antibodies. Nuclei were stained with Hoechst and are shown in blue. (C) KEGG pathway analysis using DAVID of mapped genes comparing myogenic progenitors in proliferation phase (MPCs, this study) versus activated muscle stem cells (Charville et al., 2015). The 10 most significant pathways are shown.

Figure S3 (related to Figure 6): Positive and negative controls for analysis of *in vivo* engraftment and for cell contribution to muscle regeneration *in vivo*. (A) The upper panels show sections from the *tibialis anterior* from immunodeficient mice 4 weeks after injection with PBS only. The lower panel shows sections from a human biopsy of the quadriceps femoris. Sections were analyzed by immunohistochemistry using human specific Lamin A/C, Spectrin and Dystrophin antibodies (white). (B) Upper panel represent two different locations of Lamin A/C⁺ nuclei within the same Spectrin⁺ fiber. The human nuclei on a satellite cell position are indicated with an arrow and the myonuclei with an arrowhead. Lower panel shows a PAX7⁺ (red), Lamin A/C⁺ (green) nucleus in a Laminin⁺ (grey) muscle fiber. (C) Percentage of Lamin A/C⁺ nuclei present at the muscle interstitium per section of each biological replicate. Sections that showed engraftment were used for quantification. Data are mean \pm SD (n= 2 TAs transplanted per line used. Each replicate was transplanted in different mice). All sections were counterstained with Hoechst (blue).

SUPPLEMENTARY TABLES

Table S1. Comparison of transgene-free skeletal-muscle differentiation protocols using GSK3 β inhibition

| | Purification protocol | Fold expansion | Cryopreservation | Duration | Fusion index | Engraftment |
|------------------------|---|----------------|------------------|----------|--------------|-------------------------|
| (Borchin et al., 2013) | FACS | N.R. | N.R. | 35 days | N.R. | N.R. |
| (Xu et al., 2013) | No purification, differentiation analysed in original plate | N.R. | N.R. | 36 days | N.R. | Yes, unpurified culture |
| (Shelton et al., 2014) | No purification, differentiation analysed in original plate | N.R. | N.R. | 50 days | N.R. | N.R. |
| (Chal et al., 2015) | No purification, differentiation analysed in original plate | N.R. | N.R. | 50 days | N.R. | N.R. |
| (Shelton et al., 2016) | No purification, differentiation analysed in original plate | 3x | N.R. | 50 days | N.R. | N.R. |

Supplemental Information

| | | | | | | |
|----------------------------|---|------------------------|------|---------|--------|-------------------------|
| (Choi et al., 2016) | FACS | 10 ⁵ x | Yes | 30 days | 10-15% | Yes, unpurified culture |
| (Caron et al., 2016) | Pre-plating | 1250x | N.R. | 26 days | N.R. | N.R. |
| (Chal et al., 2016) | Pre-plating | N.R. | Yes | 35 days | N.R. | N.R. |
| (Swartz et al., 2016) | No purification, differentiation analysed in original plate | N.R. | Yes | 36 days | N.R. | N.R. |
| (Kim et al., 2017) | No purification, differentiation analysed in original plate | N.R. | N.R. | 50 days | N.R. | Yes, unpurified culture |
| (van der Wal et al., 2017) | FACS | 5 x 10 ⁷ x | Yes | 35 days | 60-80% | N.R. |
| This study | FACS | 5 x 10 ¹¹ x | Yes | 35 days | 20-97% | Yes, purified culture |

N.R.: Not Reported

Table S2. Optimization of CHIR99021 concentration

Control 1

| CHIR99021 | Days | Confluency | PAX7 ⁺ cells |
|-----------|------|------------|-------------------------|
| 3 μ M | 4 | 65% | 15-10% |
| 3 μ M | 5 | 50% | 7-10% |
| 3 μ M | 8 | 40% | 1-2% |
| 3 μ M | 10 | 40% | 0% |
| 4 μ M | 4 | 85% | 30-35% |
| 4 μ M | 5 | 100% | 35-40% |
| 4 μ M | 8 | 75% | 30-35% |
| 4 μ M | 10 | 50% | 5-10% |
| 5 μ M | 4 | 60% | 15-20% |
| 5 μ M | 5 | 70% | 20-25% |
| 5 μ M | 8 | 0% | 0% |
| 5 μ M | 10 | 0% | 0% |

Control 2

| CHIR99021 | Days | Confluency | PAX7 ⁺ cells |
|-----------|------|------------|-------------------------|
| 3 μ M | 4 | 95% | 1-2% |
| 3 μ M | 5 | 100% | 2-3% |
| 3 μ M | 8 | 90% | 3-4% |
| 3 μ M | 10 | 95% | 3-4% |
| 4 μ M | 4 | 95% | 10-15% |
| 4 μ M | 5 | 95% | 9-12% |
| 4 μ M | 8 | 95% | 10-15% |
| 4 μ M | 10 | 85% | 1-2% |
| 5 μ M | 4 | 95% | 3-5% |
| 5 μ M | 5 | 95% | 15-20% |
| 5 μ M | 8 | 50% | 7-10% |
| 5 μ M | 10 | 30% | 0% |

Supplemental Information

Table S3. RNA sequencing datasets used in this study

| Data source | Accession | Abbreviation | Reference |
|-------------|------------------------|--|--------------------------|
| ENA | ERR975347 | Activated Muscle Stem Cell 2 | (Charville et al., 2015) |
| ENA | ERR975349 | Activated Muscle Stem Cell 1 P38 treated 2 | (Charville et al., 2015) |
| ENA | ERR975346 | Activated Muscle Stem Cell 1 | (Charville et al., 2015) |
| ENA | ERR975348 | Activated Muscle Stem Cell 1 P38 treated 1 | (Charville et al., 2015) |
| NCBI | GEO: <i>GSM3024344</i> | MPCs control 1A | This study |
| NCBI | GEO: <i>GSM3024345</i> | MPCs control 1B | This study |
| NCBI | GEO: <i>GSM3024346</i> | MPCs control 2A | This study |
| NCBI | GEO: <i>GSM3024347</i> | MPCs control 2B | This study |
| ENA | ERR975345 | Quiescent Muscle Stem Cell 2 | (Charville et al., 2015) |
| ENA | ERR975344 | Quiescent Muscle Stem Cell 1 | (Charville et al., 2015) |
| NCBI | GEO: <i>GSM3024348</i> | MPCs 4 days differentiated control 1A | This study |
| NCBI | GEO: <i>GSM3024349</i> | MPCs 4 days differentiated control 1B | This study |
| NCBI | GEO: <i>GSM3024350</i> | MPCs 4 days differentiated control 2A | This study |
| NCBI | GEO: <i>GSM3024351</i> | MPCs 4 days differentiated control 2B | This study |
| NCBI | GEO: <i>GSM2452280</i> | Neural stem cell 1 | (McGrath et al., 2017) |
| NCBI | GEO: <i>GSM2452281</i> | Neural stem cell 2 | (McGrath et al., 2017) |
| NCBI | GEO: <i>GSM2452282</i> | Neural stem cell 3 | (McGrath et al., 2017) |
| ENCODE | ENCBS476ENC | Dermal Fibroblast 1 | N/A |
| ENCODE | ENCBS459ENC | Mesenchymal stem cell 2 | N/A |
| ENCODE | ENCSR828TEI | Primary Myotube 1 | N/A |
| ENCODE | ENCBS018ENC | Chondocyte 1 | N/A |
| ENCODE | ENCLB014ZZZ | Cardiomyocyte | N/A |
| ENCODE | ENCBS460ENC | Mesenchymal stem cell 1 | N/A |
| ENCODE | ENCSR000CUI | Myosatellite cell 2 | N/A |
| ENCODE | ENCSR000AAG | Smooth muscle cell | N/A |
| NCBI | SRX689200 | Primary hepatocytes 2 | (Kambara et al., 2014) |
| ENCODE | ENCSR000CUI | Myosatellite cell 1 | N/A |
| ENCODE | ENCBS019ENC | Chondocyte 2 | N/A |
| ENCODE | ENCBS475ENC | Dermal Fibroblast 2 | N/A |
| ENCODE | ENCBS945YXY | Primary Kidney epithelial cell 2 | N/A |
| NCBI | SRX673854 | Primary hepatocytes 1 | (Kambara et al., 2014) |
| ENCODE | ENCBS007YZP | Primary Kidney epithelial cell 1 | N/A |
| ENCODE | ENCSR828TEI | Primary Myotube 2 | N/A |
| ENCODE | ENCSR444WHQ | Primary Myoblast 2 | N/A |
| ENCODE | ENCBS293AAA | Embryonic stem cell 1 | N/A |

Supplemental Information

| | | | |
|--------|-------------|-------------------------|-----|
| ENCODE | ENCBS624XJG | Embryonic stem cell 2 | N/A |
| ENCODE | ENCSR444WHQ | Primary Myoblast 1 | N/A |
| ENCODE | ENCBS485ENC | Hematopoietic stem cell | N/A |

Table S4. Antibodies and primers used in experiments

| Name | Dilution or Sequence 5'-3' | Company | Assay |
|-------------------------------|----------------------------|-------------------------------|-----------|
| Mouse-anti-MF20 | 1:50 | DSHB | IF |
| Rabbit-anti-Myogenin | 1:100 | Santa Cruz (sc-576) | IF |
| Mouse-anti-PAX7 | 1:100 or 1:20 | DSHB | IF or IHC |
| Mouse-anti- α -Actinin | 1:100 | Sigma-Aldrich (A7811) | IF |
| Mouse-anti-Myosin (fast) | 1:100 | Sigma-Aldrich (M4276) | IF |
| Mouse-anti-Titin | 1:50 | DSHB | IF |
| Rabbit-anti-Laminin | 1:100 | Sigma-Aldrich (L9393) | IHC |
| Mouse-anti-hSpectrin | 1:100 | Leica (SPEC1-CE) | IHC |
| Mouse-anti-hDystrophin | 1:100 | Millipore (MABT827) | IHC |
| Mouse-anti-hLaminA/C | 1:100 | Vector Laboratories (VP-L550) | IHC |
| GAA Exon 1-2 fw | AAACTGAGGCACGGAGCG | IDTDNA | RT-qPCR |
| GAA Exon 1-2 rv | GAGTGCAGCGGTGCCAA | IDTDNA | RT-qPCR |
| Set_1_fw | TCCCCAGGGCCGGTTAATGT | IDTDNA | PCR |
| Set_1_rv | GCTCTGGGCGGAGGAATATG | IDTDNA | PCR |
| Set_2_fw | CCTGAGTCCGGACCACTTTG | IDTDNA | PCR |
| Set_2_rv | CACCGGTTCAATTGCCGAC | IDTDNA | PCR |
| Set_3_fw | GTCTCTCACTCGGAAGGACAT | IDTDNA | PCR |
| Set_3_rv | TACCCGAAGAGTGAGTTGC | IDTDNA | PCR |

SUPPLEMENTARY METHODS

GAA enzyme activity assay

Differentiated myogenic progenitors were harvested with ice-cold protein lysis buffer (50 mM Tris (pH 7.5)), 100 mM NaCl, 50 mM NaF, 1% Triton X-100 and one tablet Protease Inhibitor Cocktail cOmplete, with EDTA, (Roche, Penzberg, Germany) for 10 minutes on ice. GAA enzyme activity was measured as described previously (Kroos et al., 2007). Total protein concentrations were determined with the BCA protein assay kit (Pierce, Thermo Scientific, Waltham, MA).

qRT-PCR

qRT-PCR was measured with a CFX96 real-time system (Bio-Rad, Hercules, CA). cDNA was diluted 5x or 10x times and 4 μ L was used in a qRT-PCR reaction consisting of a total volume of 15 μ L with 7.5 μ L iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules, CA), 10 pmol/ μ L forward and reverse primers (Table S4). Per plate, a standard curve was included with 5 dilutions.

Immunofluorescent analysis of *in vitro* differentiation

Myogenic progenitors were stained as described previously (van der Wal et al., 2017). Briefly, cells were permeabilized for 5 minutes with 0.1% Triton X-100 (AppliChem, Darmstadt, Germany) in PBS and blocked for 30 minutes at room temperature in blocking solution (PBS-T (0.1% Tween, Sigma-Aldrich, Irvine, UK) with 3% BSA (Sigma-Aldrich, Irvine, UK)). Primary antibodies (Table S4) were incubated for 1 hour at room temperature and diluted into 0.1% BSA in PBS-T, washed with PBS-T and incubated with secondary antibodies (1:500, Alexa-Fluor-488- α -mouse, A11001, Alexa-Fluor-594- α -rabbit, A10474, Alexa-Fluor-488- α -rabbit,

Supplemental Information

A11008, Invitrogen, Carlsbad, CA; or horse anti-mouse biotin, BA-2000, Vector Laboratories, Burlingame, CA). When a secondary biotinylated antibody was used, cells were washed three times for 5 minutes with PBS-T and incubated with Streptavidine 594 (1:500, S-32356, Invitrogen, Carlsbad, CA.). The cells were subsequently washed two times for 5 minutes with PBS and incubated for 15 minutes with Hoechst (1:15000, Thermo Scientific, Waltham, MA) before imaging.

Generation of induced pluripotent stem cells

Control iPSC lines were previously reprogrammed, characterized and cultured as described in van der Wal et al. (van der Wal et al., 2017). Healthy control 3 and healthy control 4 iPSCs were a gift from Dr. Mehrnaz and Prof. Joost Gribnau. Healthy control 2 (previously characterized in (Dambrot et al., 2013)), healthy control 5 (LUMC0004iCTRL10), and healthy control 8 (LUMC0030iCTRL12) iPSCs were gifts from Dr. Christian Freund and Prof. Christine Mummery. Using the MycoAlert™ Mycoplasma Detection Kit (Lonza, Walkersville, MD), the iPSC lines were regularly tested for contamination with mycoplasma. All results in this study were obtained with cultures that had tested negative. The identities of cell lines used in this study were confirmed by DNA sequencing.

Generation and expansion of myogenic progenitors from iPSCs

iPSC cultures in 100 mm dishes were used to initiate myogenic differentiation as described previously (van der Wal et al., 2017). Briefly, after 5 days of iPSC expansion, differentiation into myogenic progenitors was started with myogenic differentiation medium (DMEM/F12, 1% Insulin-Transferrin-Selenium-Ethanolamine (ITS-X), 1% penicillin/streptomycin/L-glutamine (P/S/G), all Gibco, Waltham, MA) supplemented with 3.5 μ M CHIR99021 (Axon Medchem, Groningen, the Netherlands) for 5 days; and changed to myogenic differentiation medium supplemented with 20 ng/ml FGF2 (Peprotech, Rocky Hill, NJ) for 14 days. For the last 16 days, cells were cultured in myogenic differentiation medium only. Myogenic progenitors were purified using FACS with anti-C-MET-APC (1:50, R&D systems, Minneapolis MN), and anti-HNK-1-FITC (1:100, Aviv Systems Biology, San Diego, CA) antibodies; and Hoechst (33258, Life Technologies, Carlsbad, CA) was added to stain live cells. The c-MET⁺/Hoechst⁺/HNK-1⁻ fraction was sorted in myogenic progenitor proliferation (MMP) medium (DMEM high-glucose supplemented with 10% fetal bovine serum, 1% penicillin/streptomycin/L-glutamine and 100 ng/ml FGF2) supplemented with 1x Revitacell supplement (Gibco, Waltham, MA) on ECM (Sigma-Aldrich, E6909)-coated dishes as described (van der Wal et al., 2017).

SUPPLEMENTARY REFERENCES

Borchin, B., Chen, J., and Barberi, T. (2013). Derivation and FACS-mediated purification of PAX3+/PAX7+ skeletal muscle precursors from human pluripotent stem cells. *Stem Cell Reports* 1, 620-631.

Caron, L., Kher, D., Lee, K.L., McKernan, R., Dumevska, B., Hidalgo, A., Li, J., Yang, H., Main, H., Ferri, G., et al. (2016). A Human Pluripotent Stem Cell Model of Facioscapulohumeral Muscular Dystrophy-Affected Skeletal Muscles. *Stem Cells Transl Med* 5, 1145-1161.

Chal, J., Al Tanoury, Z., Hestin, M., Gobert, B., Aivio, S., Hick, A., Cherrier, T., Nesmith, A.P., Parker, K.K., and Pourquie, O. (2016). Generation of human muscle fibers and satellite-like cells from human pluripotent stem cells in vitro. *Nat Protoc* 11, 1833-1850.

Chal, J., Oginuma, M., Al Tanoury, Z., Gobert, B., Sumara, O., Hick, A., Bousson, F., Zidouni, Y., Mursch, C., Moncuquet, P., et al. (2015). Differentiation of pluripotent stem cells to muscle fiber to model Duchenne muscular dystrophy. *Nat Biotechnol* 33, 962-969.

Charville, G.W., Cheung, T.H., Yoo, B., Santos, P.J., Lee, G.K., Shrager, J.B., and Rando, T.A. (2015). Ex Vivo Expansion and In Vivo Self-Renewal of Human Muscle Stem Cells. *Stem Cell Reports* 5, 621-632.

Choi, I.Y., Lim, H., Estrellas, K., Mula, J., Cohen, T.V., Zhang, Y., Donnelly, C.J., Richard, J.P., Kim, Y.J., Kim, H., et al. (2016). Concordant but Varied Phenotypes among Duchenne Muscular Dystrophy Patient-Specific Myoblasts Derived using a Human iPSC-Based Model. *Cell Rep* 15, 2301-2312.

Dambrot, C., van de Pas, S., van Zijl, L., Brandl, B., Wang, J.W., Schalij, M.J., Hoebe, R.C., Atsma, D.E., Mikkers, H.M., Mummery, C.L., et al. (2013). Polycistronic lentivirus induced pluripotent stem cells from skin biopsies after long term storage, blood outgrowth endothelial cells and cells from milk teeth. *Differentiation* 85, 101-109.

Supplemental Information

Kambara, H., Niazi, F., Kostadinova, L., Moonka, D.K., Siegel, C.T., Post, A.B., Carnero, E., Barriocanal, M., Fortes, P., Anthony, D.D., *et al.* (2014). Negative regulation of the interferon response by an interferon-induced long non-coding RNA. *Nucleic Acids Res* 42, 10668-10680.

Kim, J., Magli, A., Chan, S.S.K., Oliveira, V.K.P., Wu, J., Darabi, R., Kyba, M., and Perlingeiro, R.C.R. (2017). Expansion and Purification Are Critical for the Therapeutic Application of Pluripotent Stem Cell-Derived Myogenic Progenitors. *Stem Cell Reports* 9, 12-22.

Kroos, M.A., Pomponio, R.J., Hagemans, M.L., Keulemans, J.L., Phipps, M., DeRiso, M., Palmer, R.E., Ausems, M.G., Van der Beek, N.A., Van Diggelen, O.P., *et al.* (2007). Broad spectrum of Pompe disease in patients with the same c.-32-13T>G haplotype. *Neurology* 68, 110-115.

McGrath, E.L., Rossi, S.L., Gao, J., Widen, S.G., Grant, A.C., Dunn, T.J., Azar, S.R., Roundy, C.M., Xiong, Y., Prusak, D.J., *et al.* (2017). Differential Responses of Human Fetal Brain Neural Stem Cells to Zika Virus Infection. *Stem Cell Reports* 8, 715-727.

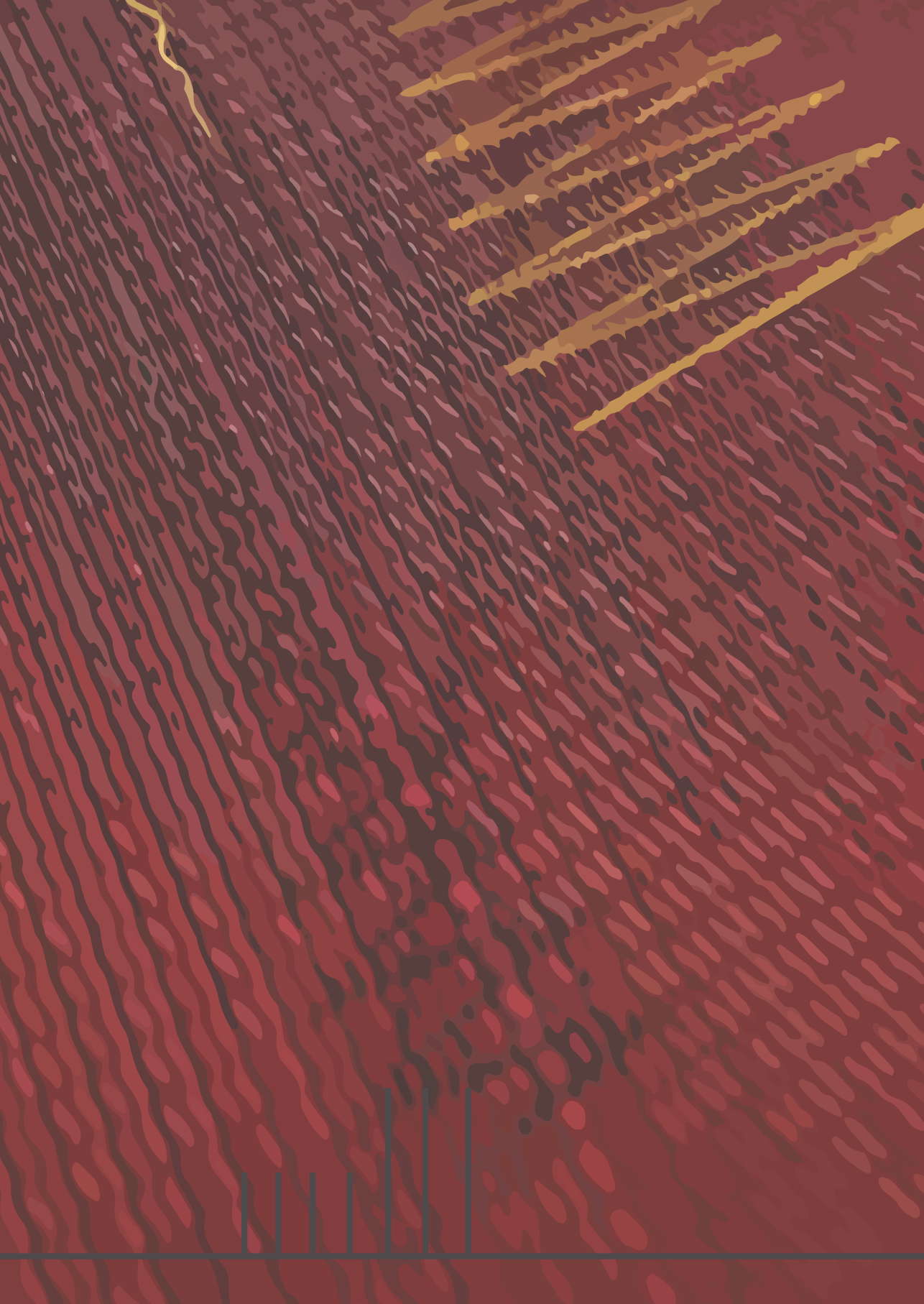
Shelton, M., Kocharyan, A., Liu, J., Skerjanc, I.S., and Stanford, W.L. (2016). Robust generation and expansion of skeletal muscle progenitors and myocytes from human pluripotent stem cells. *Methods* 101, 73-84.

Shelton, M., Metz, J., Liu, J., Carpenedo, R.L., Demers, S.P., Stanford, W.L., and Skerjanc, I.S. (2014). Derivation and expansion of PAX7-positive muscle progenitors from human and mouse embryonic stem cells. *Stem Cell Reports* 3, 516-529.

Swartz, E.W., Baek, J., Pribadi, M., Wojta, K.J., Almeida, S., Karydas, A., Gao, F.B., Miller, B.L., and Coppola, G. (2016). A Novel Protocol for Directed Differentiation of C9orf72-Associated Human Induced Pluripotent Stem Cells Into Contractile Skeletal Myotubes. *Stem Cells Transl Med* 5, 1461-1472.

van der Wal, E., Bergsma, A.J., van Gestel, T.J.M., In 't Groen, S.L.M., Zaehres, H., Arauzo-Bravo, M.J., Scholer, H.R., van der Ploeg, A.T., and Pijnappel, W. (2017). GAA Deficiency in Pompe Disease Is Alleviated by Exon Inclusion in iPSC-Derived Skeletal Muscle Cells. *Mol Ther Nucleic Acids* 7, 101-115.

Xu, C., Tabebordbar, M., Iovino, S., Ciarlo, C., Liu, J., Castiglioni, A., Price, E., Liu, M., Barton, E.R., Kahn, C.R., *et al.* (2013). A zebrafish embryo culture system defines factors that promote vertebrate myogenesis across species. *Cell* 155, 909-921.



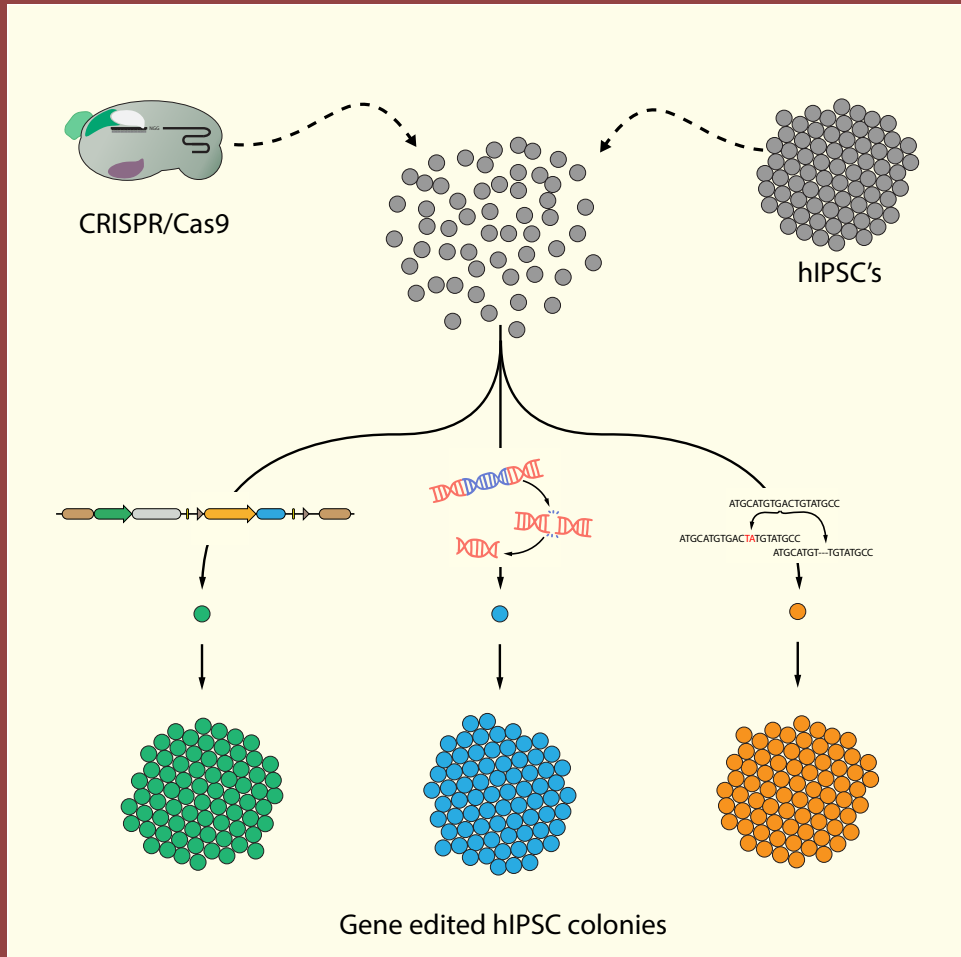
CHAPTER 7

CRISPR/CAS9-MEDIATED GENE EDITING IN HUMAN INDUCED PLURIPOTENT STEM CELLS

Stijn L.M. in 't Groen*, **Mike Broeders***, W.W.M. Pim Pijnappel

*These authors contributed equally to this work
CRISPR-Cas Methods: Volume II, Springer Nature, In press.

Graphical abstract



keywords: CRISPR/Cas9; iPSC, nucleofection; gene editing strategy; targeting plasmid; donor template integration; protocol

CRISPR/CAS9-MEDIATED GENE EDITING IN HUMAN INDUCED PLURIPOTENT STEM CELLS

Stijn L.M. in 't Groen^{1,2,3#}, Mike Broeders^{1,2,3#}, and W.W.M. Pim Pijnappel^{1,2,3*}

¹Department of Pediatrics, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands; ²Department of Clinical Genetics, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands; ³Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands

Abstract

Human induced pluripotent stem cell (iPSCs) and gene editing technologies have become broadly accessible in the last few years and are no longer confined to specialized laboratories. As a result of these developments, both techniques are becoming increasingly prominent in many fields of biomedical research. The use of the CRISPR/Cas9 platform has proven much less labour-intensive compared to alternative platforms for gene editing such as TALENs or ZFNs. However, application of CRISPR/Cas9 in human iPSCs can be cumbersome due to the relatively low efficiency of gene editing in these cells, combined with the requirement of advanced techniques for culturing human iPSCs. Here we provide protocols for CRISPR/Cas9-mediated gene editing in human iPSCs for the generation of gene knock outs, large deletions, and the introduction of a donor template in a safe harbor.

These authors contributed equally

Correspondence: W.W.M. Pim Pijnappel Department of Pediatrics,
Erasmus University Medical Center, Rotterdam, the Netherlands.

e-mail: w.pijnappel@erasmusmc.nl

1. Introduction

The discovery that human somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) has boosted research on stem cells, disease modelling, and regenerative medicine [1-3]. iPSCs can now be generated from a wide variety of somatic cells that can be obtained in a relatively easy manner, including skin, blood, urine, hair, and teeth [4-7]. Initial reprogramming protocols were quite inefficient and methods to culture human iPSCs required time consuming protocols, thus confining iPSC work to specialized laboratories. Today, improved protocols for the generation and maintenance of iPSCs have made iPSC technology more broadly accessible [2, 8, 9]. In addition, the development of improved protocols for the differentiation of iPSCs into distinct cell types is progressing [4, 5].

Research involving iPSCs is augmented by developments made in the field of gene editing. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR associated protein 9 (Cas9) has become the gene editing platform of choice in many laboratories because of its speed, low costs, and relative high efficiency compared to other gene editing methods such as transcription activator-like effector nucleases (TALENs) or zinc-finger nucleases (ZFNs). The first clinical trials involving gene editing are already ongoing [3]. Application of gene editing in iPSCs enables the introduction or removal of disease-associated variants, gene knock outs, large deletions (>1 kb), or the introduction of cDNAs in a safe harbor (a location in the genome that can be safely targeted without adverse cellular effects and that allows high expression of a transgene) for the generation of disease models and their isogenic controls, or for mechanistic studies on gene regulation [10, 11]. The generation of isogenic controls is instrumental in the correction for differences in genetic backgrounds, which appear to be very large among humans [12, 13]. Other applications of gene editing in iPSCs include the introduction of reporter constructs to monitor a biological process of interest, for example by using fluorescent proteins, and research in the field of regenerative medicine, in which patient-derived iPSCs are gene corrected and its differentiated derivatives are transplanted into disease models with the aim to replace tissue that has been lost due the disease [3, 7, 14]. Here we describe gene editing strategies applicable to human iPSCs utilizing CRISPR/Cas9 for the introduction of indels (using one single guide RNA (sgRNA)), the deletion of larger (>1 kb) regions (using two sgRNAs) and the insertion of large donor templates (using one sgRNA and a universal donor template) in a safe harbor while maintaining the integrity and differentiation potential of the iPSCs.

A general timeline for gene editing of iPSCs is shown in [Figure 1](#). The time required from target design to passaging of positive clones typically takes 19 to 33 days, depending on the application. On average, colonies can be picked around 14 days

after nucleofection. Usually, DNA can be isolated and used for genotyping after 4-5 days, before the colonies need to be passaged, but an additional passaging step may be required to obtain sufficient material for genotyping.



Figure 1. Timeline of gene editing in iPSCs. This protocol is focussed on the use of iPSCs cultured in the presence of mouse embryonic fibroblasts (MEFs). However, with minor adjustments to the protocol provided at the end of this manuscript, this strategy can also be applied to iPSCs cultured under feeder free conditions.

2. Generation of the targeting plasmid

The targeting plasmid contains the sgRNA that guides the Cas protein to the target sequence. *In silico* prediction tools should be used to identify the optimal target sequence, assessing both on and off-target activity [10, 15]. Once the optimal target sequence has been determined *in silico*, the presence of this exact sequence should be verified by Sanger sequencing in the iPSCs that will be used in the experiment. This is important because the presence of single-nucleotide polymorphisms (SNPs) in the target sequence will reduce targeting efficiency. Single strand oligonucleotides can then be ordered, annealed and inserted into the plasmid containing two BbsI restriction sites for the sgRNA cloning (Figure 2). Transcription of the sgRNA is driven by the U6 promoter. To allow efficient transcription, the 20th base of the guide sequence (5' from the PAM sequence) should be a guanine, if not substitute this base with a guanine. We have used the following targeting sequence for the insertion of a cDNA in the AAVS1 locus: 5'-GTCACCAATCCTGTCCCTAG-3', using the donor construct described in section 3 [16].

Materials:

- 10X Annealing buffer (100mM Tris, pH7.5-8.0, 500mM NaCl, 10mM EDTA)
- Forward oligonucleotide (the target sequence and ACCG overhang)
- Reverse oligonucleotide (the target sequence and CAAA overhang)
- Milli-Q water
- Cutsmart buffer (NEB, B7204S or supplied with BbsI-HF)
- BbsI-HF (NEB, R3539)
- pCRII-BbsI-sgRNA scaffold vector (Addgene, 159352)
- Agarose (Sigma, A9539)

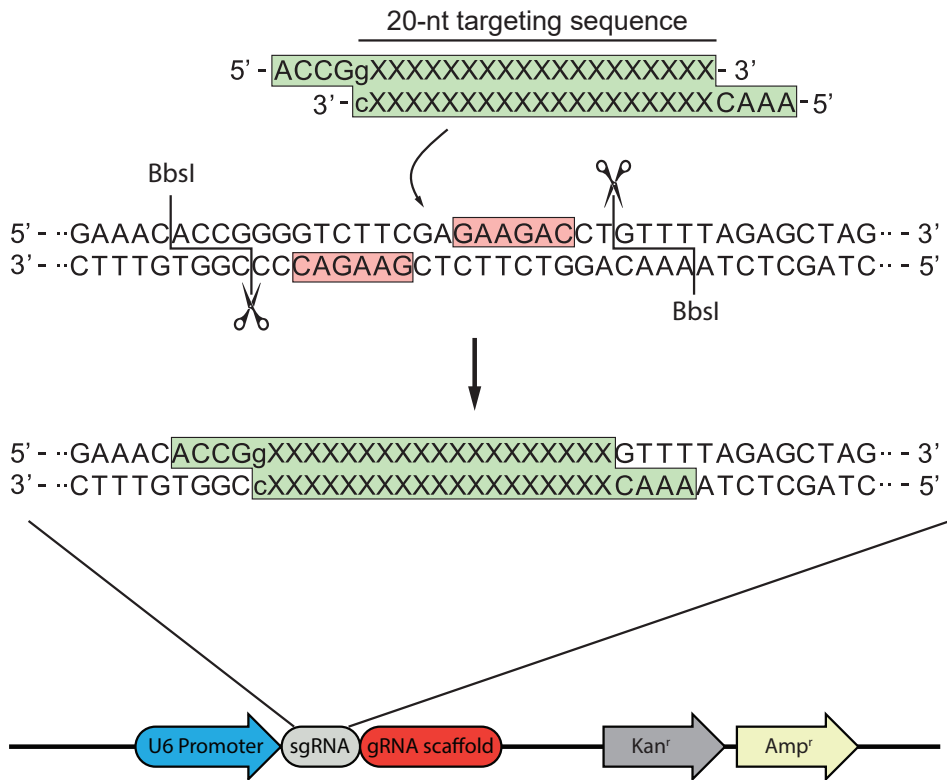


Figure 2. Scheme for the cloning of the sgRNA sequence into the targeting plasmid. The DNA oligonucleotide duplex (in green) is formed by annealing two complementary single strand oligonucleotides containing the target sequence and an overhang. The 20th base of the guide sequence should be a guanine, indicated by the small g. Using the BbsI restriction enzyme, the pCRII-BbsI-sgRNA scaffold plasmid (we designed this plasmid based on Ran et al. [17]) is digested, producing two asymmetric overhangs (indicated with the scissors). This allows the oligonucleotide duplex to be inserted unidirectionally into the pCRII-BbsI-sgRNA scaffold.

- 10x TAE Buffer (40mM Tris, 20mM acetic acid, 1mM EDTA)
- Gel extraction kit (Qiagen, 28704)
- T4 DNA Ligase (NEB, B0202S)
- T4 DNA Ligase Buffer (10X) (NEB, M0202 or supplied with T4 DNA Ligase)
- Heat shock competent cells (One Shot TOP10, Invitrogen, C4040)
- LB agar plates with 100µg/mL Ampicillin and/or 50µg/mL Kanamycin selection
- M13 forward primer (TGTAACGACGGCCAGT) or T7 sequence primer (TAATACGACTCACTATAGGG)
- Miniprep kit (Qiagen, 27106)

- Midi or Maxiprep kit (Qiagen, 740410 or 740414)

Procedure:

1. Annealing of the complementary oligonucleotides to form an oligonucleotide duplex

1. Mix the two single stranded oligonucleotides in equimolar concentrations as described below

| Oligonucleotide annealing mix for 1 reaction | |
|--|---|
| $X\mu\text{L}$ | Forward oligonucleotide (100 μM final concentration) |
| $X\mu\text{L}$ | Reverse oligonucleotide (100 μM final concentration) |
| $2\mu\text{L}$ | 10X Annealing buffer |
| $X\mu\text{L}$ | Milli-Q water |
| $20\mu\text{L}$ | Total volume |

2. Efficient annealing can be achieved by one of the two following methods

Oligonucleotide annealing method 1.

- Prepare and mix oligonucleotides in a 1.5mL microfuge tube.
- Heat to 95°C for 5 minutes in a heating block.
- Turn off the heating block and allow to slowly cool to room temperature (~45 minutes).

Oligonucleotide annealing method 2.

- Prepare and mix oligonucleotides in a PCR tube.
- Place the mixture in the thermocycler and use the following PCR program.

| Annealing program | |
|-------------------|------------------------|
| 1) | 95°C 5:00 |
| 2) | 95°C (-1°C/cycle) 2:00 |
| 3) | 20°C ∞ |
| 4) | End |

3. The resulting oligonucleotide duplex can be stored at 4°C for short term (one week) or at -20°C for long term (up to 12 months).

2. Digestion of the pCRII-BbsI-sgRNA scaffold plasmid

1. Perform a restriction reaction with the BbsI restriction enzyme on the pCRII-

BbsI-sgRNA scaffold plasmid as described below:

| <i>BbsI-HF restriction mix for 1 reaction</i> | |
|---|-----------------------------------|
| <i>X</i> μL (~2μg) | pCRII-BbsI-sgRNA scaffold plasmid |
| 5μL | Cutsmart buffer |
| <i>X</i> μL | Milli-Q water |
| 2μL | BbsI-HF |
| 50μL | Total volume |

- Incubate the reaction mix for 60 minutes at 37°C.
 - Run the restriction reaction on a 0.75% agarose TAE gel.
 - Cut the linearized plasmid (size 4407bp) from the gel using a scalpel. Note: use a low intensity UV source to visualize the DNA to prevent UV-induced damage.
 - Perform a gel extraction to isolate the product following the manufacturer's protocol.
 - Quantify the gel-purified DNA using a spectrophotometer
3. **Ligation of the oligonucleotide duplex into the pCRII-BbsI-sgRNA scaffold plasmid**
- Dilute the oligonucleotide duplex 200-fold in Milli-Q water.
 - Prepare and mix the T4 DNA ligation mix as described below:

| <i>T4 DNA Ligase reaction mix for 1 reaction</i> | |
|--|---|
| 2μL | T4 DNA Ligase Buffer (10X) |
| <i>X</i> μL (50ng) | Digested pCRII-BbsI-sgRNA scaffold plasmid |
| 2μL | 200 fold diluted duplex oligonucleotide mix |
| 15- <i>X</i> μL | Milli-Q water |
| 1μL | T4 DNA Ligase |
| 20μL | Total volume |

- Incubate for 60 minutes at room temperature, overnight at 16°C or 4°C over the weekend. (include a negative control ligation reaction that lacks the oligonucleotide duplex).
- Transform 10μL of the ligation into competent cells using the heatshock method according to manufacturer's protocol.
- Plate the transformed cells onto LB agar plates with Ampicillin (100μg/mL) and/or Kanamycin (50μg/mL) selection and incubate overnight at 37°C.
- After overnight incubation, check for the presence of colonies (typically very

few colonies should be present on the negative control and hundreds of colonies on the ligation).

7. Pick colonies from the ligation plate and perform a miniprep DNA purification according to the manufacturer's protocol.
8. Sequence the clones with the M13 forward primers or T7 primer to verify the correct insertion of the duplex oligonucleotide.
9. Colonies with a correct insertion can be used for a Midi or Maxi prep according to the manufacturer's protocol.
10. The resulting targeting plasmid will be used for further downstream applications and can be stored at 4°C for short term (one week) or at -20°C for long term (up to 12 months).

Note that once a target sequence has been introduced into the pCRII-BbsI-sgRNA scaffold plasmid, both BbsI restriction sites are no longer present. Therefore, the plasmid containing the sgRNA cannot be reused to clone an alternative sgRNA.

3. Generation of the donor construct

To generate a large knock-in, the use of a donor construct is required. The efficiency of generating a large knock-in is significantly lower than generating an indel or deletion. Therefore, utilizing a selection cassette to select for the successfully targeted clones can reduce the number of negative colonies (Figure 3).

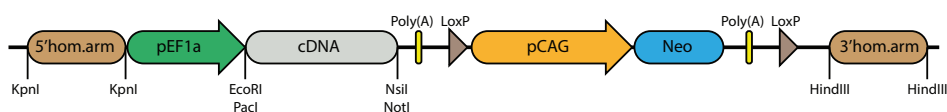


Figure 3. Map for cloning of the cDNA insert into the donor plasmid. The EF1α-cDNA-pCAG-Neo plasmid contains two KpnI recognition sites flanking the 5' homology arm and two HindIII recognition sites flanking the 3' homology arms [16]. The cDNA is expressed by the EF1α promoter and is flanked by 5' EcoRI and PacI and 3' NsiI and NotI recognition sites. The Neomycin selection is driven by the pCAG promoter and enables the selection of successful targeted cells with G418. If desired, the LoxP sites can be used to remove the Neomycin selection cassette by transient expression of Cre recombinase in the targeted iPSCs [18].

Materials:

- Milli-Q water
- Cutsmart buffer (NEB, B7204S or supplied with restriction enzyme)
- EcoRI-HF (NEB, R3101)
- PacI (NEB, R0547)

- NsiI-HF (NEB, R3127)
- NotI-HF (NEB, R3189)
- EF1 α -cDNA-pCAG-Neo vector (the plasmid containing acid alpha-glucosidase cDNA and AAVS1 target sites can be used to clone the cDNA and target sites of interest [16] and is available upon request.
- Agarose (Sigma, A9539)
- 10x TAE Buffer (40mM Tris, 20mM acetic acid, 1mM EDTA)
- Gel extraction kit (Qiagen, 28704)
- T4 DNA Ligase (NEB, B0202S)
- T4 DNA Ligase Buffer (10X) (NEB, M0202 or supplied with T4 DNA Ligase)
- Heat shock competent cells (One Shot TOP10, Invitrogen, C4040)
- LB agar plates with 100 μ g/mL Ampicillin and/or 50 μ g/mL Kanamycin selection
- Sequence primers for insert
- Miniprep kit (Qiagen, 27106)
- Midi or Maxiprep kit (Qiagen, 740410 or 740414)

Procedure:

1. Digestion of the EF1 α -cDNA-pCAG-Neo plasmid

1. Perform a restriction reaction with the restriction enzymes on the EF1 α -cDNA-pCAG-Neo plasmid as described below:

| Enzyme restriction mix for 1 reaction | |
|---------------------------------------|-------------------------------------|
| <i>X</i> μ L (~2 μ g) | EF1 α -cDNA-pCAG-Neo plasmid |
| 5 μ L | Cutsmart buffer |
| <i>X</i> μ L | Milli-Q water |
| 2 μ L | EcoRI-HF or PacI |
| 2 μ L | NsiI-HF or NotI-HF |
| 50 μ L | Total volume |

2. Incubate the reaction mix for 60 minutes at 37°C.
3. Run the restriction reaction on a 1% agarose gel.
4. Cut the linearized plasmid (size ~9100bp) from the gel using a scalpel. Note: use a low intensity UV source to visualize the DNA to prevent UV-induced damage.
5. Perform a gel extraction to isolate the product following the manufacturer's protocol.
6. Quantify the gel-purified DNA using a spectrophotometer.

2. **Ligation of the cDNA insert into the EF1α-cDNA-pCAG-Neo plasmid**

Standard cloning techniques can be used to prepare the required cDNA insert with 5' EcoRI-HF or PacI and a 3' NsiI-HF or NotI-HF overhangs (for example using PCR or ordered as gBlock (IDT)).

1. Prepare and mix the T4 ligation mix as described below:

| T4 Ligase reaction mix for 1 reaction | |
|---------------------------------------|-------------------------------------|
| 2μL | T4 DNA Ligase Buffer (10X) |
| XμL (50ng) | Digested EF1α-cDNA-pCAG-Neo plasmid |
| XμL | Digested cDNA insert |
| XμL | Milli-Q water |
| 1μL | T4 DNA Ligase |
| 20μL | Total volume |

Use the following formula to calculate the required amount of insert in ng:

$$\frac{\text{ng of vector} \times \text{size of insert in kb}}{\text{size of vector in kb}} \times \text{molar ratio of } \frac{\text{insert}}{\text{vector}} = \text{ng insert}$$

- Incubate 60 minutes at room temperature, overnight at 16°C or 4°C over the weekend. (Remember to include a negative control ligation reaction that lacks the cDNA insert).
- Transform 10μL of the ligation into competent cells using the heatshock method according to manufacturer's protocol.
- Plate the transformed cells onto LB agar plates with Ampicillin (100μg/mL) and/or Kanamycin (50μg/mL) selection and incubate overnight at 37°C.
- After overnight incubation, check for the presence of colonies on the agar plates (typically very few colonies should be present on the negative control and hundreds of colonies on the ligation).
- Pick colonies from the ligation plate and perform a miniprep according to manufacturer's protocol.
- Sequence the clones to verify the successful insertion of the cDNA insert.
- Colonies with a successful insertion can be used for a Midi or Maxi prep according to manufacturer's protocol, the resulting targeting plasmid will be used for further downstream applications and can be stored at 4°C for short term (one week) or at -20°C for long term (up to 12 months).

Note 1: The plasmid is now ready to serve as donor vector for the insertion of the cDNA of choice in the AAVS1 site. If desired, the cDNA can be inserted

at a different genomic site of choice. In this case, the homology arms should be adjusted accordingly using the KpnI and HindIII sites for the 5' and 3' homology arms, respectively. The size of the homology arms is typically 500 bp to 1kb in length. It is advised to double check that the targeting sequence is present without SNPs in the iPSCs used for the experiment, otherwise targeting efficiency will be reduced.

Note 2: This donor plasmid uses neomycin as a selection marker in iPSCs. Prior to applying selection with G148, a kill curve should be made for each cell line to determine the optimal concentration.

4. Generation of conditioned media

Conditioned medium from MEFs is used during and after nucleofection of the iPSCs. Conditioned medium contains factors that are secreted by MEFs such as growth factors and extracellular proteins and is harvested every 24 hours (Figure 4). The conditioned medium is added immediately after plating to improve the recovery of iPSCs from nucleofection.

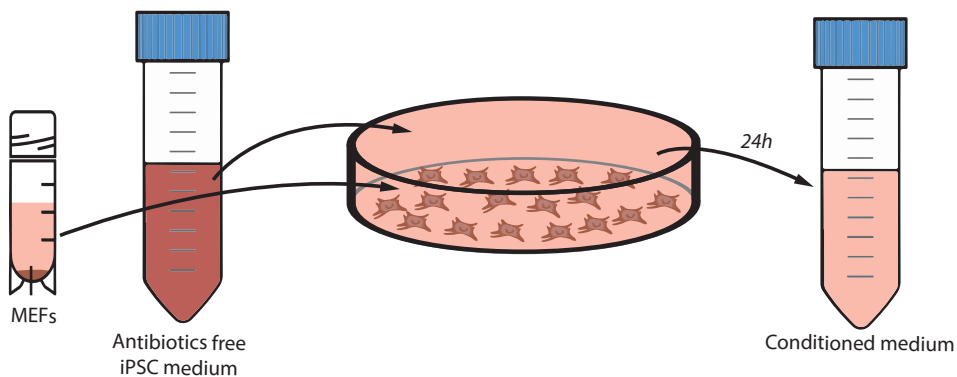


Figure 4. The generation of conditioned medium. Antibiotics-free iPSC medium is added to MEFs and harvested after 24 hours for later use.

Materials:

- Irradiated Mouse Embryonic Fibroblasts (MEFs)
- 2% gelatin solution (Sigma, G-1393)
- PBS (Gibco, 70011044)
- Fibroblast growth medium
 - DMEM high glucose (Gibco, 11965092)
 - 10% fetal bovine serum (Hyclone, 11531831)

- 1% penicillin-streptomycin-glutamine (P/S/G) (Gibco, 10378016)
- Antibiotics free iPSC medium
 - 390mL DMEM/F12 (Invitrogen, 21331046)
 - 10% KO serum replacement (Invitrogen, 10828)
 - 1% Non-essential amino acids (NEAA) (Gibco, 11140050)
 - 1% Glutamine (Gibco, 25030024)
 - 1mL β -Mercaptoethanol (Invitrogen 31350010)
 - 10ng/mL basic fibroblast growth factors (bFGF) (Preprotech, 100-18B) (Dissolved in 0.1% BSA/PBS, see manufacturer's instructions)
- 10 cm tissue culture plate (Greiner Bio-One, 664160)
- 0.45 μ m sterile cell culture filter (Millipore, SLHVR04NL)

Procedure:

1. Coat a 10cm tissue culture plate with 5mL 0.1% gelatin solution (diluted in PBS) and incubate for 15 minutes at 37°C.
2. Thaw a cryovial containing the MEFs in a 37°C water bath until almost completely thawed and gently transfer the MEFs to a 15mL tube containing 9mL fibroblast growth medium using a P1000 pipette.
3. Centrifuge 1 million MEFs at 1000 rpm for 5 minutes, remove the excess medium and resuspend the pellet in 10mL fibroblast growth medium.
4. Seed the MEFs onto the gelatin coated tissue culture plate and culture at 37°C/5% CO₂.
5. Refresh the media after 8 to 24 hours with antibiotics free iPSC medium.
6. Harvest the media after 24 hours and refresh the MEFs with antibiotics free iPSC medium. This step can be repeated for up to five days or until quality of the MEFs have been diminished as such they are no longer considered viable.
7. Filter the conditioned media with using a 0.45 μ m sterile cell culture filter.
8. Store the sterile conditioned medium at -20°C for short time storage of -80°C for long time storage.

Note: Each nucleofection reaction requires approximately 20mL of conditioned medium.

5. Preparation of the DNA prep (1 day in advance)

Depending on the method of plasmid preparation, an optional co-precipitation of the plasmids prior to transfection may be performed. This step is recommended to prevent microbial contamination.

Materials:

- 5M NaCl
- Milli-Q water
- Ice-cold 100% ethanol
- Ice-cold 70% ethanol
- sgRNA targeting plasmid (Generated in section 2)
- pCas9_GFP plasmid (Addgene, 44719)
- *Optional: Donor template plasmid (available upon request)*

Procedure:

Note: all steps should be performed in a cell culture hood to avoid contamination and maintain sterility

1. Prepare and mix the DNA prep mix to a 1.5mL Eppendorf tube as described below:

| Single sgRNA* | Double sgRNA** | Donor template insertion*** | |
|------------------|--------------------|-----------------------------|-------------------------------|
| 11.5µg (~2 pmol) | 11.5µg (~2pmol) | 8.9µg (~1.5 pmol) | Cas9 plasmid |
| 8.5µg (~3 pmol) | 4.25µg (~1.5 pmol) | 6.7µg (~2.25 pmol) | 1 st sgRNA plasmid |
| - | 4.25µg (~1.5 pmol) | - | 2 nd sgRNA plasmid |
| - | - | 4.4µg (~1 .25 pmol) | Donor plasmid |
| 2µL | 2µL | 2µL | NaCl (5M) |
| XµL | XµL | XµL | Milli-Q water |
| 100µL | 100µL | 100µL | Total volume |

* for generating one double stranded DNA break, e.g. to create a knock out

** for generating two double stranded DNA breaks, e.g. to create a large deletion

*** to insert a cDNA in a safe harbor

2. Add 250µl ice-cold 100% ethanol and precipitate the DNA for 15 minutes at -20°C.
3. Centrifuge for 10 minutes at 14.000 rcf at 4°C and remove the ethanol, the plasmids will appear as a small translucent pellet
4. Wash with 200µl ice-cold 70% ethanol and centrifuge for 10 minutes at 14.000 rcf at 4°C, repeat this once.
5. Remove the excess ethanol from the Eppendorf tube and allow the pellet to air dry for 15 minutes.
6. Add 20µl of sterile PBS onto the dry pellet, close the tube and let the DNA resuspend overnight at room temperature.

7. Transfer 2µl of the DNA prep to a new 1.5mL Eppendorf tube and quantify the DNA using a spectrophotometer.

6. Plating MEFs

After nucleofection the iPSCs are seeded on fresh MEFs (Figure 5). To allow time for the MEFs to adhere to the plate, they have to be plated at least 8 hours (preferably 24 hours) prior to seeding the nucleofected iPSCs. This protocol is for one 6-well plate but can be scaled accordingly.

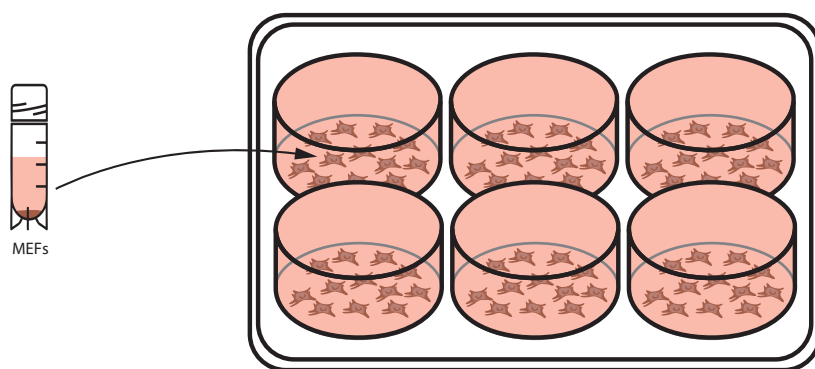


Figure 5. Plating of mouse embryonic fibroblasts (MEFs). 1 million cells are divided over all wells of a 0.1% gelatin coated 6-well plate.

Materials:

- Irradiated Mouse Embryonic Fibroblasts (MEFs)
- 2% gelatin solution (Sigma, G-1393)
- PBS (Gibco, 70011044)
- Fibroblast growth medium
 - DMEM high glucose (Gibco, 11965092)
 - 10% fetal bovine serum (Hyclone, 11531831)
 - 1% penicillin-streptomycin-glutamine (P/S/G) (Gibco, 10378016)
- 6-well tissue culture plate (Thermo Scientific, 140675)

Procedure:

1. Coat the 6-well tissue culture plate with 1mL 0.1% gelatin solution (diluted in PBS) per well and incubate for 15 minutes at 37°C.
2. Thaw the cryovial containing the MEFs in a 37°C water bath until almost completely thawed and gently transfer the MEFs to a 15mL tube containing 9mL fibroblast growth medium using a P1000 pipette.

3. Centrifuge 1 million MEFs at 1000 rpm for 5 minutes, remove the excess medium and resuspend the pellet in 12mL fibroblast growth medium.
4. Seed 2mL of the MEFs suspension per well onto the gelatin coated tissue culture plate and culture at 37°C/5% CO₂.

7. Nucleofection of iPSCs

The plasmid DNA for the sgRNA, Cas9 protein and the donor vector (optional) are introduced into the iPSCs using nucleofection (Figure 6).

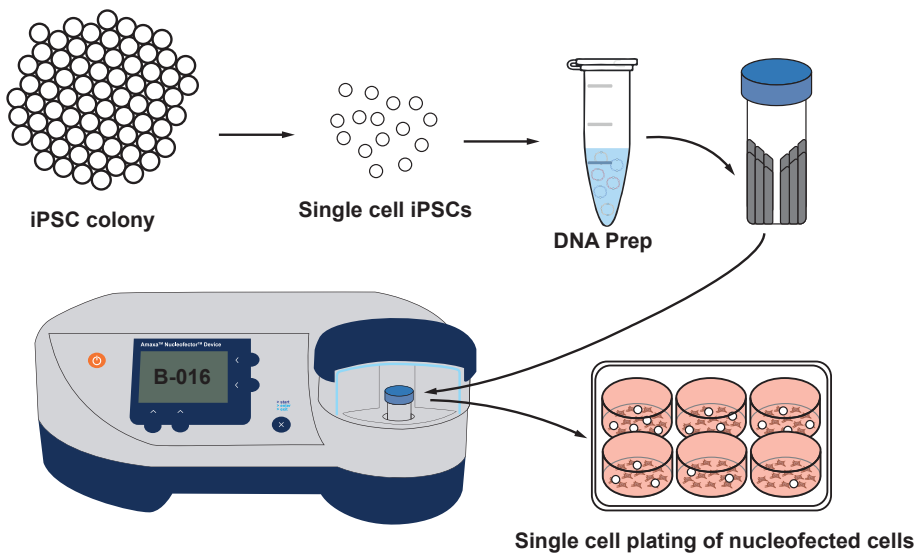


Figure 6. Procedure for nucleofection of single cell iPSCs for CRISPR/Cas9-mediated gene editing. iPSCs are dissociated into single cells and mixed with DNA. After nucleofection, cells are plated as single cells at different densities on MEFs.

Materials:

- Induced Pluripotent Stem cell medium
 - 390mL DMEM/F12 (Invitrogen, 21331046)
 - 10% KO serum replacement (Invitrogen, 10828)
 - 1% Non-essential amino acids (NEAA) (Gibco, 11140050)
 - 1% Penicillin-Streptomycin-Glutamine 100x (Gibco, 10378016)
 - 1mL β-Mercaptoethanol (Invitrogen 31350010)
 - 10ng/mL basic fibroblast growth factor (bFGF) (Preprotech, 100-18B) (Dissolved in 0.1% BSA/PBS, see manufacturer's instructions)

- Nucleofector™ 2b Device (Lonza, AAB-1001)
- Human Stem Cell Nucleofector™ Kit 2 (Lonza, VAPH-5022)
- DNA prep (prepared in section 5)
- Conditioned medium from MEFs (prepared in section 4)
- Accutase (Gibco, A11105-01) or TrypLE (Gibco, 12605010)
- PBS (Gibco, 70011044)
- ROCK inhibitor Y-27632 (Hello Bio, HB2297) or Revitacell Supplement 100X (Gibco, A2644501)
- basic fibroblast growth factor (bFGF) (Preprotech, 100-18B) (Dissolved in 0.1% BSA/PBS, see manufacturer's instructions)
- G418 (InvivoGen, ant-gn-5)

Procedure:

1. Four hours before starting the nucleofection procedure: replace the medium on the iPSCs with iPSC medium supplemented with 10μM ROCK inhibitor or 1x Revitacell Supplement.
2. 30 minutes before starting the procedure: replace the medium on the MEFs with conditioned medium supplemented with 10ng/mL bFGF and 10μM ROCK inhibitor or 1x Revitacell Supplement.
3. Transfer 9μg of the DNA prep to a new sterile 1.5mL Eppendorf tube, prepare one tube for each nucleofection reaction.
4. Remove the iPSC medium and wash the iPSCs with 2mL sterile PBS
5. Incubate the cells with 500μl of warm Accutase or TrypLE at 37°C until the cells start to detach; this should take 5-10 minutes.
6. Using a 10mL pipette, add 2mL of iPSC medium onto each well and detach the cells from the bottom of the well by pipetting the medium gently up and down the well. Transfer the cell suspension to a 50mL tube.
7. Count the cells and transfer 2 million cells into a new 50mL tube for each nucleofection reaction.
8. Centrifuge the cell suspension for 5 minutes at 1000rpm and carefully remove all medium.
9. Mix solutions A and B from the Human Stem Cell Nucleofector Kit 2; 100μl nucleofection mix is required per reaction, the rest of the mix can be stored up to 1 month at 4°C.

Note: It is recommended to perform the following steps for one reaction at a time

10. Resuspend the pellet of two million cells in 100μl of the Human Stem Cell

Nucleofector mix by pipetting up and down twice using a P1000 pipette.

11. Transfer the resuspended cells into the 1.5mL Eppendorf containing the 9µg DNA prep.
12. Mix the resuspended cells and the DNA prep by pipetting up and down four times using a P1000 pipette.
13. Carefully transfer the mix to a Human Stem Cell Nucleofector Kit 2, make sure that no air bubbles are introduced. If any bubbles do appear in the cell suspension, gently tap the bottom of the cuvette on the surface of the cell culture hood to remove them.
14. Put the lid on the cuvette and place it into the Nucleofector™ 2b. Select program [B-016] and press enter. After 2 seconds, an [OK] message appears on the display of the device indicating that the program was successfully executed. A white layer containing dead cells will be formed on one side of the cuvette.
15. Bring the cuvette back into the cell culture hood and transfer the cell suspension using the plastic transfer pipette included kit to a 50mL tube containing 2mL of conditioned medium supplemented with 10ng/mL bFGF and 10µM ROCK inhibitor or 1x Revitacell Supplement. Try to only transfer the suspension and avoid the white layer of dead cells.
16. Seed the iPSCs on the prepared MEFs in conditioned medium and culture at 37°C/5% CO₂. It is recommended to seed the iPSCs at several dilutions ($1/3$, $1/6$, $1/9$) in order to generate single cell colonies.
17. After 24 hours, replace the conditioned medium with iPSC medium and refresh the media daily.
18. The selection with G418 can be started 48 hours after nucleofection if the donor template is used.
19. 7 to 21 days after nucleofection single colonies can be picked (see section 8). The time required to obtain a colony depends on the cell density, recovery speed of the cells, and the use of selection.

8. Picking colonies

Once the colonies are large enough to be passaged, the colonies are cut with a 23 gauge needle and split manually into two wells, of which one is used for genotyping and one to continue passaging after genotyping (Figure 7).

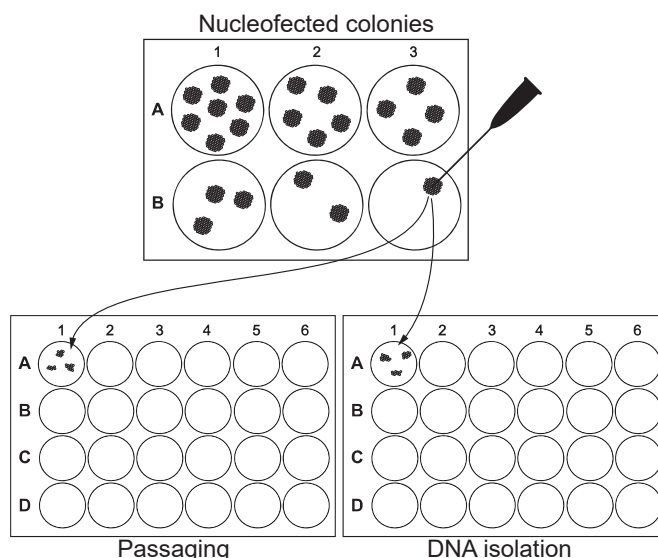


Figure 7. Picking iPSC colonies after nucleofection. iPSC colonies are dissociated, cut from the plate using a 23 gauge needle, and passaged into a plate for DNA isolation and another plate for passaging.

Materials:

- Induced Pluripotent Stem cell (iPSC) medium
 - 390mL DMEM/F12 (Invitrogen, 21331046)
 - 10% KO serum replacement (Invitrogen, 10828)
 - 1% Non-essential amino acids (NEAA) (Gibco, 11140050)
 - 1% Penicillin-Streptomycin-Glutamine 100x (Gibco, 10378016)
 - 1mL β -Mercaptoethanol (Invitrogen 31350010)
 - 10ng/mL basic fibroblast growth factors (bFGF) (Preprotech, 100-18B) (Dissolved in 0.1% BSA/PBS, see manufacturer's instructions)
- Irradiated Mouse Embryonic Fibroblasts (MEFs)
- 2% gelatin solution (Sigma, G-1393)
- PBS (Gibco, 70011044)
- Fibroblast growth medium
 - DMEM high glucose (Gibco, 11965092)
 - 10% fetal bovine serum (Hyclone, 11531831)
 - 1% penicillin-streptomycin-glutamine (P/S/G) (Gibco, 10378016)
- 48-well tissue culture plate (Greiner bio-one, 677180)
- 1mg/mL Collagenase IV (Invitrogen 17104-019) in KO DMEM/F12 (Invitrogen 21331046) (Dissolve at 37°C for 10-15 minutes, filter through 0.2 μ M sterile filter (Millipore, SLFGR04NL))
- 23 gauge needle

Procedure:

1. 24 hour prior to picking: prepare two 48-well plates per nucleofection with MEFs (scale down from section 6). One will be used for DNA isolation and one for passaging of the colonies.
2. Before picking colonies: Rinse MEFs with 1mL PBS and add iPSC medium.
3. Remove the iPSC medium and wash with 1mL PBS.
4. Add 1mL collagenase IV (1mg/ml) solution into each well.
5. Incubate the plate at 37°C for 5-15 minutes.
 - a. monitor the cell detachment under microscope as more time might be needed.
 - b. the edge of detached colonies should look slightly “curled” comparing to the attached ones.
6. Add 1mL iPSC medium into each well.
7. Cut the selected single colony into small pieces with a 23 gauge needle:
 - a. hold needle in an upward direction of needle opening.
 - b. scrape gently, avoid cutting the plastic surface.
8. Dissociate the selected colony from the plate with a P1000 and divide over two wells, one on each plate.
9. Repeat until all the selected single colonies are picked.
10. Refresh the media and monitor the colonies daily until passaging or harvesting DNA for genotyping.

9. Genotyping

Normally, the genotyping can be finished before the sister colony has to be passaged, but if required it can be also performed on cells of a later passage. The method of genotyping will differ depending on the gene editing strategy (Figure 8). For small indels and large deletions, a generic PCR can be performed using primers flanking the target sequence(s), and the genomic alteration of the target site can be determined by Sanger sequencing of the PCR product(s). For large deletions, a dual-PCR strategy can be used to determine the mono-allelic or bi-allelic presence of the deletion. For cDNA insertions mediated by a donor construct, a dual-PCR strategy can be used to determine whether the donor template has integrated at the target location. To determine copy number variations and to further examine genomic changes, Southern blotting and/or qPCR analysis of genomic DNA can be performed. Finally, it is important to monitor any off-target events resulting from the CRISPR reaction. This is usually done by analyzing predicted off-target sites using PCR and Sanger sequencing, although this might not always be sufficient. For a more extensive discussion on off-target effects, see [10, 19-21].

Below we describe examples for the genotyping of indels, deletions and template knock-ins.

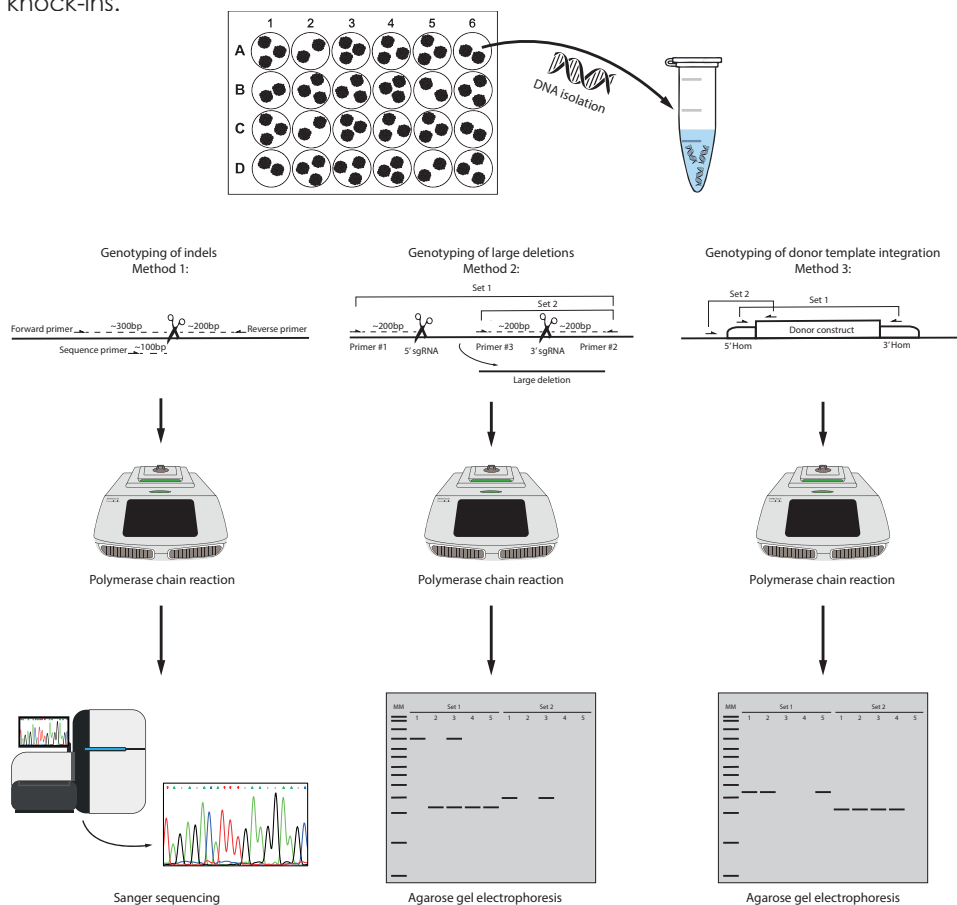


Figure 8. Genotyping methods to detect indels, large deletions and template integrations. DNA is isolated from the iPSC colonies and used for a PCR-based genotyping strategy. Method 1 uses Sanger sequencing to determine the indels, methods 2 and 3 use agarose gel electrophoresis to identify successfully targeted colonies. Typical results for agarose gel electrophoresis are shown.

DNA isolation

Materials:

- Lysis buffer (Tris pH 8.5 0.1M, EDTA 5mM, SDS 0.2%, NaCl 0.2M; add 100µg/mL fresh protease K)
- NaCl (5M)
- Isopropanol (Sigma, 59300)
- 70% Ethanol (Sigma, 72032221)
- Milli-Q water

Procedure:

1. Remove the iPSC medium and wash the iPSCs once with PBS.
2. Add 500µl lysis buffer to each well and incubate at 37°C for 1-18 hours (in cell culture incubator).
3. Transfer the cell lysate to a 1.5mL tube (Optional: Store lysate at -20°C). Subsequent steps are performed at room temperature, unless stated otherwise.
4. Add 260µl NaCl (5M) and shake (do not vortex to avoid breaking the genomic DNA), a white protein precipitate forms.
5. Centrifuge for 5 minutes at maximum speed and transfer the supernatant to a new 1.5mL tube, without touching the white pellet.
6. Add 532µl (0.7x volume) isopropanol, shake (do not vortex), a small piece of DNA should appear, if not shake again.
7. Centrifuge 5 minutes at maximum speed.
8. Remove the supernatant and wash the pellet with 500µl 70% ethanol and centrifuge 5 minutes at maximum speed.
9. Remove the excess ethanol from the Eppendorf tube and allow the pellet to air dry for 15 minutes.
10. Dissolve the pellet in 40µl Milli-Q water and incubate for 1 hour at 65°C.
11. Quantify the extracted DNA using a spectrophotometer.

Genotyping method 1: introduction of indels (nested DNA sequencing)

For the genotyping of indels a PCR reaction with primers flanking the targeted area is used to amplify the DNA, this is subsequently sequenced to verify the introduction of indels (*Figure 9*).

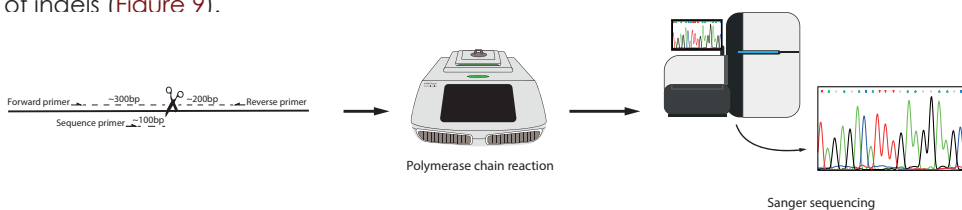


Figure 9. Genotyping of indels. Primers flanking the indel amplify the region, after which the indel can be determined by Sanger sequencing.

Materials:

- Forward primer, located ~300bp upstream the sgRNA sequence (stock: 10µM in 10mM Tris)
- Reverse primer, located ~200bp downstream the sgRNA sequence (stock: 10µM in 10mM Tris)

- Sequence primer, located ~100bp upstream the sgRNA sequence (stock: 10μM in 10mM Tris)
- Isolated DNA (isolated previously from individual iPSC colonies)
- Milli-Q water
- FastStart™ Taq DNA Polymerase (Roche, 12032902001)
- 10x PCR buffer + MgCl₂ (supplied with faststart taq polymerase)
- dNTPs (Invitrogen, 10297-018) (stock: 10mM in 10mM Tris pH 8.5 for each nucleotide)
- BigDye™ Terminator v3.1 Cycle Sequencing Kit (Thermo Scientific, 4337458)
- Exosap
- 5x sequencing buffer
- BigDye® Terminator v3.1 (BDT)

Procedure:

1. Dilute the DNA samples to the required DNA concentration using Milli-Q water.
2. Perform a PCR reaction as described below. Use a DNA sample from the unedited cell as a negative control.

| PCR reaction mix for 1 reaction | | PCR Program | |
|---------------------------------|------------------------------------|-------------|-----------------------|
| 1.5μl | 10x PCR buffer + MgCl ₂ | 1) | 96°C 4:00 |
| 0.5μl | dNTPs (10mM) | 2) | 96°C 0:20 |
| 0.5μl | Forward primer (10μM) | 3) | 55-65°C 0:30 |
| 0.5μl | Reverse primer (10μM) | 4) | 72°C 1:00 |
| 1μl | DNA (<25ng) | 5) | Go to step 2 34x |
| 0.5μl | FastStart Taq | 6) | 72°C 5:00 |
| 10.5μl | Milli-Q water | 7) | 10°C ∞ |
| 15μl | Total volume | 8) | End |

3. Add 1μl exosap to the PCR product and perform the exosap reaction as described below.

| Exosap Program: | |
|-----------------|------------|
| 1) | 37°C 45:00 |
| 2) | 80°C 15:00 |
| 3) | 10°C ∞ |
| 4) | End |

4. Use 4μl of the product to perform a BDT reaction as described below.

| BDT reaction mix for 1 reaction | | BDT reaction Program | |
|---------------------------------|------------------------|----------------------|----------------------|
| 3.5µl | 5x sequence buffer | 1) | 96°C 0:45 |
| 1µl | Sequence primer (10µM) | 2) | 96°C 0:10 |
| 4µl | PCR product | 3) | 58°C 0:05 |
| 0.5µl | BDT | 4) | 60°C 3:00 |
| 1µl | Milli-Q water | 5) | Goto step 2 25x |
| 10µl | Total volume | 6) | 10°C ∞ |
| | | 7) | End |

5. Perform Sanger sequencing

Results:

A successful CRISPR/Cas9 reaction will result in the appearance of heterozygous sequence calls starting around the target sequence. Note that also in the case of both alleles being targeted, the this will still appear as heterozygous callings as the indels introduced will likely differ between the two alleles.

Genotyping method 2: Introduction of large deletions

Two PCR reactions with different sets of primers are used to genotype the introduction of large deletions. The product size for set 1 will decrease if the large deletion is successful. A long-range PCR protocol could be necessary for the reaction depending on the size of the deletion. A reaction with primer set 2 will not result in a product if the deletion is successful as the forward primer is in the deleted region. The combined results will provide information on the targeting of both alleles and will exclude false positive results (Figure 10).

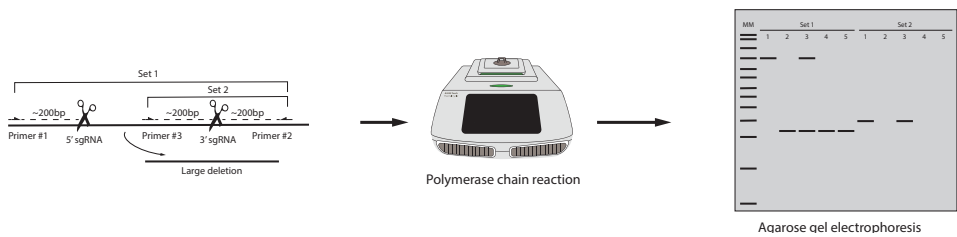


Figure 10. Genotyping of large deletions. In primer set 1, PCR primers flank the desired deletion to amplify that region. In primer set 2, the forward primer is located inside the desired deletion to detect unedited alleles. A typical result is shown: sample #1 = no deletion, sample #3 = mono-allelic deletion, sample # 2, 4 and 5 = bi-allelic deletion.

Materials:

- Primer 1: forward primer located ~200bp upstream the 5' sgRNA (stock: 10 μ M in 10mM Tris)
- Primer 2: reverse primer located ~200bp downstream the 3' sgRNA (stock: 10 μ M in 10mM Tris)
- Primer 3: forward primer located ~200bp upstream the 3' sgRNA (stock: 10 μ M in 10mM Tris)
- Isolated DNA (isolated previously from individual iPSC colonies)
- Milli-Q water
- FastStart™ Taq DNA Polymerase (Roche, 12032902001)
- 10x PCR buffer + MgCl₂ (supplied with faststart taq polymerase)
- dNTPs (initrogen, 10297-018) (stock: 10mM in 10mM Tris pH8.5 for each nucleotide)
- 1.5% agarose TAE gel
 - 1x TAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA)
 - Agarose (Sigma, A9539)

Procedure:

1. Dilute the DNA samples to the required DNA concentration using Milli-Q water
2. Perform two PCR reactions as listed below: one using set 1 (primers 1 & 2) and one using set 2 (primers 2 & 3).

| PCR reaction mix for 1 reaction | | PCR Program | |
|---------------------------------|------------------------------------|-------------|----------------------|
| 1.5 μ l | 10x PCR buffer + MgCl ₂ | 1) | 96°C 4:00 |
| 0.5 μ l | dNTPs (10mM) | 2) | 96°C 0:20 |
| 0.5 μ l | Forward primer (10 μ M) | 3) | 55-65°C 0:30 |
| 0.5 μ l | Reverse primer (10 μ M) | 4) | 72°C (1:00 per 1 kb) |
| 1 μ l | DNA (<25ng) | 5) | Go to step 2 34x |
| 0.5 μ l | FastStart Taq | 6) | 72°C 5:00 |
| 10.5 μ l | Milli-Q water | 7) | 10°C ∞ |
| 15 μ l | Total volume | 8) | End |

3. Run and visualize both PCR reactions on a 1.5% agarose TAE gel.

Results:

A large product size for set 1 indicates the presence of the full-sized product, a small product size indicates the introduction of a large deletion. The absence of the large product indicates successful targeting on both alleles. The presence of both products indicates successful targeting of one allele.

Set 2 will verify the results, as a product will only be present if the original DNA sequence is present on one or both alleles.

| | No deletion | Mono-allelic deletion | Bi-allelic deletion |
|--------------|--------------------|-------------------------|---------------------|
| Primer set 1 | Only large product | Large and small product | Only small product |
| Primer set 2 | Product present | Product present | Product absent |

Genotyping method 3: Knock-in with donor template integration

The genotyping of a template integration can be performed using two primer sets. Set 1 uses primers flanking the integration. Set 2 uses a forward primer upstream from the 5' homology arm, the reverse primer is only present in the insert. Set 1 provides information on the presence of the integration, whereas set 2 indicates if the integration is at the desired location. The combined results will provide information on the taractina of both alleles (Figure 11).

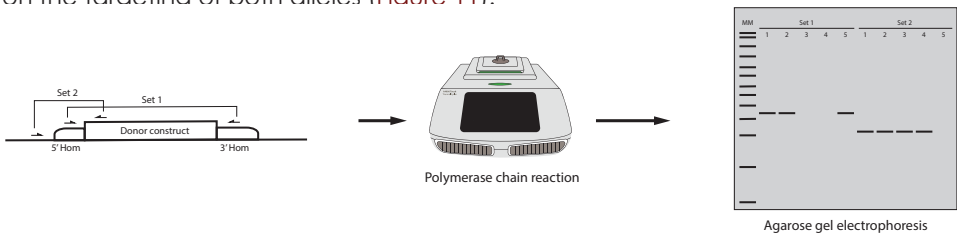


Figure 11. Genotyping of donor template integration. In primer set 1, PCR primers are located within the 5' and 3' homology arms, and under the PCR conditions employed will only amplify the DNA if the donor is not integrated. In primer set 2, PCR primers are located at a 5' upstream location and within the donor construct, and will only result in a correct PCR product if the template has been integrated at the desired location. A typical result is shown: sample #1 and #2 = mono-allelic integration, sample #3 and #4 bi-allelic integration, sample #5 no integration.

Materials:

- Primer set 1: (located in the homology arms)
 - Forward primer, located ~200bp upstream the target site (stock: 10μM in 10mM Tris)
 - Reverse primer, located ~200bp downstream the target site (stock: 10μM in 10mM Tris)
- Primer set 2:
 - Forward primer, located ~200bp upstream 5' homology arm (stock: 10μM in 10mM Tris)
 - Reverse primer, located in the insert, downstream the homology arm (stock: 10μM in 10mM Tris)
- Isolated DNA (isolated previously from individual iPSC colonies)

- Milli-Q water
- FastStart™ Taq DNA Polymerase (Roche, 12032902001)
- 10x PCR buffer + MgCl₂ (supplied with faststart™ taq DNA polymerase)
- dNTPs (Invitrogen, 10297-018) (stock: 10mM in 10mM Tris pH8.5 for each nucleotide)
- 1.5% agarose TAE gel
 - 1x TAE buffer (40mM Tris, 20mM acetic acid, 1mM EDTA)
 - Agarose (Sigma, A9539)

Procedure:

1. Dilute the DNA samples to the required DNA concentration using Milli-Q water
2. Perform two PCR reactions as listed below: one using set 1 and one using set 2

| PCR reaction mix for 1 reaction | | PCR Program | |
|---------------------------------|------------------------------------|-------------|-----------------------|
| 1.5μl | 10x PCR buffer + MgCl ₂ | 1) | 96°C 4:00 |
| 0.5μl | dNTPs (10mM) | 2) | 96°C 0:20 |
| 0.5μl | Forward primer (10μM) | 3) | 55-65°C 0:30 |
| 0.5μl | Reverse primer (10μM) | 4) | 72°C 1:00 |
| 1μl | DNA (<25ng) | 5) | Go to step 2 34x |
| 0.5μl | FastStart Taq | 6) | 72°C 5:00 |
| 10.5μl | Milli-Q water | 7) | 10°C ∞ |
| 15μl | Total volume | 8) | End |

3. Run and visualize both PCR reactions on a 1.5% agarose TAE gel.

Results:

Set 1 only results in a product in the absence of the integration on one or both alleles. Upon successful integration, the product becomes too large to amplify under given PCR conditions. Absence of a product for set 1 indicates successful bi-allelic integration. A product for set 2 indicates the integration of the donor template at the desired location, since the primer design ensures that a PCR product can only be formed if the correct integration has occurred.

| | No integration | Mono-allelic integration | Bi-allelic integration |
|--------------|-----------------|--------------------------|------------------------|
| Primer set 1 | Product present | Product present | Product absent |
| Primer set 2 | Product absent | Product present | Product present |

10. Protocol adjustments for feeder free iPSCs

Recently, iPSC culture protocols have been developed that do not require the presence of MEFs but use adjusted cell culture media. The use of these protocols provide iPSCs with a more stable environment and results in a higher rate of proliferation compared to feeder-dependent cultures. It is possible to use this CRISPR/Cas9 strategy for iPSCs in feeder free conditions with a few minor adaptations to the protocol.

1. Generation of conditioned media:

This section of the protocol does not need to be performed.

2. Plating frozen MEFs:

This section of the protocol does not need to be performed.

3. Nucleofection of iPSCs:

This section of the protocol has some minor adjustments with regard to the cell culture conditions of the iPSCs without MEFs.

Materials:

- MTSER PLUS medium (Stem Cell Technologies, 05825)
- Penicillin/streptomycin (P/S) (Gibco, 15140122)
- Nucleofector™ 2b Device (Lonza, AAB-1001)
- Human Stem Cell Nucleofector™ Kit 2 (Lonza, VAPH-5022)
- DNA prep (prepared previously)
- Vitronectin XF (Stem cell Technologies, 07180)
- CellAdhere™ Dilution Buffer (Stem cell Technologies, 07183)
- 6-well suspension plates (Greiner bio-one, 657185)
- Revitacell Supplement 100X (Gibco, A2644501)
- Accutase (Gibco, A11105-01) or TrypLE (Gibco, 12605010)
- PBS (Gibco, 70011044)

Procedure:

1. Four hours before starting the nucleofection procedure: replace the medium on the iPSCs with mTSEr PLUS medium supplemented with 1% P/S and 1x Revitacell Supplement.
2. One hour before starting the nucleofection procedure: for each nucleofection reaction, coat two wells of a 6-well suspension plates with Vitronectin XF (1:25 diluted in CellAdhere™ Dilution Buffer) and incubate for

one hour at room temperature. After incubation, remove the coating and add 2mL mTSER PLUS medium to each well.

Perform steps 3 – 15 as stated in section 7, continue the protocol as follows:

16. Seed the iPSCs on the Vitronectin XF coated 6-well (2/3 and 1/3 of the cell suspension respectively).
17. 23 hours after nucleofection: coat one full 6-well suspension culture plate with Vitronectin XF (one plate for each nucleofection reaction).
18. 24 hours after nucleofection: from one well, detach the iPSCs from the plate using 500µl TrypLE (preferably the well with the highest confluency, unless the confluency exceeds 80%).
19. Centrifuge the cell suspension at 1000rpm for 5 minutes.
20. Resuspend the iPSCs in 1mL mTSER PLUS medium supplemented with 1% P/S and 1x Revitacell Supplement.
21. Using a P1000 pipette, reseed the iPSCs to a confluency of 5% in the first well, and continue diluting the iPSCs by a factor 2 for each following well in order to be able to obtain single cells colonies.
22. 48 hours after transfection: replace the medium on the iPSCs with mTSER PLUS medium supplemented with 1% P/S and refresh daily. Selection with G418 can be started now if a donor template is used.
23. 6 to 14 days after nucleofection: single colonies can be picked, depending on the cell density, recovery speed of the cells, and use of selection.

Note: a new kill curve may be required to determine the optimal G418 concentration in feeder free iPSCs.

4. Picking colonies:

This section of the protocol follows a different procedure.

Materials:

- mTSER PLUS medium (Stem Cell Technologies, 05825)
- Penicillin/streptomycin (P/S) (Gibco, 15140122)
- Vitronectin XF (Stem cell Technologies, 07180)
- CellAdhere™ Dilution Buffer (Stem cell Technologies, 07183)
- 48-well suspension plates (Greiner bio-one, 677102)

Procedure:

1. One hour before picking: for each nucleofection reaction, coat two full

48-well suspension plates with Vitronectin XF (1:25 diluted in CellAdhere™ Dilution Buffer) and incubate for one hour at room temperature. One plate will be used for DNA isolation and one plate for passaging.

2. Before picking colonies: prepare 1 x 48 wells plate by remove the coating from the wells and add 500µl mTSER PLUS medium supplemented with 1% P/S and 1x Revitacell Supplement to all wells of the plate.
3. Rinse the iPSCs with 1mL PBS and add 2mL mTSER PLUS medium supplemented with 1% P/S.
4. Using P1000 pipette, gently scrape the selected colony with the pipet tip, once partly detached use the P1000 pipette to transfer the iPSCs to a well containing 500µl medium.
5. Repeat until all the selected single colonies are picked.
6. Remove the coating from the wells from the second plate.
7. Dissociate the picked colonies by pipetting up and down using a P1000 pipette. Transfer 250µl of the suspended cells to a well on the second plate and leave the remaining 250µl in the original plate. Culture both plates at 37°C/5% CO₂.
8. Refresh the media and monitor the colonies daily until passaging or harvesting DNA for genotyping.

11. Troubleshooting

| Problem | Possible cause and suggestions |
|--|---|
| Low number of colonies after nucleofection | <i>Number of cells used in the nucleofection reaction was too low</i> <ul style="list-style-type: none">• Increase the number of cells used for nucleofection |
| | <i>Cells died during the nucleofection</i> <ul style="list-style-type: none">• Pretreat the iPSCs with Revitacell Supplement• Use conditioned media without antibiotics• Be gentle during the handling of single cell iPSCs• Reduce the amount of time that the cells spend in the Human Stem Cell Nucleofector mix to a minimum |

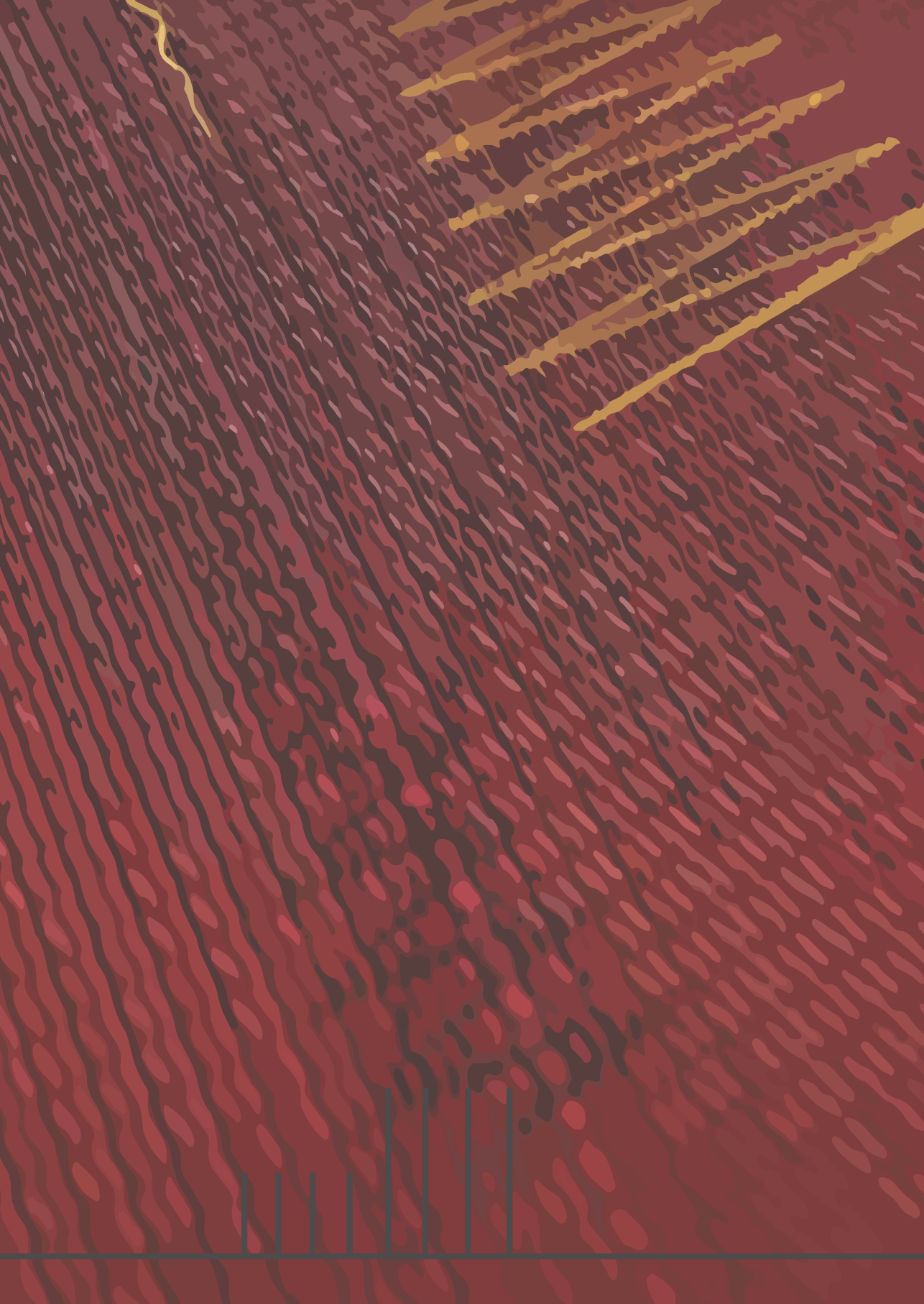
| | |
|---|--|
| Only untargeted colonies | <i>Poor plasmid quality</i> <ul style="list-style-type: none"> • Check plasmid integrity |
| | <i>Human Stem Cell Nucleofector mix is expired</i> <ul style="list-style-type: none"> • Use freshly made Human Stem Cell Nucleofector mix |
| | <i>sgRNA target not present</i> <ul style="list-style-type: none"> • Sequence the target site for variants, design multiple sgRNA per targeting |
| | <i>Targeting of the gene is lethal to the cells</i> |
| | <i>Bad sgRNA design, too many off-target events after nucleofection</i> <ul style="list-style-type: none"> • Use in silico prediction algorithms to predict off-target effects |
| Poor morphology of colonies after nucleofection | <i>Poor morphology of cells before nucleofection</i> <ul style="list-style-type: none"> • Use iPSC with a low passage number • Remove all differentiation before nucleofection |
| | <i>Differentiation after nucleofection</i> <ul style="list-style-type: none"> • Add fresh bFGF to the conditioned media • Refresh the iPSC media daily |
| No single cell colonies | <i>Plating density after nucleofection too high</i> <ul style="list-style-type: none"> • Decrease plating density after nucleofection • Use multiple plating densities |
| | <i>No single cell passaging during nucleofection</i> <ul style="list-style-type: none"> • Ensure iPSCs are single cells before nucleofection |
| Only mono allelic targeted cells | <i>Bi allelic targeting is lethal to the cells</i> |
| No colonies after G418 selection | <i>G418 concentration too high</i> <ul style="list-style-type: none"> • Perform a kill curve to determine the optimal G418 concentration |
| High number of negative colonies after G418 selection | <i>G418 concentration too low</i> <ul style="list-style-type: none"> • Perform a kill curve to determine the optimal G418 concentration |

| | |
|--|---|
| MEF viability decreases during selection with G418 | MEFs not resistant to G418 |
| | <ul style="list-style-type: none">• Use G418 resistant MEFs |
| | Optimal G418 concentration for iPSCs is too high for MEFs <ul style="list-style-type: none">• Add fresh MEFs to the well when the density of MEFs drops below 50% |

12. References

1. Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors*. Cell, 2006. **126**(4): p. 663-76.
2. Liu, G., et al., *Advances in Pluripotent Stem Cells: History, Mechanisms, Technologies, and Applications*. Stem Cell Rev Rep, 2020. **16**(1): p. 3-32.
3. Ernst, M.P.T., et al., *Ready for Repair? Gene Editing Enters the Clinic for the Treatment of Human Disease*. Mol Ther Methods Clin Dev, 2020. **18**: p. 532-557.
4. Malik, N. and M.S. Rao, *A review of the methods for human iPSC derivation*. Methods Mol Biol, 2013. **997**: p. 23-33.
5. Shi, Y., et al., *Induced pluripotent stem cell technology: a decade of progress*. Nat Rev Drug Discov, 2017. **16**(2): p. 115-130.
6. Raab, S., et al., *A Comparative View on Human Somatic Cell Sources for iPSC Generation*. Stem Cells Int, 2014. **2014**: p. 768391.
7. Singh, V.K., et al., *Induced pluripotent stem cells: applications in regenerative medicine, disease modeling, and drug discovery*. Front Cell Dev Biol, 2015. **3**: p. 2.
8. Hockemeyer, D. and R. Jaenisch, *Induced Pluripotent Stem Cells Meet Genome Editing*. Cell Stem Cell, 2016. **18**(5): p. 573-86.
9. Hotta, A. and S. Yamanaka, *From Genomics to Gene Therapy: Induced Pluripotent Stem Cells Meet Genome Editing*. Annu Rev Genet, 2015. **49**: p. 47-70.
10. Broeders, M., et al., *Sharpening the Molecular Scissors: Advances in Gene-Editing Technology*. iScience, 2019. **23**(1): p. 100789.
11. Gaj, T., C.A. Gersbach, and C.F. Barbas, 3rd, *ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering*. Trends Biotechnol, 2013. **31**(7): p. 397-405.
12. Bock, C., et al., *Reference Maps of human ES and iPSC cell variation enable high-throughput characterization of pluripotent cell lines*. Cell, 2011. **144**(3): p. 439-52.
13. Boulting, G.L., et al., *A functionally characterized test set of human induced pluripotent stem cells*. Nat Biotechnol, 2011. **29**(3): p. 279-86.
14. Jang, Y., et al., *Development of immunocompatible pluripotent stem cells via CRISPR-based human leukocyte antigen engineering*. Exp Mol Med, 2019. **51**(1): p. 1-11.

15. Listgarten, J., et al., *Prediction of off-target activities for the end-to-end design of CRISPR guide RNAs*. Nat Biomed Eng, 2018. **2**(1): p. 38-47.
16. van der Wal, E., et al., *Large-Scale Expansion of Human iPSC-Derived Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies*. Stem Cell Reports, 2018. **10**(6): p. 1975-1990.
17. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system*. Nat Protoc, 2013. **8**(11): p. 2281-2308.
18. Anastassiadis, K., et al., *A practical summary of site-specific recombination, conditional mutagenesis, and tamoxifen induction of CreERT2*. Methods Enzymol, 2010. **477**: p. 109-23.
19. Kim, D., et al., *Evaluating and Enhancing Target Specificity of Gene-Editing Nucleases and Deaminases*. Annu Rev Biochem, 2019. **88**: p. 191-220.
20. Pattanayak, V., J.P. Guilinger, and D.R. Liu, *Determining the specificities of TALENs, Cas9, and other genome-editing enzymes*. Methods Enzymol, 2014. **546**: p. 47-78.
21. Manghwar, H., et al., *CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects*. Adv Sci (Weinh), 2020. **7**(6): p. 1902312.



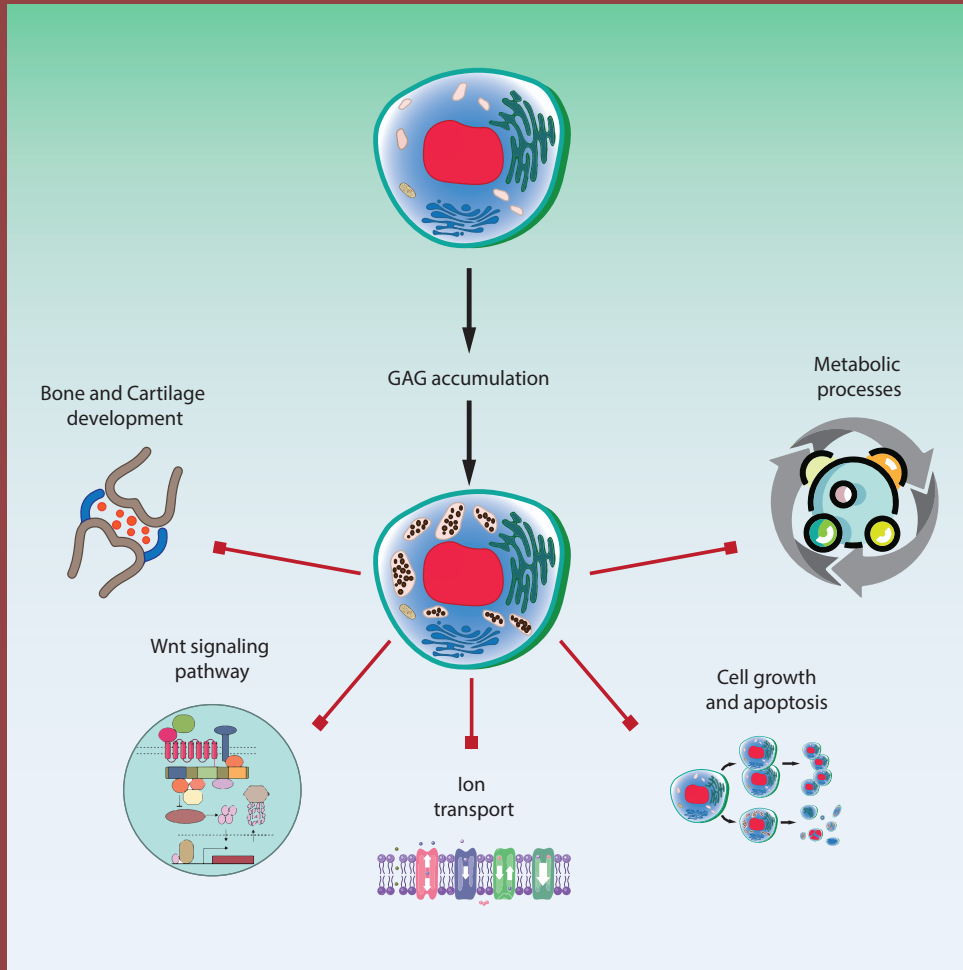
CHAPTER 8

GENERATION OF A DISEASE MODEL FOR CARTILAGE PATHOLOGY IN MPS VI USING PATIENT-DERIVED AND ISOGENIC GENE-CORRECTED HIPSCS

Mike Broeders, Jeroen G.J. van Rooij, Esmee Oussoren,
Tom J.M. van Gestel, Christopher A. Smith, Susan J. Kimber, Rob M. Verdijk,
Margreet A.E.M. Wagenmakers, Hannerieke J.M.P. van den Hout,
Ans T. van der Ploeg, Roberto Narcisi, W.W.M. Pim Pijnappel

Submitted

Graphical abstract



keywords: CRISPR/Cas9; iPSC, nucleofection; gene editing strategy; targeting plasmid; donor template integration; protocol

Generation of a disease model for cartilage pathology in MPS VI using patient-derived and isogenic gene-corrected hiPSCs

M Broeders¹, JGJ van Rooij², E Oussoren³, TJM van Gestel¹, CA Smith⁴, SJ Kimber⁴, RM Verdijk⁵, MAEM Wagenmakers⁶, JMP van den Hout², AT van der Ploeg², R Narcisi⁷, WWM Pijnappel^{1*}

- 1 Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, Netherlands; Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, Netherlands; Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands.
- 2 Department of Internal Medicine, Erasmus Medical Center, 3015 GE Rotterdam, Netherlands
- 3 Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, Netherlands; Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands.
- 4 Division of Cell Matrix Biology and Regenerative Medicine, School of Biological Sciences, Faculty of Biology Medicine and Health, University of Manchester, Manchester M13 9PT, U.K.
- 5 Department of Pathology, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands
- 6 Department of Internal Medicine, Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, 3015 GE Rotterdam, The Netherlands.
- 7 Department of Orthopaedics and Sports Medicine, Erasmus MC University Medical Center, 3015 GE Rotterdam, Netherlands.

* **Correspondence:** W.W.M. Pim Pijnappel Department of Pediatrics, Erasmus University Medical Center, Rotterdam, the Netherlands.

e-mail: w.pijnappel@erasmusmc.nl

Abstract

Mucopolysaccharidosis type VI (MPS VI) is a metabolic disorder caused by disease-associated variants in the Arylsulfatase B (ARSB) gene, resulting in ARSB enzyme deficiency, lysosomal glycosaminoglycan accumulation, and cartilage and bone pathology. We generated a disease model for cartilage pathology in MPS VI using patient-derived human induced pluripotent stem cells (hiPSCs), and we generated isogenic controls by inserting the ARSB cDNA in the AAVS1 safe harbor locus using CRISPR/Cas9. Using an optimized chondrogenic differentiation protocol, we found Periodic acid–Schiff positive inclusions in chondrogenic cells with MPS VI. Genome-wide mRNA expression analysis showed deregulated expression in MPS VI of genes involved in bone and cartilage development, Wnt signaling, cell growth and apoptosis, metabolism, and ion transport. This model provides insight in the early stages of cartilage pathology in patients with MPS VI, and demonstrates the power of using isogenic controls to correct for differences in genetic background among humans.

Introduction

Mucopolysaccharidoses type VI (MPS VI) is an autosomal recessive disorder caused by Arylsulfatase B (ARSB) enzyme deficiency leading to intralysosomal accumulation of the glycosaminoglycans (GAGs) dermatan sulfate (DS) and chondroitin sulfate (CS). MPS VI is a multisystemic disease with GAG accumulation in connective tissues and organs which ultimately leads to a cascade of symptoms such as corneal clouding, hepatosplenomegaly and bone and cartilage pathology. In all MPS VI patients the hips are frequently and severely affected resulting in limitations in mobility and pain with impact on quality of life [1].

Currently, treatment of MPS VI consists of enzyme replacement therapy (ERT) with intravenous administration of recombinant ARSB enzyme. Although ERT attenuates disease progression and resolves hepatosplenomegaly, the therapeutic effect on bone and cartilage pathology is limited. Cartilage in particular is nonresponsive to ERT, likely due to the poor vascularization of cartilage resulting in impaired delivery of the enzyme from the circulation to chondrocytes.

The development of cartilage pathology in patients with MPS VI remains poorly understood. Several studies in animal models for MPS VI and related types of MPS have been performed, and have provided insight in the cascade of events resulting in cartilage pathology. This has shown that chondrocytes in MPS VI are abnormal with a swollen appearance and containing membrane-bound inclusions/

vacuoles [2]. In the growth plate, there is loss of columnar structure and excess of calcified cartilage [3]. In newborn mice abnormal anlagen of tracheal and articular cartilage was reported, indicating that cartilage pathology can already start during development [4]. Studies from human cartilage biopsies of MPS I, II and III patients showed severe cartilage damage and abnormal chondrocytes [5, 6], but limited information is available on the development of cartilage pathology in human patients with MPS VI.

At the molecular level, primary lysosomal accumulation of GAGs can result in secondary accumulation of cholesterol and the gangliosides GM2 and GM3, indicating cross talk between metabolic pathways in MPS and other lysosomal diseases [7]. GAGs can also function as receptor ligands to activate Toll like receptors, resulting in inflammation. In addition GAGs can also activate growth factors such as BMPs, involved in cartilage and bone development [5]. In mice with the related disorder MPS VII, the growth plate was enlarged, disorganized, and contained fewer chondrocytes. It was suggested that chondrocytes displayed a delayed exit from G1 into M and a delayed terminal differentiation, possibly mediated by elevated expression of PTHrP and Wnt5a [8]. How these processes operate in human chondrocytes with MPS VI remains largely unknown.

In humans, the high variability of genetic background between individuals complicates the downstream analysis of disease progression [9, 10]. It has become clear that differences in gene expression and functional parameters can be considerable between individuals, which introduces a large amount of (seemingly random) variation to the analyses. Recent advancements in gene editing using CRISPR/Cas9 enable the generation of isogenic controls to reduce these effects and to decipher the robust molecular mechanisms of disease by generating isogenic controls, i.e. diseased and healthy versions with the same genetic background. In contrast to mesenchymal stem cells [11], hiPSCs have high capacity for self-renewal and are suitable for gene editing, which makes them more suitable for disease modeling of cartilage pathology.

Here, we generated isogenic pairs from four patient-derived hiPSC lines using CRISPR-Cas9 to generate a model for the cartilage pathology in MPS VI. We improved the protocol for chondrogenic differentiation of hiPSCs, and applied this to characterize the pathology and molecular changes in human chondrogenic cells with MPS VI.

Results

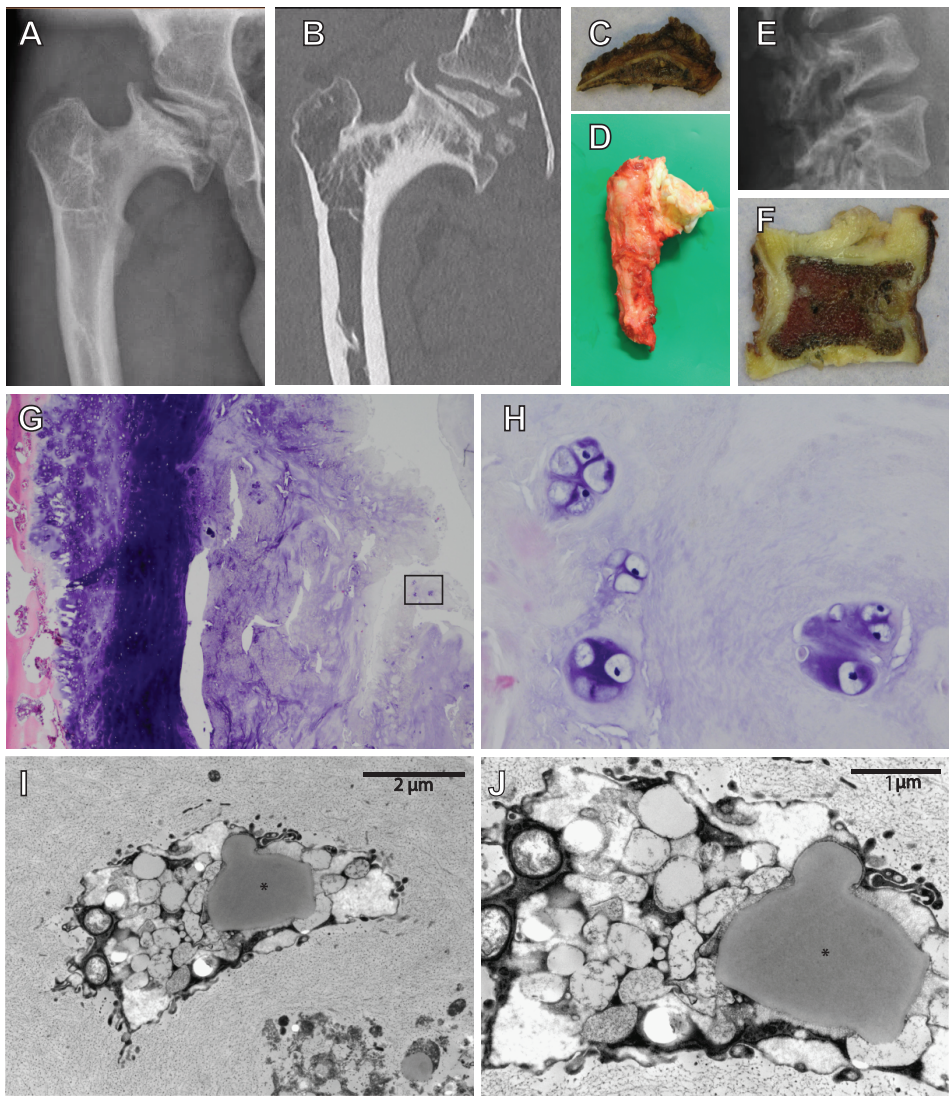
Cartilage pathology in a MPS VI Patient

To study cartilage pathology in MPS VI, a femoral head and acetabulum were conducted by autopsy of a 25-year-old MPS VI patient that died from disease-related symptoms. X-ray and CT analysis of the hip showed severe osteoarthritis of the femoral head with total destruction and abnormal ossification, and with an abnormal steep and dysplastic acetabulum, which is also shown in the macroscopic pictures (**Figure 1A-D**). The vertebrae were abnormal shaped, with an abnormal degenerative cartilage endplate (**Figure 1E-F**). Chondrocytes within the vertebrae cartilage showed multiple proliferative clones, which reflects the pathological response, and vacuolar changes in HE stained sections (**Figure 1G-H**). Electron microscopy analysis of the same cartilage showed examples of chondrocytes with different types of vacuoles and in addition lipid storage (**Figure 1I-J**). We conclude that this patient displayed severe degeneration of cartilage and that chondrocytes showed very severe vacuolization and lipid storage.

Generation of patient-derived hiPSCs and gene editing to generate isogenic controls

To model cartilage pathology in human MPS VI, we selected four patients with a rapid disease progression. Diagnosis was based on ARSB enzyme activity, accumulation of urinary GAGs (uGAGs), the presence of disease-associated variants in the ARSB gene, and clinical symptoms (**Figure 1K**). Patient #1, #2 and #3 were homozygous for ARSB c.1142+2T>C, c.995T>G and c.937C>T, respectively. No disease-associated variant for patient #4 had been identified, but we recently confirmed the molecular diagnosis at the RNA level [12]. hiPSC lines were generated from primary fibroblasts using lentiviral expression of Oct4, Sox2, Klf-4, and C-Myc as described [13].

To restore expression of ARSB, a healthy copy was introduced into a safe location of the genome, also known as a safe harbor. We chose the widely used *PPP1R12C* gene in the *AAVS1* locus located on chromosome 19 (**Figure 2A**). A donor construct previously described by us [13, 14] was adapted to introduce a healthy copy of the ARSB gene using CRISPR/cas9 via the homology directed repair (HDR) pathway. This construct (**Figure 2B**) contained the ARSB cDNA driven by the constitutively active EF1a promoter, and the neomycin resistance gene to enable selection for successful integrations, which is required due to the highly inefficient HDR pathway. To validate ARSB expression by the donor construct, HeLa TK- cells were transfected and selection was started 24 hours after transfection. After selection, the cells showed an ARSB enzyme activity of ~1400 nmol/mg/hr, which was above the range seen in healthy control fibroblasts (**Figure S1**). No ARSB enzyme activity



| K | | Allele 1 | | | | Allele 2 | | | | Type of disease progression |
|-----------|--------|-------------------------------|----------|------------------------------|-------------------------------|----------|------------------------------|-------|--|-----------------------------|
| | | ΔDNA (cDNA HGVS nomenclature) | Location | Δprotein (HGVS nomenclature) | ΔDNA (cDNA HGVS nomenclature) | Location | Δprotein (HGVS nomenclature) | | | |
| Patient # | Line # | | | | | | | | | |
| 1 | 1 | c.1142+2T>C | Intron 5 | p.? | c.1142+2T>C | Intron 5 | p.? | Rapid | | |
| 2 | 2 | c.995T>G | Exon 5 | p.V332G | c.995T>G | Exon 5 | p.V332G | Rapid | | |
| 3 | 3 | c.937C>T | Exon 5 | p.P313S | c.937C>T | Exon 5 | p.P313S | Rapid | | |
| 4 | 4 | * | | | * | | | Rapid | | |

*Diagnosis confirmed at RNA level (r.691_898del) and enzyme deficiency

Legend on next page

was detected in HeLa TK- cells transfected with a pCAG-NEO control vector. This confirmed that the *ARSB* donor construct was functional.

To generate a control, a gene targeting procedure that did not change *ARSB* expression, we deleted the *ARSB* cDNA from the targeting vector (**Figure 2C**). hiPSCs were gene edited using either the *ARSB* cDNA vector (GE) or the empty targeting vector (EV) by co-transfection with an expression vector for human codon-optimized Cas9 nuclease, an expression vector for the guide RNA targeting the AAVS1 locus, and either targeting vector. After selection with G418, colonies were picked and genotyped using two PCR strategies; set 1 yields a product of 749 bp from the untargeted allele and 7132 bp from the targeted allele. This 7132 bp product is too large to be detected under the PCR conditions employed. Set 2 only yields a product of 980 bp if the *ARSB* cDNA is inserted in the targeted site (**Figure 2D**). A typical result is shown in **Figure 2E**, where the results of set 1 show that 8 out of 15 colonies still have the endogenous sequence at target site on one or both alleles. The results of set 2 show that 12 out of 15 colonies have the *ARSB* insertion at the targeted site. From these results we conclude that 6 colonies have both alleles targeted, 6 colonies have 1 allele targeted and 3 colonies have no insertion of the *ARSB* cDNA at the target site. The gene editing results of all the targeted patient-derived hiPSCs are shown in **Figure 2F**, with a mono-allelic insertion ranging from 40-80% and a bi-allelic insertion ranging from 6-40%, depending on the hiPSC line. Quantitative RNA expression analysis showed that *ARSB* mRNA expression after gene editing was 4-5-fold above the average expression in healthy controls (**Figure 2G**).

Differentiation of hiPSC to chondrogenic cells

Gene edited hiPSCs (GE and EV) were differentiated into chondrogenic cells using a 14-day differentiation protocol that was modified from Oldershaw et al. [15]. This protocol uses a chemically defined culture media and matrix-coated substrates, and is based on the sequential activation of signaling pathways that operate during development of cartilage (**Figure 3A, Table S1**). With this modified protocol we achieved successful chondrogenic differentiation in every hiPSC tested (data not shown). Genome-wide mRNA expression analysis was performed

Figure 1. Cartilage pathology in a 25 year old diseased MPS VI patient. A: X-ray and B: CT scan of the right hip. Macroscopic pictures from autopsy of the C: right acetabulum and D: femoral head. E: X-ray of the lumbar vertebrae. F: A macroscopic picture from autopsy of one vertebra. G: HE stain of cartilage from this vertebra (Magnification 100x), H: details from the inset in G (Magnification 400x). I, J: Electron microscopy analysis showing examples of chondrocytes with different types of vacuoles and in addition lipid storage (indicated with *). K: Characteristics of included patients.

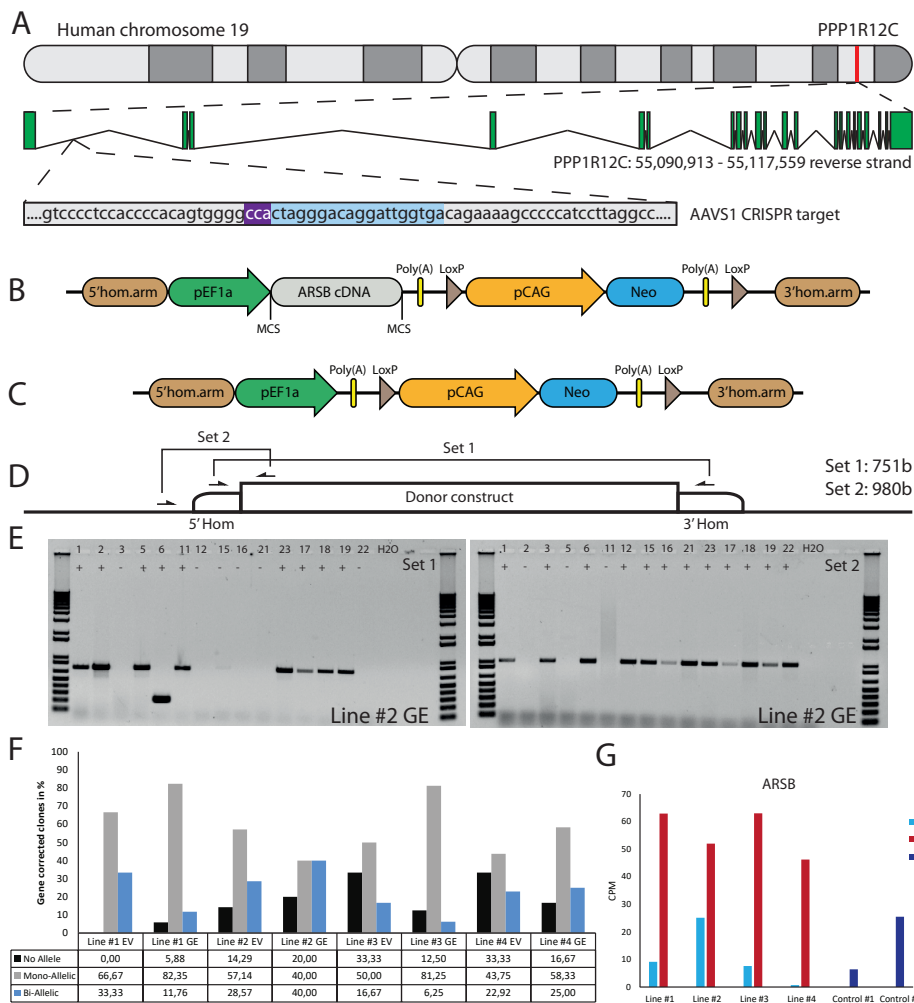


Figure 2. Gene editing in hiPSCs. A: The gene editing strategy was designed to express ARSB from the AAVS1 locus. PAM sequence indicated in purple, gRNA target indicated in blue. B: The donor construct generated for insertion of the ARSB cDNA in the AAVS1 safe harbor by CRISPR/Cas9-mediated gene editing. The neomycin cassette enables G418 selection of targeted colonies. C: The empty vector construct used to generate isogenic controls. D: Strategy for the PCR based genotyping of targeted colonies, primer set 1 spans the insertion site and only gives a product in the absence of targeting, primer set 2 amplifies a product only in the presence of the construct at the target site. E: Typical genotyping result of picked colonies. With primer set 2, 12/15 colonies show successful mono-allelic or bi-allelic targeting. With primer set 1, 6/15 lack a PCR product, indicating a loss of the endogenous sequence. In combination colonies 3, 12, 15, 16, 21 and 22 show bi-allelic insertion of the construct. F: Quantification of mono-allelic, bi-allelic and unsuccessfully targeted colonies. G: ARSB mRNA expression in the selected bi-allelic targeted clones, determined using RNAseq. GE: Gene edited with ARSB cDNA, EV: Gene edited with empty vector control.

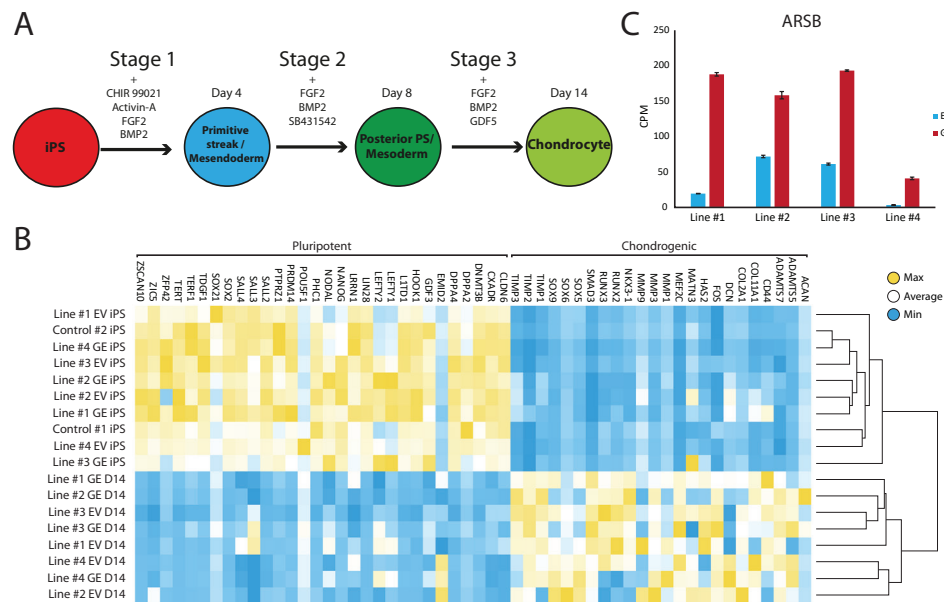


Figure 3. Differentiation of hiPSCs into chondrogenic cells. A: Schematic representation of the refined 14-day differentiation protocol. Different factors are added to the media during the three stages. Cells are passaged at the end of stage 1 and 2 at day 4 and 8, respectively. B: Expression analysis of pluripotent and chondrogenic genes show a loss of pluripotency and an increase in expression of chondrogenic genes after the chondrogenic differentiation. C: ARSB gene expression remained high after chondrogenic differentiation ($n=3$). GE: Gene edited with ARSB cDNA, EV: Gene edited with empty vector control.

using RNA sequencing to compare expression of genes involve in pluripotency and chondrogenic differentiation between hiPSCs and chondrogenic cells. High expression of pluripotency genes was observed in hiPSC cells, while these genes were not expressed at day 14 of differentiation. Conversely, chondrogenic genes were not expressed at the hiPSC stage but were upregulated at day 14 of differentiation (**Figure 3B**). Key chondrogenic genes that were expressed after 14 days of differentiation included *COL2A1*, and *SOX9*. ARSB RNA expression in chondrogenic cells was 2 to 12-fold increased in the GE compared to the EV cells (**Figure 3C**).

Chondrogenic cells from MPS VI patients accumulate GAGs

To assess whether hiPSC-derived chondrogenic cells accumulated carbohydrate macromolecules, we analyzed cytopspins of GE and EV gene edited patient-derived cells using periodic acid–Schiff (PAS) staining (**Figure 4**). In EV cells, PAS-positive extranuclear areas were observed in the form of small to larger dots that in some

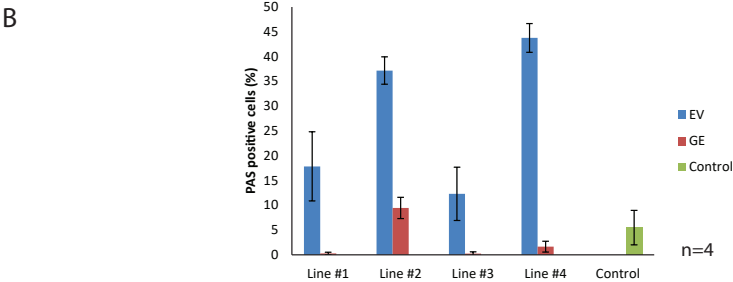
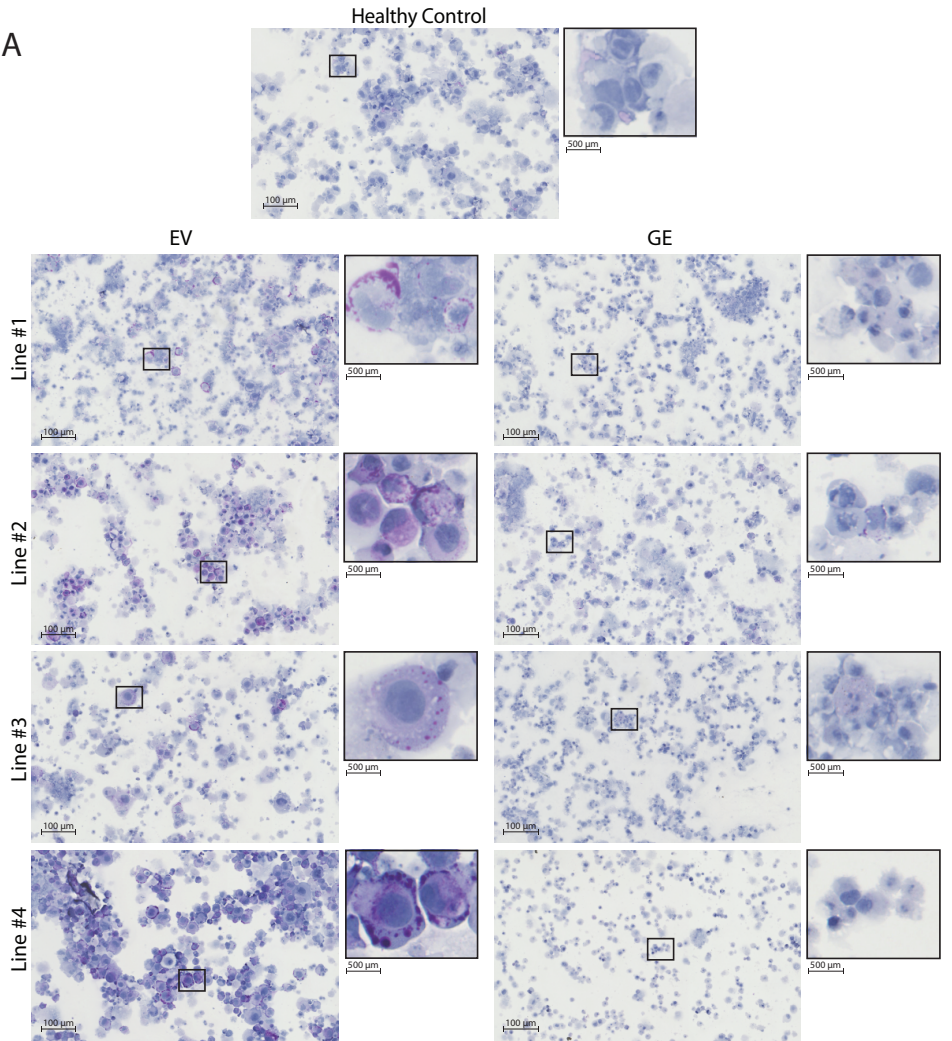
cases seemed fused (see insets in **Figure 4A**). These PAS-positive areas were rare or absent in GE cells and in cells from a healthy control (**Figure 4A**). Quantification of the number of PAS positive cells confirmed this (**Figure 4B**). We conclude that hiPSC-derived chondrogenic cells generated from MPS VI patients, differently from the GE hiPSC, store PAS-reactive material.

Genome-wide mRNA expression analysis of isogenic pairs

RNA sequence analysis was used in a unbiased approach to compare genome-wide changes in mRNA expression between chondrogenic cells from all 4 isogenic pairs. Biological triplicates clustered together (with one exception) and in general the differences between patients were larger than differences caused by gene correction, highlighting the need for isogenic controls (**Figure S2**). This showed a total number of 703 differentially expressed genes when all 4 GE cells were compared to all 4 EV cells based on a fold change (FC) of >1.5, false discovery rate (FDR) of <0.05 and counts per million (CPM) >0 (**Figure S3-S4**). The subsets of upregulated genes in GE (420) and EV (283) were analyzed for enrichment of gene ontology (GO) terms using PANTHER (pantherdb.org). Analysis of biological processes revealed that 120 biological processes were upregulated in GE and 99 were upregulated in EV cells. The 25 most significant terms are shown for the upregulated genes of both sets in **Figure 5A**. Five over- or underrepresented biological processes were interesting in the light of cartilage and lysosomal homeostasis: cell growth and apoptosis, bone and cartilage development, Wnt signaling pathway, ion transport and regulation of metabolic processes. Analysis of GO terms for molecular function showed that 9 groups were upregulated in GE and 12 that were upregulated in EV cells, mainly related to enzyme activity and ion channels (**Figure 5B**). Analysis of the GO terms for cellular component showed an upregulation of 12 components in GE cells, related to the Golgi apparatus and nuclear membrane, and 16 in EV cells, related to the spindle and vesicles (**Figure 5C**).

Bone and Cartilage development

Several key genes involved in bone and cartilage development were differently expressed between GE and EV cells (**Figure 6A**): 28 genes were upregulated in GE and 6 in EV cells. Fold changes ranged from 1.5 for *IGFBP5* to 7.6 for *ACAN* (**Figure 7A**). Genes such as *TGFB2*, *BMP6*, *GDF6*, *PTN* and members of the *CCN* family *CYR61*, *NOV* and *CTGF* are all involved in the BMP/TGF- β signaling pathway and important in bone and cartilage development [16]. *CTGF*, a BMP-7 inhibitor, was highly expressed in GE cells, and may explain the decreased expression of *BMP7* in EV cells. In addition, other proteins involved in extracellular matrix formation such as *ACAN*, *COMP* [17, 18], *FBLN5* [19, 20], and *SULF1* were upregulated in GE cells.



Legend on next page

These results suggest an dysregulated development of chondrogenic cells and extracellular matrix formation in MPS VI.

Wnt signaling pathway

GO analysis indicated a dysregulation of the canonical Wnt pathway as several signaling inhibitors such as *AXIN2* [21-24], *NKD1* [21, 22, 25], *NKD2* [21, 26], *APCDD1* [27, 28], *DRAXIN* [29], *KREMEN1* [30, 31], *WNT9A* [32] and *LEF1* [33] were upregulated in EV (**Figure 6B**). Fold changes ranged from 1,6 for *KREMEN1* to 3,9 and 3,7 for *NKD2* and *NKD1*, respectively (**Figure 7B**). Apart from *WNT9A*, all other Wnt genes were either too lowly expressed or did not show a fold change above 1.5. *WNT2B* did show a > 1.5-fold change upregulation in GE cells, but this was not significant with an FDR of 0,24 (data not shown). Upregulation of inhibitory genes may either indicate low Wnt levels and inactive β -catenin, or active Wnt signaling as inhibitors can be part of a feedback inhibition following pathway activation.

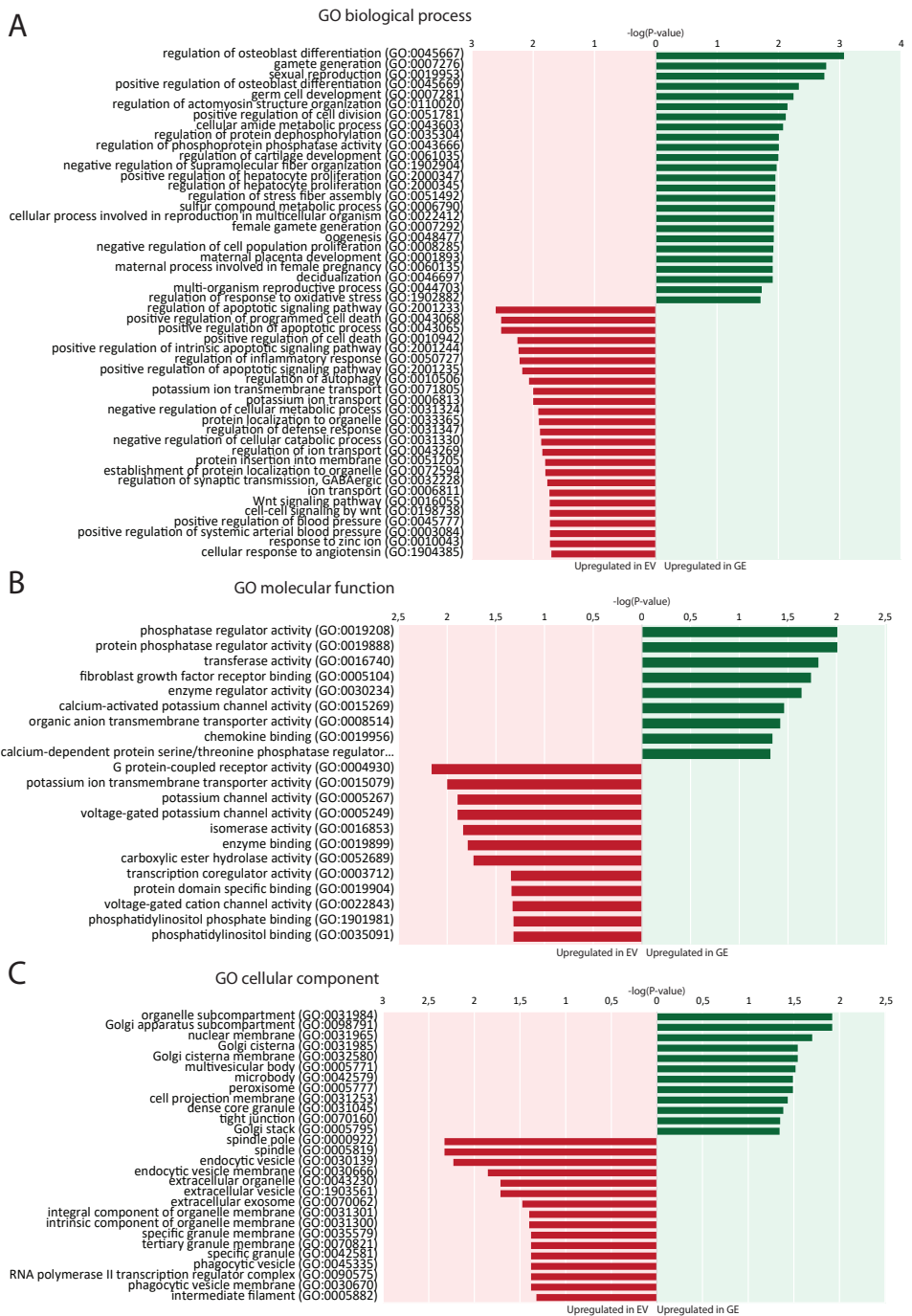
Cell growth and apoptosis

GO analysis further indicated a dysregulation of cell growth and apoptosis in EV cells: 37 genes associated with negative regulation of cell growth and positive regulation of apoptosis were upregulated in EV cells (**Figure 6C**). Genes upregulated in EV cells that inhibit cell growth and promote apoptosis included: *MSX1* [34-36], *HMOX1* [37], *PHLDA1* [38], *BMP7* [39], *EGR3* [40], *SEMA6A* [41], *BCL2L1* [42, 43], *BMF* [43], and *G0S2* [44]. Fold changes ranged from 1.5 for *PHLDA1* to 3.7 for *NKD1* (**Figure 7C**). These results suggest that the EV cells may have a decreased cell growth and a higher apoptotic rate compared to GE cells.

Metabolic processes

The category metabolic processes contained 52 genes that were upregulated in EV cells and 11 that were upregulated in GE cells (**Figure 6D**). Fold changes ranged from 1.5 for *TIPARP* to 5.1 for *ARSB* itself (**Figure 7D**). Although most genes found in the analysis are factors indirectly involved in metabolic processes, several genes upregulated in EV cells are of particular interest. These include genes linked to the ECM such as *SPOCK2*, *MMP9*, *ITIH5* and the collagens *COL6A3* and *COL7A1*. Several

Figure 4. Normalization of enlarged lysosomes after gene correction. A: periodic acid–Schiff (PAS) staining of day 14 differentiated chondrogenic cells showed the accumulation of enlarged lysosomes in EV cells, indicated by the increase of purple signal. The insert shows a higher magnification of selected cells to visualize the enlarged lysosomes. B: Quantification of PAS positive cells showed the normalization of enlarged lysosomes after gene correction. Data represent means \pm SD of 4 biological replicas. GE: Gene edited with *ARSB* cDNA, EV: Gene edited with empty vector control.



Legend on next page

genes related to GAG metabolism were upregulated in GE cells, including *SULF1*, *ITIH3*, *CHSY3*, involved in synthesis of CS, and *ST3GAL1*, involved in the synthesis of gangliosides including GM2 and GM3. Two E3 ubiquitin ligases, *RNF125* and *CBLB*, were upregulated in EV cells and are involved in proteasome-mediated protein degradation. Other genes upregulated in EV cells included *HEY2*, *SMPD3* and *XDH*. *HEY2* represses transcription by interaction with a histone deacetylase complex and is induced by the Notch signaling pathway [45]. *SMPD3* catalyzes sphingomyelin hydrolysis. *XDH* catalyzes oxidative metabolism of purines. Other genes of interest upregulated in GE cells include *UGCG* and *FAR2*, involved in biosynthesis of glycosphingolipids and reduction of saturated fatty acids to fatty alcohols, respectively. These results indicate that ARSB deficiency may indirectly deregulate other metabolic processes besides GAG degradation.

Ion transport

GO analysis showed upregulation of 28 genes in EV cells that are involved in ion transport or regulation of ion transport (**Figure 6E**). Fold changes ranged from 1.5 for *SLC4A8* to 3 for *SCN3A* (**Figure 7E**). In addition, GO terms for molecular function analysis indicated the upregulation of potassium ion channel activity in EV cells (**Figure 5B**). Calcium-activated potassium channel activity was found to be upregulated in GE cells based on the molecular function analysis.

Discussion

In this study, we have generated a patient-derived *in vitro* disease model for MPS VI with isogenic controls, and used this to obtain insight in early stages of pathology in chondrogenic cells. We have applied this model to four patients and characterized cell biological and gene expression changes. This confirmed clinical information and previous reports using animal models, and provided novel insight in the early molecular changes in chondrocytes that are associated with MPS VI.

We used PAS to assess whether chondrocytes accumulated glycoprotein macromolecules. PAS positive vesicles were detected at day 14 of differentiation in disease chondrogenic cells relative to their isogenic controls. At present, the identity of the PAS positive material is unknown. We used a short protocol for PAS

Figure 5. Enrichment of gene ontology (GO) terms. A: 25 top dysregulated GO biological processes in EV and GE cells. B: Significantly dysregulated GO molecular functions. C: Significantly dysregulated GO cellular components. Upregulated processes in GE cells are indicated in green. Upregulated processes in EV cells are indicated in red. GE: Gene edited with ARSB cDNA, EV: Gene edited with empty vector control.

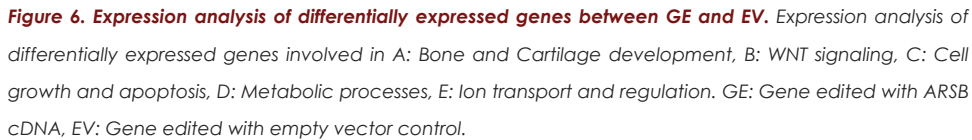


Figure 6. Expression analysis of differentially expressed genes between GE and EV. Expression analysis of differentially expressed genes involved in A: Bone and Cartilage development, B: WNT signaling, C: Cell growth and apoptosis, D: Metabolic processes, E: Ion transport and regulation. GE: Gene edited with ARSB cDNA, EV: Gene edited with empty vector control.

staining of cytopins that is unlikely capable of detecting charged GAGs such as DS and CS. Additional analysis using red oil staining ruled out that these were lipid droplets (data not shown). We hypothesize that they might represent secondary storage products, possibly GM2 or GM3, which are known to accumulate in MPS VI [7]. Interestingly, the EM analysis of a cartilage biopsy from an MPS VI patient (see **Figure 1**) showed very heterogeneous shapes and forms of vesicles, suggesting that they might contain a mixture of storage products, which should be confirmed in future experiments. The fact that PAS positive material was already detected at day 14 of chondrogenic differentiation suggests that chondrogenic cells have an early onset pathology, which is in line with clinical information on hip abnormalities in MPS VI that likely develop before the age of 1 year [1, 5].

Early disease onset in chondrogenic cells was also suggested by RNAseq analysis. Key genes in bone and cartilage development were differentially expressed between EV and GE cells. Genes involved in the TGF- β /BMP signaling, which plays a fundamental role in cartilage homeostasis, were downregulated in EV cells. The TGF- β /BMP signaling pathway is known to be influenced by GAGs, but it is not fully understood how GAG accumulation in MPS influences its activity [5]. Heparan sulfate containing proteoglycans (HSPG) can modulate BMPs and their antagonists [46] and chondroitin sulfate is linked to BMPs and TGF- β activity and intracellular localization [47, 48]. TGF- β 2 is one of the members of the TGF- β superfamily and was downregulated in EV cells. Downregulated members of the BMP family in EV cells included *BMP6* and *GDF6* [16, 49, 50]. *BMP6* stimulates differentiation of MSCs into chondrocytes and promotes the synthesis of chondrocytes [51-53]. In addition, *BMP6* induces proteoglycan synthesis, which is particularly interesting in the context of the GAG accumulation in MPS VI [52, 54, 55]. *GDF6* is a member of the BMP family and a regulatory protein associated with the growth and differentiation cartilaginous tissue [56]. *PTN* is a TGF- β dependent heparin-binding growth-associated molecule and downregulated in EV cells, it is involved in a variety of processes in bone formation and stimulates chondrocyte proliferation [57, 58]. The CCN family members *CYR61*, *NOV* and *CTGF* were upregulated in GE cells and are known to induce the expression of chondrogenic markers and to regulate TGF- β and BMP [59-62]. *CYR61*, is a direct target of canonical Wnt/beta-catenin signaling and is involved in osteogenic differentiation and bone healing [63]. *NOV* is involved in the maintenance of articular cartilage and may inhibit osteoarthritis [64, 65]. *CTGF* is involved in the regulation of chondrogenesis and is a BMP-7 inhibitor, possibly mediating the downregulation of *BMP7* in GE cells [59-61]. *COMP* was downregulated in EV and encodes an extracellular matrix protein which plays an important role in cartilage matrix organization through interaction with other extracellular matrix proteins [17, 18, 66]. Disease-associated variants in *COMP*

cause impaired cartilage development leading to skeletal abnormalities such as short stature and skeletal dysplasias [67-73]. All together, these results suggest that the accumulation of intralysosomal GAGs in MPS VI impairs early chondrogenic development.

Several Wnt inhibitors of the Wnt/ β -catenin signaling pathway were upregulated in EV cells including *AXIN2*, *NKD1*, *NKD2*, *APCDD1*, *DRAXIN*, *KREMEN1*, *WNT9A* and *LEF1*. *AXIN2* is a scaffold protein of the β -catenin destruction complex and promotes phosphorylation of β -catenin and its consequent degradation. *NKD1* and *NKD2* interact with *AXIN2* to inhibit β -catenin in the canonical Wnt signaling pathway [74]. *APCDD1* can co-precipitate with Wnt3A and LRP5 and likely inhibits Wnt/ β -catenin signaling by preventing formation of the Wnt receptor complex [27]. The zebrafish homologue of *DRAXIN*, *Neucrin*, has been suggested to inhibit the stabilization of β -catenin during canonical Wnt signalling [75]. *KREMEN1* is a transmembrane protein that acts as a receptor for Dickkopf-1 with which it functionally cooperates to inhibit the Wnt/ β -catenin signaling. *WNT9A* has been suggested as a non-canonical ligand to inhibit β -catenin [32]. *LEF1* is a key target in the Wnt signaling pathway and can negatively regulate the expression of Wnt signaling genes by binding to Groucho-related corepressors [33]. This upregulation of Wnt inhibitors suggests that the Wnt signaling pathway is dysregulated in chondrogenic cells with MPS VI. Interestingly, a link between GAGs and Wnt signaling has been made previously: biglycan, a protein core with two CS or DS GAG chains, affects the Wnt signaling pathway through a direct interaction of the core protein with the LRP6 receptor and its Wnt ligand [76]. The Wnt signaling pathway plays a crucial role in bone and cartilage development [77-79], and downregulation of Wnt2 and beta-catenin can inhibit cell proliferation and induce apoptosis [80-82], suggesting an impaired chondrogenic development and reduced cell proliferation in MPS VI.

Additional evidence for increased apoptosis and decreased cell growth in MPS VI was obtained by upregulation of 37 genes associated with these processes in EV cells. These included *MSX1*, *HMOX1*, *PHLDA1*, *BMP7*, *EGR3*, *SEMA6A*, *BCL2L11*, *BMF* and *G0S2*. *MSX1* has been shown to inhibit cell growth and induce apoptosis by inhibiting the Notch signaling pathway and maintaining cyclin D1 expression [34-36, 83]. Upregulation of *PHLDA1* expression leads to reduced cell growth and increased apoptosis [38, 84, 85]. *BMP7* contributes to cell cycle arrest at the G1/S checkpoint and stimulates apoptosis via Smad1-dependent and Smad1-independent pathways [39, 86]. *SEMA6A* induces apoptosis via the cytosolic region of the SEMA

Figure 7. The 2logFC of genes involved in A: Bone and Cartilage development, B: WNT signaling, C: Cell growth and apoptosis, D: Metabolic processes, E: Ion transport and regulation. GE: Gene edited with ARSB cDNA, EV: Gene edited with empty vector control.

domain through association with Fas-associated protein with death domain (FADD) [41]. *EGR3* inhibits cell proliferation and induces apoptosis via upregulation of Fas ligand [40]. *BCL2L11*, also known as *BIM*, *BMF* and *G0S2* interact together to promote apoptosis [87]. Increased apoptosis in MPS VI chondrocytes is likely due to the GAG accumulation, which is consistent with findings from Simonaro et al [88] who showed that GAG accumulation leads to apoptosis of articular chondrocytes. However, in this case chondrocyte apoptosis was due to the activation of the TLR-4 signaling pathway followed by the release of proinflammatory cytokines. In contrast, we showed increased apoptosis in the absence of proinflammatory cytokines, suggesting the existence of an intrinsic non-inflammatory mechanism leading to apoptosis.

Interestingly, several genes linked to GAG metabolism were dysregulated in EV cells. *SPOCK2*, *MMP9* and *ITIH5* were upregulated and *SULF1*, *ITIH3*, *CHSY3* and *ST3GAL1* were downregulated compared to GE cells. *SPOCK2* encodes the protein Testican-2, a proteoglycan that binds to glycosaminoglycans and forms part of the extracellular matrix. *MMP9* encodes an enzyme belonging to the matrix metalloproteinase family with substrates including aggrecan and collagen and non-ECM substrates such as TGF- β 1 [89]. *ITIH5* is involved in extracellular matrix stabilization. *SULF1* encodes a heparan sulfate endosulfatase which removes sulfate groups from chains of heparan sulfate proteoglycans. *ITIH3* encodes a subunit of the pre-alpha-trypsin inhibitor complex which binds to bind hyaluronic acid and stabilizes the extracellular matrix. *CHSY3*, also known as *CSS3* or chondroitin sulfate synthase, is a glycosyltransferase that is involved in transferring GlcUA and GalNAc to the nonreducing terminus of chondroitin sulfate. Its downregulation in MPS VI may indicate a feedback mechanism induced by CS accumulation to reduce the synthesis of CS. *ST3GAL1* encodes a sialyltransferase that catalyzes the transfer of sialic acid from cytidine monophosphate-sialic acid to galactose-containing substrates such as GAGs [90]. It is also involved in the synthesis of GM2 and GM3, known secondary storage products in many lysosomal storage disorders including MPS VI [7]. These findings indicate that MPS VI chondrogenic cells have disturbed metabolism of GAGs and other ECM components, including feedback inhibition of CS synthesis and disturbed ganglioside metabolism.

Genes involved in ion transport of calcium and potassium and encoding other ion channels were upregulated in EV cells. Genes involved in potassium transport included *KCNF1*, *KCNJ15*, *KCNS1*, *KCNA5*, *KCNQ2* and *SLC12A5*. Genes linked to calcium transport include *FLVCR2*, *CACNA1D*, *TMEM37* and *STC1*. Other genes upregulated in EV cells included *CHRNA9*, *SCN3A*, and *P2RX6*, a ligand-gated ionic channel, a voltage-gated sodium channels and ATP-gated ion, respectively. Changes in ion channels have been linked previously to osteoarthritic cartilage [91,

92]. In addition, extracellular GAG deposits in tissue causes increased absorption of water which results in inflated tissue [93]. The intralysosomal accumulation of GAGs in EV cells might cause osmotic stress as GAGs can retain up to 100 times up to its weight. The dysregulation of ion channels might be a regulatory mechanism to compensate for the difference in cell osmolarity [94-96]. In addition, accumulation of GAGs is suggested to compromise the integrity of lysosomal membranes [93], which can lead to a dysregulation of ion content in the lysosome.

In summary, we have generated a disease model for the cartilage pathology in MPS VI based on hiPSCs and isogenic controls generated using CRISPR/Cas9. This has proven to be a powerful approach to detect the early cellular and molecular changes in chondrogenic cells that ultimately lead to the severe cartilage pathology in MPS VI patients. As the approach is generic, it should be generally applicable to model genetic disorders that affect cartilage. Future work should focus on further development of the model into actual cartilage that is subject to mechanical loading that mimics natural loading of weight bearing joints that occurs in human individuals.

Methods

Ethics Approval and Consent to Participate

The Institutional Review Board approved the study protocol, and all patients provided written informed consent.

hiPSC

hiPSCs were generated from patient-derived primary fibroblasts as described before [13]. In short, patient-derived fibroblast cells were reprogrammed using a polycistronic lentiviral vector of Oct4, Sox2, Klf4, and c-Myc (LV-OSKM) and cultured on γ -irradiated mouse embryonic feeders. After gene correction, the hiPSCs were transferred to a feeder free culture with Vitronectin XF (Stem Cell) as coating and mTeSR™ Plus (Stem Cell) as media. Healthy control hiPSCs were obtained from the HIPSCI database.

Gene correction

The donor construct to express the healthy copy of the ARSB gene was generated as described previously [13, 14], with the ARSB cDNA instead of the GAA cDNA. Approximately 14 days after initiation of selection, single IPS colonies were picked and genotyped using primers from **Table S2**.

Chondrogenic differentiation

hiPSCs were differentiated into chondrogenic cells using a 14-day protocol based on Oldershaw et al. [15]. Single cells were generated from hiPSC colonies by incubation with TrypLE (Thermo Fisher Scientific), and 5×10^5 cells were plated on a Vitronectin XF (Stem cell) coated plate. This protocol consists of activation of the WNT signaling using $2 \mu\text{M}$ CHIR99021 (R & D Systems) and 50 to 10 ng/mL Activin-A (Peprotech) from day 1 to 3. 40 ng/mL FGF2 (Peprotech) was added from day 2 and sustained during the whole differentiation. 5 ng/mL BMP2 (R&D Systems) was added from day 3 to 10. $1 \mu\text{M}$ SB431542 (Torc1s) was sustained from day 5 to 8. GDF5 was added in increasing concentrations from day 9 onwards: 20 ng/mL on day 9 and 10, 40 ng/mL from day 11 until day 14. Cells were split on day 5 in a 1/5 to 1/8 ratio and on day 8 in a 1/4 to 1/6 ratio. Medium was supplemented with Revitacell on the day of splitting. Cell culture plates were coated with Vitronectin XF on days 1-7 and from day 8 to 14 with a Vitronectin XF and 0,1% gelatin mix (Sigma-Aldrich).

Periodic acid–Schiff staining

Day 14 chondrogenic cells were harvested by incubating with TrypLE (Thermo Fisher Scientific) and single cells were centrifuged on a slide using the Cytospin. The samples were fixed in a formaldehyde ethanol solution for 5 minutes and rinsed with water. Samples were then incubated in 1% Periodic acid (Supelco) for 10 minutes and rinsed with water for 10 minutes. The samples were placed in Schiff's reagent (Sigma-Aldrich) for 15 minutes at 37°C and rinsed with water for 15 minutes. The samples were then placed in Mayer's hemalum solution (Sigma-Aldrich) for 5 minutes and rinsed with water. They were air-dried and mounted with Entellan (Sigma-Aldrich). Images were obtained using a NanoZoomer 2.0HT C9600-12 (Hamamatsu).

ARSB Enzyme assay

The ARSB enzyme activity was measured as described previously [97]. In short, cells were lysed in TNE-1% (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 50 mM NaF, and 1% Triton X100, containing a protease inhibitor cocktail (Roche) and added to the paranitrocatecholdisulfate substrate. The conversion of paranitrocatecholdisulfate to paranitrocatechol was measured on a spectrophotometer.

RNA isolation and sequencing

RNA was isolated using the RNAeasy miniprep kit with on-column DNase treatment (Qiagen) according to manufacturer's protocol. Sequencing was performed after poly-A selection and TruSeq library prep at the Human Genomics facility (glindna.org) on a HiSeq2000 at paired-end 150 bp. Data were processed per sample using STAR (v2.3.0) [98, 99], picard (v1.90), and GATK (v2.8). Transcript quantification

was performed using featureCounts (v1.4.3) against all 57,820 gene features in GENCODE (version date; 2013-12-05) [100, 101].

Data analysis

Raw counts per gene were normalized using the edgeR (v3.8.6) trimmed mean of M-values method to counts per million values. Principal components (PCs) were calculated using “prcomp” in R, and then plotted to visually identify sample outliers. Statistical analysis was performed per gene using the glmFit function in edgeR, correcting stratifying for the four cell lines. Selected groups of genes are clustered and visualized using z-transformation per gene set and subsequent euclidean hierarchical clustering in the TIBCO Spotfire package (v7.14).

GO Analysis

Differential expressed genes among isogenic controls were analyzed by the PANTHER Classification System [102]. Selection for genes was based on a CPM of >0 for all samples and replicates, a log2 FC of 0.59 or -0.59 for upregulated and downregulated genes, respectively, and an FDR of ≤0.05. The flowchart for the selection of genes is shown in **Figure S3**.

Author Contributions

M.B. and W.W.M.P.P. conceived and designed the study and drafted the manuscript. M.B. and T.J.M.V.G. performed the experiments. W.W.M.P.P. supervised the study. M.B., J.G.J.v.R., T.J.M.v.G., C.A.S., S.J.K., E.O., M.A.E.M.W., J.M.P.v.d.H., A.T.v.d.P., R.N., W.W.M.P. were involved in data interpretation and approved the final manuscript

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

This work has received funding from Zeldzame Ziekten Fonds/WE Foundation, Metakids (project number 2018-082), and Stofwisselkracht. RN is part of the Erasmus Postgraduate School Molecular Medicine and Medical Delta Regenerative Medicine 4D program. We thank G.J.V.M (Gerjo) van Osch for scientific input.

References

1. Oussoren, E., J. Bessems, V. Pollet, et al., A long term follow-up study of the development of hip disease in Mucopolysaccharidosis type VI. *Mol Genet Metab*, 2017. 121(3): p. 241-251.
2. Haskins, M.E., G.D. Aguirre, P.F. Jezyk, et al., The pathology of the feline model of mucopolysaccharidosis VI. *Am J Pathol*, 1980. 101(3): p. 657-74.
3. Abreu, S., J. Hayden, P. Berthold, et al., Growth plate pathology in feline mucopolysaccharidosis VI. *Calcif Tissue Int*, 1995. 57(3): p. 185-90.
4. Evers, M., P. Saffig, P. Schmidt, et al., Targeted disruption of the arylsulfatase B gene results in mice resembling the phenotype of mucopolysaccharidosis VI. *Proc Natl Acad Sci U S A*, 1996. 93(16): p. 8214-9.
5. Oussoren, E., M.M. Brands, G.J. Ruijter, et al., Bone, joint and tooth development in mucopolysaccharidoses: relevance to therapeutic options. *Biochim Biophys Acta*, 2011. 1812(11): p. 1542-56.
6. Oussoren, E., M. Wagenmakers, B. Link, et al., Hip disease in Mucopolysaccharidoses and Mucopolipidoses: A review of mechanisms, interventions and future perspectives. *Bone*, 2021. 143: p. 115729.
7. Ballabio, A. and V. Gieselmann, Lysosomal disorders: from storage to cellular damage. *Biochim Biophys Acta*, 2009. 1793(4): p. 684-96.
8. Jiang, Z., A.L.K. Derrick-Roberts, C. Reichstein, et al., Cell cycle progression is disrupted in murine MPS VII growth plate leading to reduced chondrocyte proliferation and transition to hypertrophy. *Bone*, 2020. 132: p. 115195.
9. Hockemeyer, D. and R. Jaenisch, Induced Pluripotent Stem Cells Meet Genome Editing. *Cell Stem Cell*, 2016. 18(5): p. 573-86.
10. Soldner, F., J. Laganier, A.W. Cheng, et al., Generation of isogenic pluripotent stem cells differing exclusively at two early onset Parkinson point mutations. *Cell*, 2011. 146(2): p. 318-31.
11. Kolf, C.M., E. Cho, and R.S. Tuan, Mesenchymal stromal cells. *Biology of adult mesenchymal stem cells: regulation of niche, self-renewal and differentiation*. *Arthritis Res Ther*, 2007. 9(1): p. 204.
12. Broeders, M., K. Smits, B. Goynuk, et al., A generic assay to detect aberrant ARSB splicing and mRNA degradation for the molecular diagnosis of MPS VI. *Molecular Therapy - Methods & Clinical Development*, 2020.
13. van der Wal, E., P. Herrero-Hernandez, R. Wan, et al., Large-Scale Expansion of Human iPSC-Derived Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies. *Stem Cell Reports*, 2018. 10(6): p. 1975-1990.
14. In 't Groen, S.L.M., M. Broeders, and W.W.M.P. Pijnappel, CRISPR/Cas9-mediated gene editing in human induced pluripotent stem cells. *CRISPR-Cas Methods: Volume II*, Springer Nature, In press.
15. Oldershaw, R.A., M.A. Baxter, E.T. Lowe, et al., Directed differentiation of human embryonic stem cells toward chondrocytes. *Nat Biotechnol*, 2010. 28(11): p. 1187-94.
16. Chang, C., Agonists and Antagonists of TGF-beta Family Ligands. *Cold Spring Harb Perspect Biol*, 2016. 8(8).

17. Coustry, F., K.L. Posey, T. Maerz, et al., Mutant cartilage oligomeric matrix protein (COMP) compromises bone integrity, joint function and the balance between adipogenesis and osteogenesis. *Matrix Biol*, 2018. 67: p. 75-89.
18. Acharya, C., J.H. Yik, A. Kishore, et al., Cartilage oligomeric matrix protein and its binding partners in the cartilage extracellular matrix: interaction, regulation and role in chondrogenesis. *Matrix Biol*, 2014. 37: p. 102-11.
19. de Vega, S., T. Iwamoto, and Y. Yamada, Fibulins: multiple roles in matrix structures and tissue functions. *Cell Mol Life Sci*, 2009. 66(11-12): p. 1890-902.
20. Timpl, R., T. Sasaki, G. Kostka, et al., Fibulins: a versatile family of extracellular matrix proteins. *Nat Rev Mol Cell Biol*, 2003. 4(6): p. 479-89.
21. Katoh, M. and M. Katoh, WNT signaling pathway and stem cell signaling network. *Clin Cancer Res*, 2007. 13(14): p. 4042-5.
22. Larraguibel, J., A.R. Weiss, D.J. Pasula, et al., Wnt ligand-dependent activation of the negative feedback regulator Nkd1. *Mol Biol Cell*, 2015. 26(12): p. 2375-84.
23. Kikuchi, A., Roles of Axin in the Wnt signalling pathway. *Cell Signal*, 1999. 11(11): p. 777-88.
24. Jho, E.H., T. Zhang, C. Domon, et al., Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Mol Cell Biol*, 2002. 22(4): p. 1172-83.
25. Angonin, D. and T.J. Van Raay, Nkd1 functions as a passive antagonist of Wnt signaling. *PLoS One*, 2013. 8(8): p. e74666.
26. Zhao, S., L. Kurenbekova, Y. Gao, et al., NKD2, a negative regulator of Wnt signaling, suppresses tumor growth and metastasis in osteosarcoma. *Oncogene*, 2015. 34(39): p. 5069-79.
27. Shimomura, Y., D. Agalliu, A. Vonica, et al., APCDD1 is a novel Wnt inhibitor mutated in hereditary hypotrichosis simplex. *Nature*, 2010. 464(7291): p. 1043-7.
28. Cruciat, C.M. and C. Niehrs, Secreted and transmembrane wnt inhibitors and activators. *Cold Spring Harb Perspect Biol*, 2013. 5(3): p. a015081.
29. Hutchins, E.J. and M.E. Bronner, Draxin acts as a molecular rheostat of canonical Wnt signaling to control cranial neural crest EMT. *J Cell Biol*, 2018. 217(10): p. 3683-3697.
30. Mao, B., W. Wu, G. Davidson, et al., Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin signalling. *Nature*, 2002. 417(6889): p. 664-7.
31. Causeret, F., I. Sumia, and A. Pierani, Kremen1 and Dickkopf1 control cell survival in a Wnt-independent manner. *Cell Death Differ*, 2016. 23(2): p. 323-32.
32. Ali, I., B. Medegan, and D.P. Braun, Wnt9A Induction Linked to Suppression of Human Colorectal Cancer Cell Proliferation. *Int J Mol Sci*, 2016. 17(4): p. 495.
33. Steinke, F.C. and H.H. Xue, From inception to output, Tcf1 and Lef1 safeguard development of T cells and innate immune cells. *Immunol Res*, 2014. 59(1-3): p. 45-55.
34. Yue, Y., Y. Yuan, L. Li, et al., Homeobox protein MSX1 inhibits the growth and metastasis of breast cancer cells and is frequently silenced by promoter methylation. *Int J Mol Med*, 2018. 41(5): p. 2986-2996.
35. Yue, Y., K. Zhou, J. Li, et al., MSX1 induces G0/G1 arrest and apoptosis by suppressing Notch signaling

- and is frequently methylated in cervical cancer. *Onco Targets Ther*, 2018. 11: p. 4769-4780.
36. Lallemand, Y., V. Bensoussan, C.S. Cloment, et al., *Msx genes are important apoptosis effectors downstream of the Shh/Gli3 pathway in the limb*. *Dev Biol*, 2009. 331(2): p. 189-98.
 37. Hill, M., V. Pereira, C. Chauveau, et al., *Heme oxygenase-1 inhibits rat and human breast cancer cell proliferation: mutual cross inhibition with indoleamine 2,3-dioxygenase*. *FASEB J*, 2005. 19(14): p. 1957-68.
 38. Zimnicka, A.M., T. Sharma, M. Regan, et al., *Knockout of the PHLDA1 gene in breast cancer cells reveals multiple roles for PHLDA1 in cancer phenotypes*. *The FASEB Journal*, 2017. 31(S1): p. 178.8-178.8.
 39. Piscione, T.D., T. Phan, and N.D. Rosenblum, *BMP7 controls collecting tubule cell proliferation and apoptosis via Smad1-dependent and -independent pathways*. *Am J Physiol Renal Physiol*, 2001. 280(1): p. F19-33.
 40. Zhang, S., C. Xia, C. Xu, et al., *Early growth response 3 inhibits growth of hepatocellular carcinoma cells via upregulation of Fas ligand*. *Int J Oncol*, 2017. 50(3): p. 805-814.
 41. Shen, C.Y., Y.C. Chang, L.H. Chen, et al., *The extracellular SEMA domain attenuates intracellular apoptotic signaling of semaphorin 6A in lung cancer cells*. *Oncogenesis*, 2018. 7(12): p. 95.
 42. Koenig, M.N., E. Naik, L. Rohrbeck, et al., *Pro-apoptotic BIM is an essential initiator of physiological endothelial cell death independent of regulation by FOXO3*. *Cell Death Differ*, 2014. 21(11): p. 1687-95.
 43. Pinon, J.D., V. Labi, A. Egle, et al., *Bim and Bmf in tissue homeostasis and malignant disease*. *Oncogene*, 2008. 27 Suppl 1: p. S41-52.
 44. Welch, C., M.K. Santra, W. El-Assaad, et al., *Identification of a protein, G0S2, that lacks Bcl-2 homology domains and interacts with and antagonizes Bcl-2*. *Cancer Res*, 2009. 69(17): p. 6782-9.
 45. Weber, D., C. Wiese, and M. Gessler, *Hey bHLH transcription factors*. *Curr Top Dev Biol*, 2014. 110: p. 285-315.
 46. Khan, S.A., M.S. Nelson, C. Pan, et al., *Endogenous heparan sulfate and heparin modulate bone morphogenetic protein-4 signaling and activity*. 2008. 294(6): p. C1387-97.
 47. Kluppel, M., T.N. Wight, C. Chan, et al., *Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis*. *Development*, 2005. 132(17): p. 3989-4003.
 48. Alliston, T., *Chondroitin sulfate and growth factor signaling in the skeleton: Possible links to MPS VI*. *J Pediatr Rehabil Med*, 2010. 3(2): p. 129-38.
 49. Xu, X., L. Zheng, Q. Yuan, et al., *Transforming growth factor-beta in stem cells and tissue homeostasis*. *Bone Res*, 2018. 6: p. 2.
 50. Tan Timur, U., M. Caron, G. van den Akker, et al., *Increased TGF-beta and BMP Levels and Improved Chondrocyte-Specific Marker Expression In Vitro under Cartilage-Specific Physiological Osmolarity*. *Int J Mol Sci*, 2019. 20(4).
 51. Ye, F., H. Xu, H. Yin, et al., *The role of BMP6 in the proliferation and differentiation of chicken cartilage cells*. *PLoS One*, 2019. 14(7): p. e0204384.
 52. Thielen, N.G.M., P.M. van der Kraan, and A.P.M. van Caam, *TGFbeta/BMP Signaling Pathway in Cartilage Homeostasis*. *Cells*, 2019. 8(9).

53. Rahman, M.S., N. Akhtar, H.M. Jamil, et al., TGF-beta/BMP signaling and other molecular events: regulation of osteoblastogenesis and bone formation. *Bone Res*, 2015. 3: p. 15005.
54. Chubinskaya, S., D. Segalite, D. Pikovsky, et al., Effects induced by BMPs in cultures of human articular chondrocytes: comparative studies. *Growth Factors*, 2008. 26(5): p. 275-83.
55. Bobacz, K., R. Gruber, A. Soleiman, et al., Expression of bone morphogenetic protein 6 in healthy and osteoarthritic human articular chondrocytes and stimulation of matrix synthesis in vitro. *Arthritis Rheum*, 2003. 48(9): p. 2501-8.
56. Wei, A., B. Shen, L.A. Williams, et al., Expression of growth differentiation factor 6 in the human developing fetal spine retreats from vertebral ossifying regions and is restricted to cartilaginous tissues. *J Orthop Res*, 2016. 34(2): p. 279-89.
57. Pufe, T., G. Groth, M.B. Goldring, et al., Effects of pleiotrophin, a heparin-binding growth factor, on human primary and immortalized chondrocytes. *Osteoarthritis Cartilage*, 2007. 15(2): p. 155-62.
58. Tare, R.S., R.O. Oreffo, N.M. Clarke, et al., Pleiotrophin/Osteoblast-stimulating factor 1: dissecting its diverse functions in bone formation. *J Bone Miner Res*, 2002. 17(11): p. 2009-20.
59. Abreu, J.G., N.I. Ketpura, B. Reversade, et al., Connective-tissue growth factor (CTGF) modulates cell signalling by BMP and TGF-beta. *Nat Cell Biol*, 2002. 4(8): p. 599-604.
60. Nguyen, T.Q., P. Roestenberg, F.A. van Nieuwenhoven, et al., CTGF inhibits BMP-7 signaling in diabetic nephropathy. *J Am Soc Nephrol*, 2008. 19(11): p. 2098-107.
61. Maeda, A., T. Nishida, E. Aoyama, et al., CCN family 2/connective tissue growth factor modulates BMP signalling as a signal conductor, which action regulates the proliferation and differentiation of chondrocytes. *J Biochem*, 2009. 145(2): p. 207-16.
62. Holbourn, K.P., K.R. Acharya, and B. Perbal, The CCN family of proteins: structure-function relationships. *Trends Biochem Sci*, 2008. 33(10): p. 461-73.
63. Si, W., Q. Kang, H.H. Luu, et al., CCN1/Cyr61 is regulated by the canonical Wnt signal and plays an important role in Wnt3A-induced osteoblast differentiation of mesenchymal stem cells. *Mol Cell Biol*, 2006. 26(8): p. 2955-64.
64. Huang, X., B. Ni, Z. Mao, et al., NOV/CCN3 induces cartilage protection by inhibiting PI3K/AKT/mTOR pathway. *J Cell Mol Med*, 2019. 23(11): p. 7525-7534.
65. Janune, D., T. Abd El Kader, E. Aoyama, et al., Novel role of CCN3 that maintains the differentiated phenotype of articular cartilage. *J Bone Miner Metab*, 2017. 35(6): p. 582-597.
66. Maly, K., E. Andres Sastre, E. Farrell, et al., COMP and TSP-4: Functional Roles in Articular Cartilage and Relevance in Osteoarthritis. *Int J Mol Sci*, 2021. 22(5).
67. Briggs, M.D., S.M. Hoffman, L.M. King, et al., Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nat Genet*, 1995. 10(3): p. 330-6.
68. Hecht, J.T., L.D. Nelson, E. Crowder, et al., Mutations in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat Genet*, 1995. 10(3): p. 325-9.
69. Unger, S. and J.T. Hecht, Pseudoachondroplasia and multiple epiphyseal dysplasia: New etiologic developments. *Am J Med Genet*, 2001. 106(4): p. 244-50.
70. McKeand, J., J. Rotta, and J.T. Hecht, Natural history study of pseudoachondroplasia. *Am J Med*

- Genet, 1996. 63(2): p. 406-10.
71. Posey, K.L. and J.T. Hecht, The role of cartilage oligomeric matrix protein (COMP) in skeletal disease. *Curr Drug Targets*, 2008. 9(10): p. 869-77.
 72. Hecht, J.T., M. Deere, E. Putnam, et al., Characterization of cartilage oligomeric matrix protein (COMP) in human normal and pseudoachondroplasia musculoskeletal tissues. *Matrix Biol*, 1998. 17(4): p. 269-78.
 73. Hecht, J.T., E. Hayes, R. Haynes, et al., COMP mutations, chondrocyte function and cartilage matrix. *Matrix Biol*, 2005. 23(8): p. 525-33.
 74. Miller, B.W., G. Lau, C. Grouios, et al., Application of an integrated physical and functional screening approach to identify inhibitors of the Wnt pathway. *Mol Syst Biol*, 2009. 5: p. 315.
 75. Miyake, A., Y. Takahashi, H. Miwa, et al., Neucrin is a novel neural-specific secreted antagonist to canonical Wnt signaling. *Biochem Biophys Res Commun*, 2009. 390(3): p. 1051-5.
 76. Berendsen, A.D., L.W. Fisher, T.M. Kilts, et al., Modulation of canonical Wnt signaling by the extracellular matrix component biglycan. *Proc Natl Acad Sci U S A*, 2011. 108(41): p. 17022-7.
 77. Kobayashi, Y., S. Uehara, N. Udagawa, et al., Regulation of bone metabolism by Wnt signals. *J Biochem*, 2016. 159(4): p. 387-92.
 78. Usami, Y., A.T. Gunawardena, M. Iwamoto, et al., Wnt signaling in cartilage development and diseases: lessons from animal studies. *Lab Invest*, 2016. 96(2): p. 186-96.
 79. Narcisi, R., M.A. Cleary, P.A. Brama, et al., Long-term expansion, enhanced chondrogenic potential, and suppression of endochondral ossification of adult human MSCs via WNT signaling modulation. *Stem Cell Reports*, 2015. 4(3): p. 459-72.
 80. Pu, P., Z. Zhang, C. Kang, et al., Downregulation of Wnt2 and beta-catenin by siRNA suppresses malignant glioma cell growth. *Cancer Gene Ther*, 2009. 16(4): p. 351-61.
 81. Zimmerman, Z.F., R.M. Kulikauskas, K. Bomsztyk, et al., Activation of Wnt/beta-catenin signaling increases apoptosis in melanoma cells treated with trail. *PLoS One*, 2013. 8(7): p. e69593.
 82. You, L., B. He, Z. Xu, et al., Inhibition of Wnt-2-mediated signaling induces programmed cell death in non-small-cell lung cancer cells. *Oncogene*, 2004. 23(36): p. 6170-4.
 83. Hu, G., H. Lee, S.M. Price, et al., Msx homeobox genes inhibit differentiation through upregulation of cyclin D1. *Development*, 2001. 128(12): p. 2373-84.
 84. Nagai, M.A., Pleckstrin homology-like domain, family A, member 1 (PHLDA1) and cancer. *Biomed Rep*, 2016. 4(3): p. 275-281.
 85. Neef, R., M.A. Kuske, E. Prols, et al., Identification of the human PHLDA1/TDAG51 gene: down-regulation in metastatic melanoma contributes to apoptosis resistance and growth deregulation. *Cancer Res*, 2002. 62(20): p. 5920-9.
 86. Klose, A., Y. Waerzeggers, P. Monfared, et al., Imaging bone morphogenetic protein 7 induced cell cycle arrest in experimental gliomas. *Neoplasia*, 2011. 13(3): p. 276-85.
 87. Ramjaun, A.R., S. Tomlinson, A. Eddoudi, et al., Upregulation of two BH3-only proteins, Bmf and Bim, during TGF beta-induced apoptosis. *Oncogene*, 2007. 26(7): p. 970-81.
 88. Simonaro, C.M., M. D'Angelo, X. He, et al., Mechanism of glycosaminoglycan-mediated bone and joint disease: implications for the mucopolysaccharidoses and other connective tissue diseases. *Am J*

- Pathol*, 2008. 172(1): p. 112-22.
89. Ma, Y., R.P. Iyer, L.E. de Castro Brás, et al., Chapter 4 - Cross Talk Between Inflammation and Extracellular Matrix Following Myocardial Infarction, in *Inflammation in Heart Failure*, W.M. Blankesteyn and R. Altara, Editors. 2015, Academic Press: Boston. p. 67-79.
 90. Wu, X., J. Zhao, Y. Ruan, et al., Sialyltransferase ST3GAL1 promotes cell migration, invasion, and TGF- β 1-induced EMT and confers paclitaxel resistance in ovarian cancer. *Cell Death & Disease*, 2018. 9(11): p. 1102.
 91. Bertram, K.L., U. Banderali, P. Taylor, et al., Ion channel expression and function in normal and osteoarthritic human synovial fluid progenitor cells. *Channels (Austin)*, 2016. 10(2): p. 148-57.
 92. Lewis, R. and R. Barrett-Jolley, Changes in Membrane Receptors and Ion Channels as Potential Biomarkers for Osteoarthritis. *Front Physiol*, 2015. 6: p. 357.
 93. Hampe, C.S., J.B. Eisengart, T.C. Lund, et al., Mucopolysaccharidosis Type I: A Review of the Natural History and Molecular Pathology. *Cells*, 2020. 9(8).
 94. Erickson, G.R., L.G. Alexopoulos, and F. Guilak, Hyper-osmotic stress induces volume change and calcium transients in chondrocytes by transmembrane, phospholipid, and G-protein pathways. *J Biomech*, 2001. 34(12): p. 1527-35.
 95. Erickson, G.R., D.L. Northrup, and F. Guilak, Hypo-osmotic stress induces calcium-dependent actin reorganization in articular chondrocytes. *Osteoarthritis Cartilage*, 2003. 11(3): p. 187-97.
 96. Kurita, T., H. Yamamura, Y. Suzuki, et al., The ClC-7 Chloride Channel Is Downregulated by Hypoosmotic Stress in Human Chondrocytes. *Mol Pharmacol*, 2015. 88(1): p. 113-20.
 97. Baum, H., K.S. Dodgson, and B. Spencer, The assay of arylsulphatases A and B in human urine. *Clin Chim Acta*, 1959. 4(3): p. 453-5.
 98. Bolger, A.M., M. Lohse, and B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 2014. 30(15): p. 2114-20.
 99. Dobin, A., C.A. Davis, F. Schlesinger, et al., STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 2013. 29(1): p. 15-21.
 100. Harrow, J., A. Frankish, J.M. Gonzalez, et al., GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*, 2012. 22(9): p. 1760-74.
 101. Liao, Y., G.K. Smyth, and W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 2014. 30(7): p. 923-30.
 102. Mi, H., A. Muruganujan, X. Huang, et al., Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat Protoc*, 2019. 14(3): p. 703-721.

Supplementary information

Table S1. Chondrogenic differentiation protocol for stages 1–3. Revitacell; RevitaCell™ Supplement (100X); VN: Vitronectin XF; VN:Gel0,1%: Vitronectin XF/Gelatin 0,1%

| Day | CHIR 99021 | Activin-A | FGF2 | BMP2 | SB431542 | GDF5 | Revitacell | Media | Coating | Split |
|-------------|------------|-----------|----------|---------|-----------|----------------------------|---------------|-------|------------|------------|
| 0 | | | | | | Plate out 500k to 1M cells | | | | |
| 1 | 2 μ M | 50 ng/mL | | | | | | 2 mL | VN | |
| 2 | 2 μ M | 25 ng/mL | 40 ng/mL | | | | | 2 mL | VN | |
| 3 | 2 μ M | 10 ng/mL | 40 ng/mL | 5 ng/mL | | | | 2 mL | VN | |
| 4 | | | 40 ng/mL | 5 ng/mL | | | | 2 mL | VN | |
| 5 | | | 40 ng/mL | 5 ng/mL | 1 μ M | | 10 μ L/mL | 2 mL | VN | 1/5 to 1/8 |
| 6 | | | 40 ng/mL | 5 ng/mL | 1 μ M | | | 2 mL | VN | |
| 7 | | | 40 ng/mL | 5 ng/mL | 1 μ M | | | 2 mL | VN | |
| 8 | | | 40 ng/mL | 5 ng/mL | 1 μ M | | 10 μ L/mL | 2 mL | VN:Gel0,1% | 1/4 to 1/6 |
| 9 | | | 40 ng/mL | 5 ng/mL | | 20 ng/mL | | 2 mL | VN:Gel0,1% | |
| 10 | | | 40 ng/mL | 5 ng/mL | | 20 ng/mL | | 2 mL | VN:Gel0,1% | |
| 11 | | | 40 ng/mL | 5 ng/mL | | 40 ng/mL | | 2 mL | VN:Gel0,1% | |
| 12 | | | 40 ng/mL | 5 ng/mL | | 40 ng/mL | | 2 mL | VN:Gel0,1% | |
| 13 | | | 40 ng/mL | 5 ng/mL | | 40 ng/mL | | 2 mL | VN:Gel0,1% | |
| 14 | | | 40 ng/mL | 5 ng/mL | | 40 ng/mL | | 2 mL | VN:Gel0,1% | |
| Termination | | | | | | | | | | |

We adapted the chondrogenic differentiation protocol for hiPSCs based on our previous protocol (Oldershaw et al., 2010) to increase its robustness. The highly unstable Wnt3a was replaced by the small molecule CHIR-99021, a selective GSK-3 α/β inhibitor and a Wnt/ β -catenin signaling pathway activator. For activin inhibition in stage two of the protocol, Follistatin was replaced by the small molecule SB431542, which is a more selective activin inhibitor (Chen et al., 2012). BMP4 was replaced with BMP2, which has previously been shown to be more potent at driving chondrogenic differentiation (Wang et al., 2019).

Supplementary information**Table S2.** Primers used for genotyping gene edited hiPSCs.

| Name | Sequence 5'-3' | Company |
|---------------|-----------------------|---------|
| Set 1 forward | TCCCAGGGCCGGTTAATGT | IDTDNA |
| Set 1 reverse | GCTCTGGGCGGAGGAA-TATG | IDTDNA |
| Set 2 forward | CCTGAGTCCGGACCACTTTG | IDTDNA |
| Set 2 reverse | CACCGGTTCAATTGCCGAC | IDTDNA |

LV-OSKM lentivirus production

HEK293T cells were cultured in 10 cm culture dishes with DMEM high glucose (Gibco) supplemented with 10% fetal bovine serum (Thermo Scientific) and 100 U/ml Penicillin/Streptomycin/Glutamine (Gibco). At 80% confluency, cells were transfected with 3 µg LV-OSKM reprogramming vector, 2 µg psPAX2 and 1 µg pVSV using Fugene 6 transfection reagent according to manufacturer's protocol (Promega). Medium was filtered with 0,45 µm PDFV filters (Millipore) 72 hours post transfection and concentrated by centrifugation for 2 hours at 20k rpm with a Beckman Coulter Ultracentrifuge with SW32 Ti rotor. The virus was dissolved in 100 µl concentration in DMEM low glucose (Gibco) and stored at -80 °C.

hiPSCc

hiPSCs were generated from patient-derived primary fibroblasts as described before (van der Wal et al., 2018). Fibroblast cells from four patients with a rapid disease progression were reprogrammed using a polycistronic lentiviral vector of Oct4, Sox2, Klf4, and c-Myc (LV-OSKM). hiPSCs were cultured on γ -irradiated mouse embryonic feeder (MEF) cells with hiPSC culture medium consisting of DMEM/F12 medium (Invitrogen), 20% knock-out serum replacement (Invitrogen), 1% non-essential amino acids (Gibco), 1% penicillin/streptomycin/L-glutamine (100x, Gibco), 2 mM β -mercaptoethanol (Invitrogen) and 20 ng/ml basic fibroblast growth factor 2 (Peprotech). After gene correction, the hiPSCs were transferred to a feeder free culture with Vitronectin XF (Stem Cell) as coating and mTeSR™ Plus (Stem Cell) as media. Healthy control hiPSCs were obtained from the HIPSCI database and cultured with

Supplementary information

Vitronectin XF (Stem Cell) as coating and mTeSR™ Plus (Stem Cell) as media. Mycoplasma tests were routinely performed on all cell lines using the MycoAlert™ Mycoplasma Detection Kit (Lonza) and were found negative.

Gene correction

The donor construct to express the healthy copy of the ARSB gene was generated as described previously (van der Wal et al., 2018, In 't Groen et al., In press), with the ARSB cDNA instead of the GAA cDNA. CRISPR/Cas9 was used to insert the transgene into the AAVS1 locus. Gene editing was performed as described previously (In 't Groen et al., In press, van der Wal et al., 2018). In short, optimal target sites for the AAVS1 locus were selected using the CRISPRscan program (Moreno-Mateos et al., 2015). A vector containing the U6 promoter (addgene: 41824) was used to express the CCACTAGGGACAGGATTGGTGA sgRNA. hiPSCs on feeders were pretreated with RevitaCell (Thermo Fisher Scientific) 4 hr prior nucleofection. Single cells were generated from hiPSC colonies by incubating with TrypLE (Thermo Fisher Scientific). A DNA mix was prepped consisting of 4 mg of pCAG-hCAS9-GFP (addgene: 44719), 3 mg of TOPO-sgRNA, and 2 mg of donor vector. 2×10^6 cells were nucleofected with the DNA mix using Amaxa Human Stem Cell Nucleofector Kit2 (VPH-5022, Lonza) with program B-016. Cells were plated in hiPSC-conditioned medium (hiPSC medium incubated for 24 hr on feeder cells) supplemented with 20 ng/mL FGF2 (Preprotech) and RevitaCell to improve survival. 100 mg/mL G-418 (Invivogen) selection was initiated 48 hr after nucleofection. Approximately 14 days after initiation of selection, single IPS colonies were picked and genotyped using primers from **Table S2**.

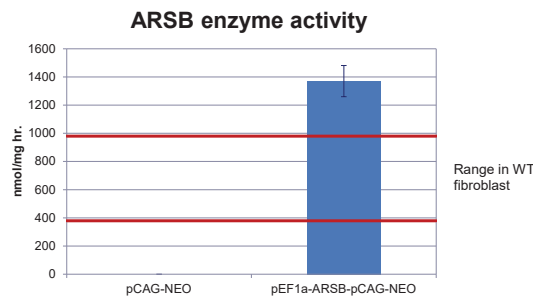


Figure S1. ARSB enzyme activity in HeLa Tk- cells following transfection of pEF1a-ARSB-pCAG-NEO. ARSB enzyme activity was measured in HeLa Tk- cells after selection with G418. No background activity was measured after transfection with pCAG-Neo as a negative control. The red solid lines indicate the wild type ARSB enzyme activity levels in fibroblasts.

Supplementary information

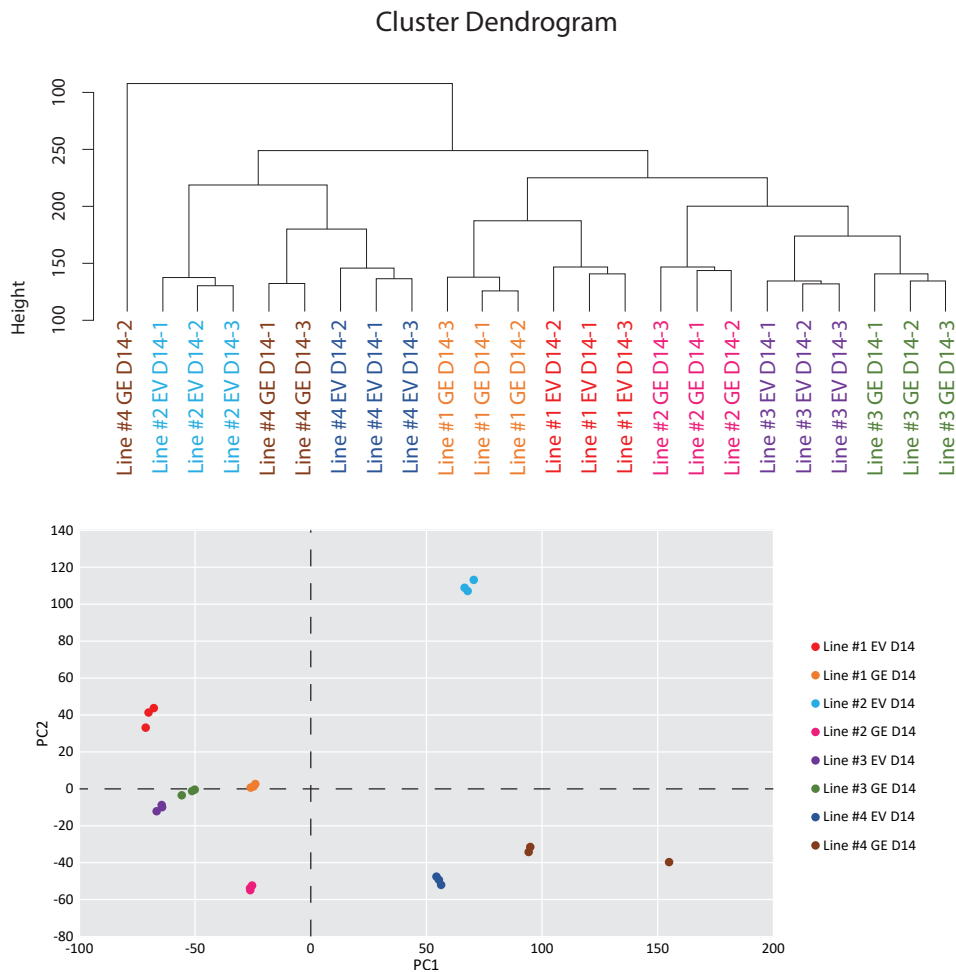


Figure S2. Cluster analysis and principal component analysis of all lines. Colors represent the triplicate for each line, GE: Gene edited, EV: Empty vector control.

Supplementary information

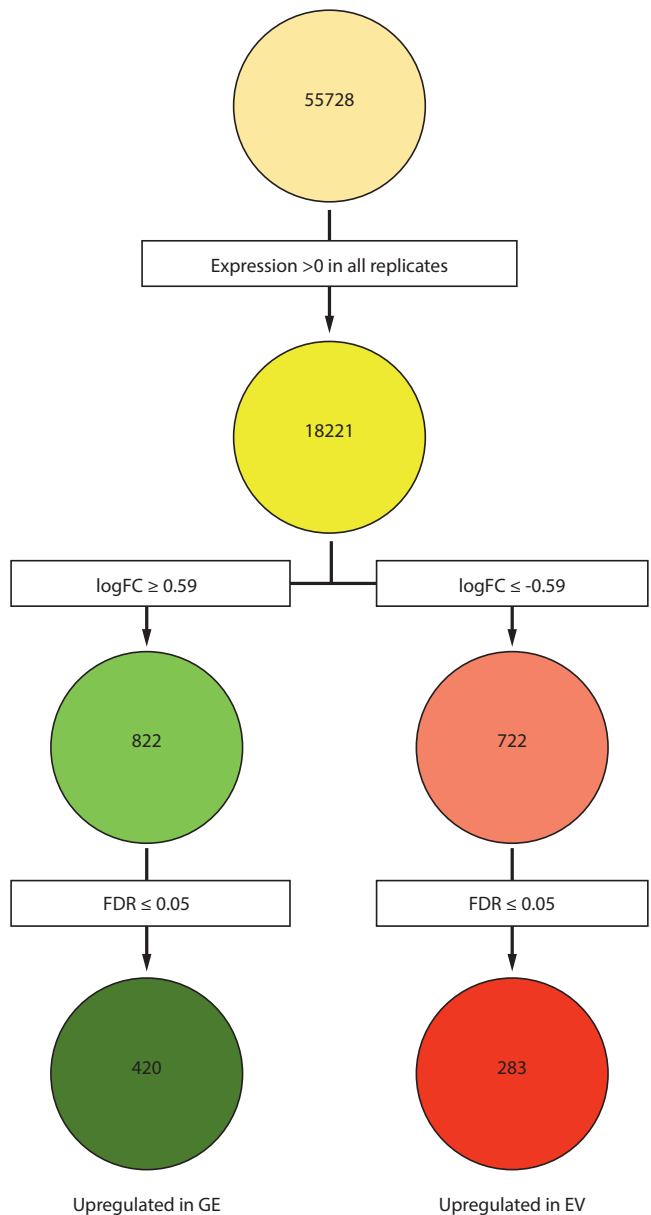


Figure S3. Flowchart of data analysis. Genes with a CPM of 0 in one or more samples were filtered out. Selection of genes with a $2^{\log_{FC}}$ of >0.59 and an FDR of <0.05 resulted in the identification of 420 genes that were upregulated in gene edited cells. Selection of genes with a $2^{\log_{FC}}$ of <-0.59 and an FDR of <0.05 resulted in the identification of 283 upregulated genes in empty vector cells. GE: Gene edited, EV: Empty vector control.

Supplementary information

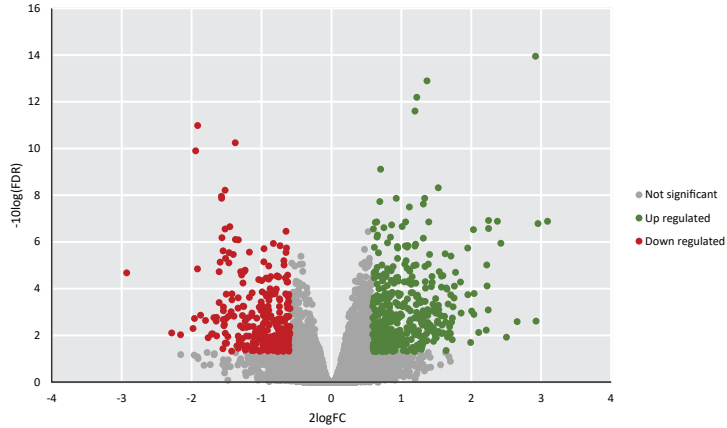
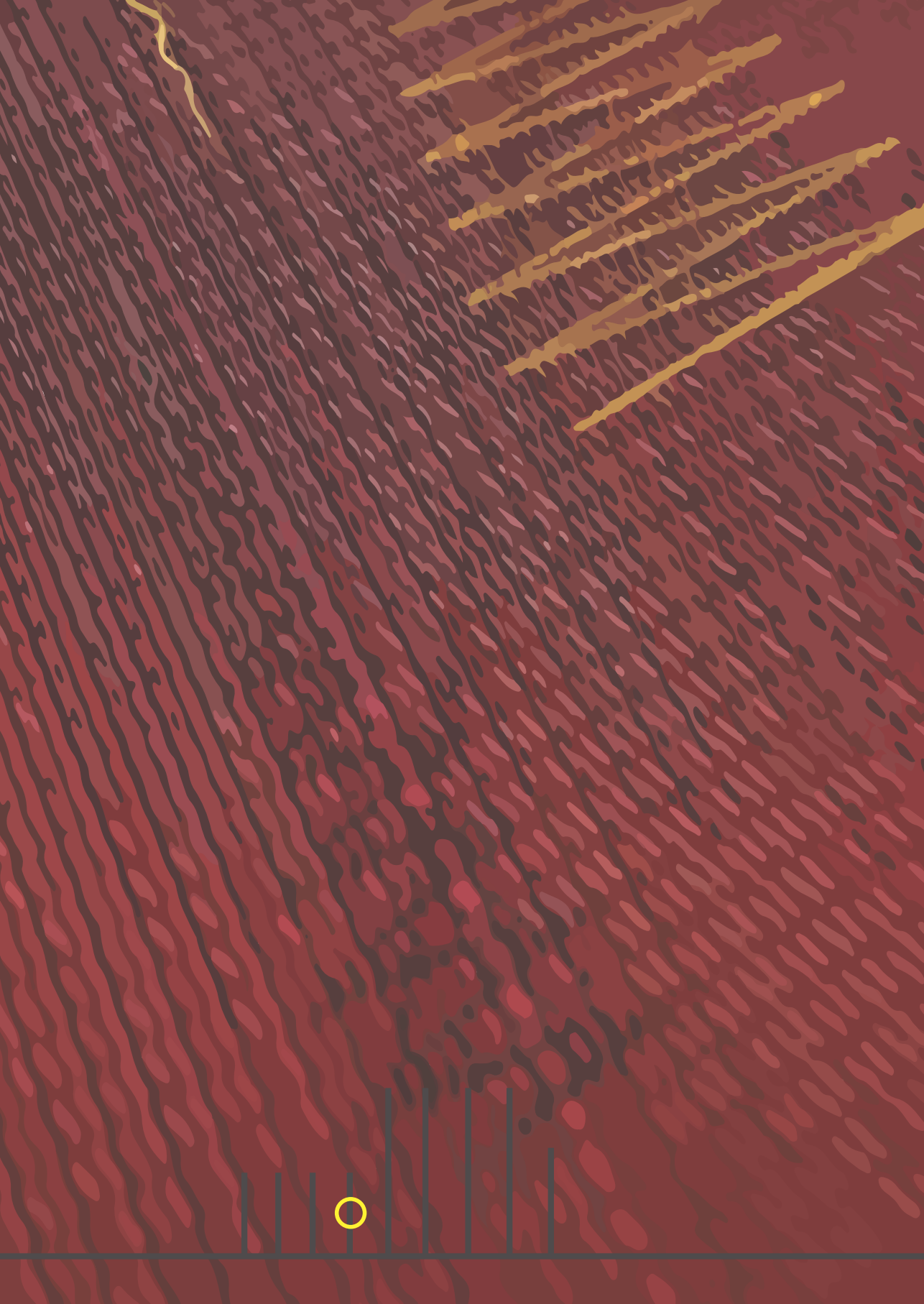
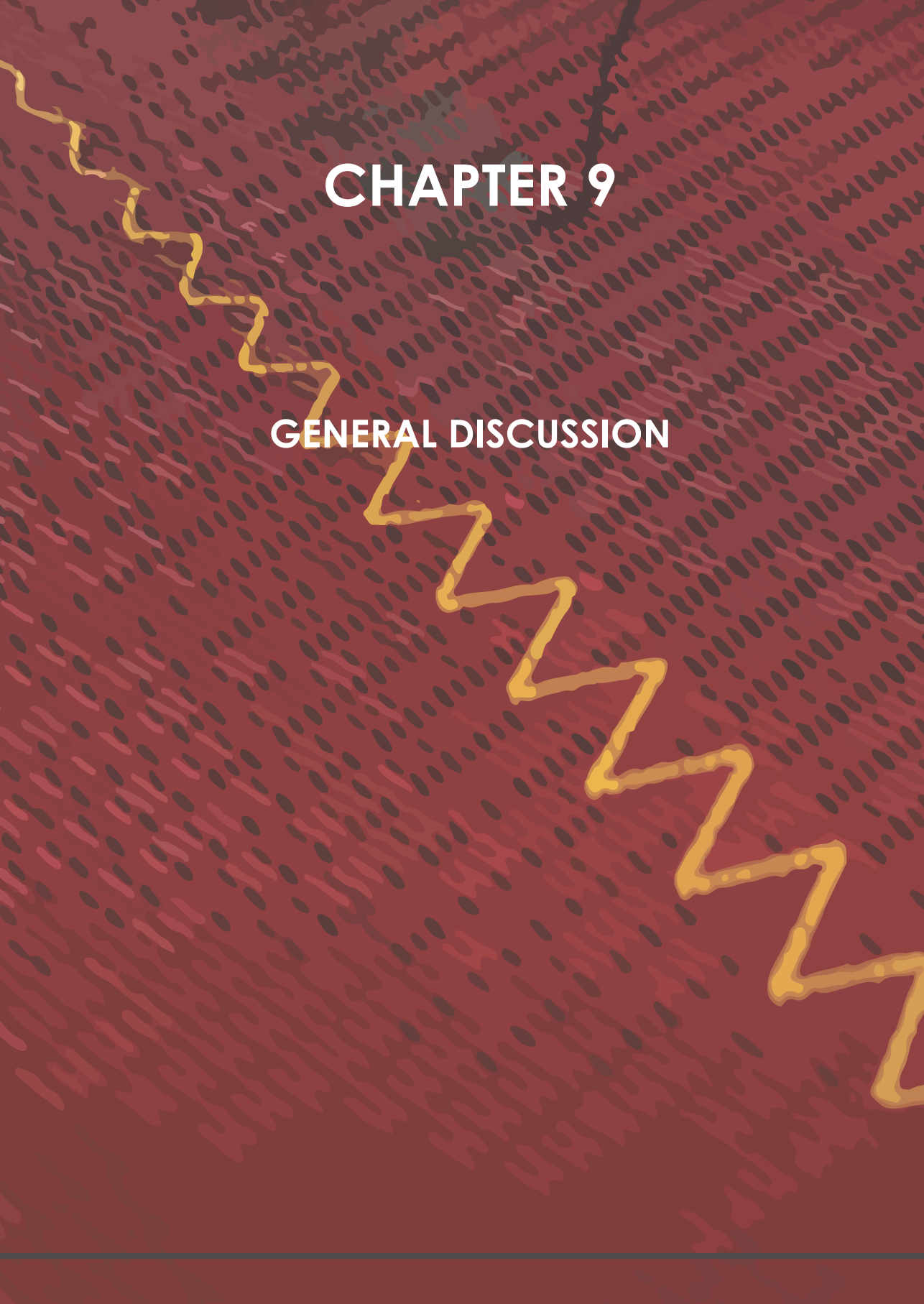


Figure S4. Volcano plots of all samples together and individual isogenic pairs. Each dot represents a single gene, those in red have a $2^{\log_{FC}}$ of <-0.59 and an FDR of <0.05 , those in green have a $2^{\log_{FC}}$ of >0.59 and an FDR of <0.05 . Not significantly changed genes are displayed in gray.



CHAPTER 9

GENERAL DISCUSSION



General Discussion

In this thesis we have addressed several aspects of research on MPS VI and genetic disorders, including: improved diagnostics, modeling of cartilage pathology, development of methods for CRISPR/Cas9 mediated gene editing, technological developments of gene editing, applications of gene editing in the clinic, and methods for splicing analysis and modulation for therapies for genetic disorders.

Splicing analysis for improved diagnostics and novel therapies

The identification of the disease-associated variant is an important aspect of diagnosis, predicting disease severity and genetic counseling. In the current diagnostic settings, genomic DNA sequence analysis of the coding and partial intronic regions of the gene of interest are used to identify the disease-associated variants of monogenic disorders.

To discriminate between disease-associated variants and other variants, databases with characterized variants and *in silico* predictions are used. *In silico* predictions include the effect of variants on protein folding and stability, and pre-mRNA splicing. The algorithms used for splicing can predict the effect of variants on the 5' and 3' splice sites, as well as on splice silencer and enhancer elements. Given the fact that splicing is complex and hard to predict, these algorithms might fail to predict the correct effect of variants on splicing. In addition, even with a clear effect on splicing, the splicing outcome is often not clear. For instance, the loss of a splice donor can result in the skip of that exon or the inclusion of (part) of the intron. As these outcomes have a profound effect on the mRNA stability and residual enzyme activity, determining the functional effect of variants is crucial in characterizing the variants. As MPS VI disease progression can vary from slow to rapid disease progression, a full characterization of disease-associated variants for MPS VI is important to unravel the genotype to phenotype correlation. This is particularly important as newborn screening is now available for multiple types of MPS, including MPS VI [1].

Currently, the diagnosis is confirmed on protein level using functional diagnostic assays such as enzymatic activity measurements or Western-blot. However, these assays do not provide additional insight on the functional outcome of disease-associated variants on RNA level. With the introduction of whole exome sequencing or whole genome sequencing, more variants will be identified, but it

will remain a challenge to confirm the diagnosis without functional data on novel variants. Splicing variants are likely underrepresented in MPS VI with only ~5% of all ARSB variants linked to splicing. Implementation of functional assays to assess variants at the level of pre-mRNA is required in the diagnostic setting. Full characterization of all variants is important, as a more complete comprehensive analysis of all variants in ARSB will improve diagnosis. In addition, aberrant splicing could provide a further insight into disease pathology and could be a novel target for therapy.

In **chapter 3**, we show a functional assay to detect aberrant splicing in MPS VI and used this to characterize all genotypes in the Dutch population. By combining a splicing assay with a cycloheximide treatment to prevent nonsense mediated decay, novel information on multiple disease-associated variants was obtained and the diagnosis was confirmed on RNA level for one patient. In addition, several aberrant splicing products were detected regardless of the presence of a disease-associated variant. These aberrant products showed inclusion of a pseudo exon from intron 5 or 6, and might provide a target for an RNA based therapy. Further discussion on splicing analysis for improved diagnostics is provided in **chapter 2**.

Modulation of splicing

Splicing can be regulated by using antisense oligonucleotides (AONs) to block certain regions in the pre-mRNA. This approach is currently used in several pre-clinical studies and has shown to improve the splicing outcome [2-4]. The use of AONs to block the inclusion of the pseudo exons might increase the level of correctly spliced products. Initial experiments blocking the splice donor and acceptor of the pseudo exons with AONs led to a decrease of pseudo exon inclusion, although no overall increase in mRNA levels was observed. This might be explained by the large intron size of 46kb and 57kb for intron 5 and 6, respectively. Blocking the splice sites might lead to splicing to other cryptic splice sites, thereby not improving overall mRNA levels. As the presence of large introns in genes increase the complexity of splicing, strategies such as blocking one or two sites with AONs might fail to improve the splicing outcome. To improve the overall splicing to clinically relevant levels will be especially challenging, since the levels of aberrant splicing were very low. This approach might be an interesting therapy if the overall splicing and enzyme activity could be improved through modulation by AONs, due to their long-term stability and their potential for modification. Using AONs for specific disease-associated variants that affect splicing might be an alternative approach. However, this kind of therapy will only be applicable to a small number of patients and has to be tailored

to each variant. Thus, application of AONs to modulate splicing is dependent on the disease. In MPS VI, achieving clinically relevant levels of canonical splicing using AONs directed to natural pseudo exons seems difficult. However, in general splicing modulation provides a very attractive way to treat genetic disorders, either by correcting a frequent disease-associated variant or by targeting natural aberrant splicing. Besides AONs, splicing may be modulated by small molecules [5], or via gene therapy. An in-depth discussion on this topic is provided in **chapter 2**.

Modeling of MPS VI cartilage pathology

Although cartilage is one of the majorly affected tissues in MPS VI patients, the development of cartilage pathology remains poorly understood. In **chapter 8**, we have developed an *in vitro* model system to study the MPS VI disease pathology in chondrogenic cells. As the high variability of gene expression and functional parameters between individuals can cause problems in the analysis, isogenic pairs were generated from patient-derived hiPSCs using CRISPR/Cas9 using a method developed in **chapter 6**. The insertion of a healthy version of the ARSB gene lead to an overexpression of the ARSB enzyme. The isogenic pairs were differentiated in a 14-day protocol and a clear increase of enlarged lysosomes was observed in MPS VI chondrogenic cells. RNA sequencing showed the dysregulation of several signal transduction pathways involved in bone and cartilage development, Wnt signaling, cell growth and apoptosis, ion transport and metabolic processes. The dysregulation of cartilage development and homeostasis may provide novel targets for therapy, which may be tested for efficacy and safety in personalized hiPSC-derived disease models for MPS VI.

Among the most promising druggable targets that we identified was the WNT signaling pathway. WNT signalling is involved in cell proliferation, differentiation and apoptosis and therefore seems like a promising target. However, no specific Wnt-targeting drugs are FDA approved yet and several WNT specific therapeutic agents have only recently entered clinical trials [6]. A few FDA-approved drugs do affect WNT signalling, although in a nonspecific fashion. Lithium salts are used to treat psychiatric disorders such as manic depression, and stabilize β -catenin and activate WNT signalling by GSK3 inhibition [7]. Non-steroidal anti-inflammatory drugs such as Sulindac and Celecoxib act on COX, key enzyme in the arachidonic acid cascade, and lead to WNT inhibition [8-12]. Pyrvinium is FDA approved as an anthelmintic drug and has been identified as a strong WNT inhibitor through β -catenin stabilisation and inhibition of Axin degradation [13-15]. Future research is

required to test the response of MPS VI cartilage to the specific and nonspecific WNT therapeutic agents for putative treatments.

Although a clear pathology was observed in chondrogenic cells in the *in vitro* model, these chondrogenic cells are cultured in a monolayer that does not resemble the native milieu of chondrocytes in cartilage tissue. It would therefore be of great interest to develop a 3-D culture method to form cartilage tissue using these patient-derived chondrogenic cells. Cartilage pellets can be generated from MSCs and are widely used in cartilage research [16, 17]. We have used a similar protocol to generate cartilage using the hiPSC-derived chondrogenic cells. Although we could generate healthy cartilage (data not shown), the reproducibility was poor and further refinement of the protocols is necessary to enable the comparison between lines and isogenic pairs. Different differentiation protocols can be explored for their potential in generating cartilage pellets. Promising protocols that show successful generation of cartilage pellets focus on the direct differentiation of iPSC to chondrogenic cells using growth factors [18, 19], differentiation through embryonic body formation [20] or through the generation of induced mesenchymal stem cells [21].

The generation of cartilage pellets would resemble the native milieu of the chondrocytes more closely. However, they lack multiple factors important for cartilage homeostasis, such as mechanical stimulation. The development of more complex 3-D culturing systems such as a cartilage-on-chip models would enable the introduction of other factors present in native cartilage, such as mechanical stimuli. Recently, several groups have developed cartilage-on-chip models to study osteoarthritis. Rosser et al. show an equine arthritis cartilage-on-chip model that mimics articular chondrocyte morphology, cell distribution, metabolism and gene expression using a physiologic nutrient diffusion gradient [22]. Occhetta et al. use human articular chondrocytes for a cartilage-on-chip model with mechanical stimulation to study osteoarthritis pathophysiology [23]. A different group is developing a cartilage-on-chip model to include mechanical stimulations on chondrocyte-laden hydrogels [24]. Lin et al. used hiPSC-derived mesenchymal progenitor cells to develop a micro-physiological osteochondral tissue chip with a layer of bone and a layer of cartilage tissue [25]. The development of a cartilage-on-chip model for MPS VI would not only enable to study the disease pathology in a more native milieu, it would also enable to study the effect of mechanical stimuli on MPS VI cartilage tissue. In addition, the introduction of inflammatory cytokines to a cartilage-on-chip model would allow us to study the effect of inflammation on the development and progression of the MPS VI cartilage pathology. Advanced and

personalized 2D cell models such as described in **chapter 8** and future advanced 3D models are expected to be of great value to mimic the patient's response. The higher predictive value of such disease models may prevent failure in phase III clinical trials and predict success of novel treatment options.

Current and future therapies for MPS VI

The current therapy for MPS VI consists of enzyme replacement therapy (ERT), in which recombinant ARSB enzyme is administered intravenously. Although some symptoms are alleviated, ERT is ineffective in treating the cartilage pathology. As cartilage is a non-vascularized tissue, all treatments that rely on intravenous delivery will have difficulties in reaching the chondrocytes in the cartilage. Hence, therapies such as intravenously administered AONs, retroviral gene therapy [26] or AAV gene therapy [27, 28] (currently in clinical trial) might be able to reduce urinary GAG levels and GAG accumulation in visceral tissues, but will likely be ineffective in treating the cartilage pathology. Chondrocytes get their nutrients from synovial fluid and the exchange of nutrients from the blood to the synovium is limited. An intra-articular therapy might be more effective in treating the cartilage pathology.

Intra-articular administration in combination with the long-term stability of AONs might prove a promising strategy to treat cartilage pathology. In addition, AONs can be modified to improve the diffusion through the ECM of the cartilage. However, finding a general target for AON treatment is challenging and specific AON treatment for disease-associated variants that affect splicing will exclude the majority of patients. Other limitations of AONs for clinical therapy include delivery and toxicity [29]. However, great improvements are made in recent years to improve delivery and reduce toxicity [2, 30] and the first AON based therapies have been approved by the FDA for duchenne muscular dystrophy and spinal muscular atrophy [31], showing the great potential of AON based therapy for the future.

Intra-articular injections of ERT showed promising results in pre-clinical studies, clearing GAG accumulation in MPS VI feline model [32, 33]. However, clinical application might be challenging as it would involve frequent intra-articular injections over a patient's lifetime. Intra-articular gene therapy might be a better approach for a novel therapy. Intra-articular lentiviral gene therapy showed a widespread transduction of synovial membrane cells, although *in vivo* transduction of chondrocytes was poor [34]. Supported by the results with intra-articular injections of the recombinant enzyme, localized overproduction of ARSB in the joint

might facilitate diffusion of enzyme into cartilage and uptake by chondrocytes, but no functional data was shown to support this. Given the lower risks involved, intra-articular AAV gene therapy might be another potential strategy. Pre-clinical studies showed that AAV gene therapy is able to transduce cells in the synovium and cartilage in mice as well as in equine joints [36-38]. As these studies were not performed in the context of MPS VI, it would be interesting to test the efficacy of delivering ARSB enzyme to joints of a MPS VI animal model.

Regenerative medicine would be an alternative approach to treat the cartilage pathology by replacing diseased tissue with healthy tissue. In **chapter 8**, we have shown that the pathology can be reversed in patient-derived hiPSCs by gene correction and chondrogenic cells can be generated through a refined differentiation protocol. This form of *ex vivo* gene editing provides a promising strategy in general for the replacement of disease tissue, not only in cartilage but also in other tissues such as bone marrow, immune cells and skeletal muscle. Various aspects of gene editing with respect to technological improvements, applications in the clinic, and a protocol for application in hiPSCs are discussed in **chapters 4, 5 and 7**.

Regenerative medicine using iPSCs in combination with gene editing tools holds great promise and can revolutionize the way patients are treated. For most genetic disorders, current therapy is based on surgical intervention or preventing and slowing down the symptoms caused by defects in the DNA. Gene editing allows the correction of the cause of these diseases, the DNA. With the recent technological advances in gene editing, such as CRISPR/Cas9, DNA itself becomes a “druggable” target. This is especially true if combined with iPSCs, as these allow for a virtually inexhaustible pool of cells with the ability to differentiate into any cell type of the human body. In addition, due to the high expansion capacity, a single iPSC cell can generate enough cells for a transplantation. This allows for a full genomic screening of cells to ensure safety. As the permanent DNA modifications can have adverse effects, gene corrected cells have to be validated to minimize the risk of genomic off-target effects. Genome wide DNA sequencing offers a platform to characterize these cells on DNA level. However, due to the high costs involved in the generation and culturing of patient derived iPSCs under good manufacturing practice (GMP), the labour-intensive safety validation of each clone and the transplantation procedure, the costs of an iPSC-based therapy could be a big burden on healthcare. Using HLA matched iPSCs to treat multiple patients would significantly lower costs and would notably reduce the time required to generate cells for transplantation into the patient [39]. The generation of a databank with gene

corrected iPSCs for different HLA matches would therefore greatly benefit a iPSC based regenerative therapy.

Engraftment of the cells is crucial for regenerative therapy. For the treatment of cartilage using stem cells, Cheng et al. have shown that cells that undergo a highly similar differentiation route can engraft into an osteochondral defect and contribute to the formation of healthy new cartilage [40]. As ARSB is naturally secreted, overexpression of the ARSB transgene does not only correct the targeted cells, but may also result in cross-correction of neighbouring chondrogenic cells in the cartilage upon engraftment of gene-corrected cells. Engraftment of gene-corrected chondrogenic cells would not only stop or delay the disease progression through cross correction, but also improve the cartilage pathology by the formation of new cartilage tissue. However, engraftment of the corrected chondrogenic cells would require surgical intervention into each affected joint, which might pose an obstacle for clinical application. In addition, chondrocytes rely on mechanical forces to distribute nutrients from the synovial fluid through the ECM of the cartilage, the lack of movement of MPS VI patients due to the poor joint mobility might accelerate the disease progression and hamper the delivery of enzyme through the synovial fluid to the chondrocytes.

Conclusion

The screening of variants in the current diagnostic setting provides limited insight into the effects of variants on splicing. An overview of improved diagnostics and novel treatment options for alternative splicing is shown in **chapter 2**. In **chapter 3** we developed a functional assay to detect aberrant splicing in MPS VI and characterize the effects of variants on ARSB splicing. These insights are important, because they can provide a novel therapeutic target. Implementation of the splicing assay in the diagnostic setting is recommended and will result in a better characterization of variants and their effects on splicing. In **chapter 4 and 5** we provide an overview of the recent technological advances in gene editing and the clinical trials involving gene editing. In **chapter 7** we have developed a gene editing method using CRISPR/Cas9 for hiPSC, which we applied to Pompe disease in **chapter 6**. In **chapter 8** we have developed an *in vitro* model system to study the MPS VI disease pathology in chondrogenic cells using hiPSCs and gene editing. Using isogenic hiPSC pairs, we generated an *in vitro* model system to study the cartilage pathology in patient derived chondrogenic cells. We showed a clear cellular pathology, which affects processes such as cell growth, apoptosis and

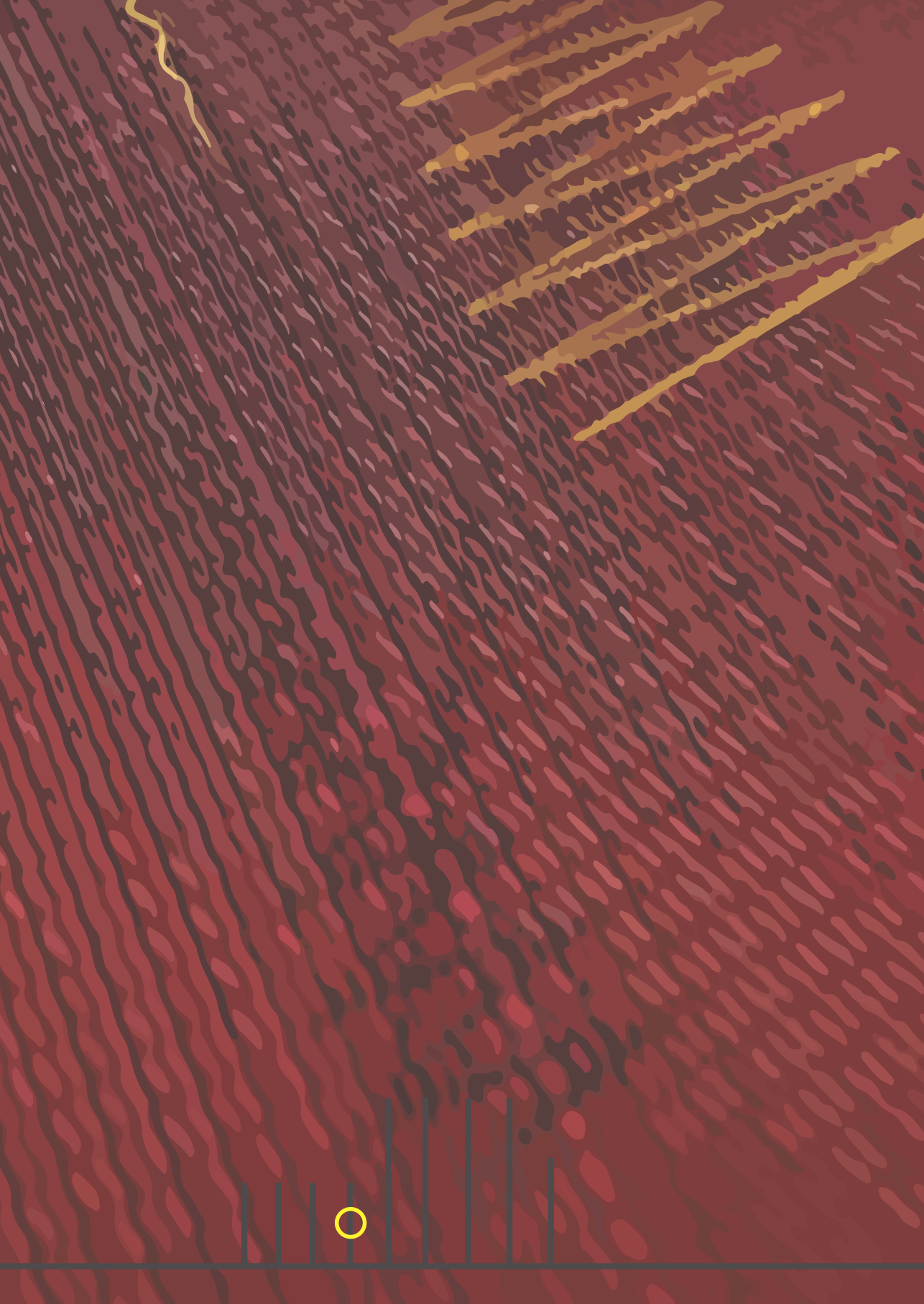
development of bone and cartilage. The generated gene corrected hPSCs also offer the potential for the development of a stem cell-based therapy, which would produce the ARSB enzyme locally in the affected joint. However, the engraftment of these cells into the affected tissue should be investigated further.

References

1. Chien, Y.H., et al., Newborn screening for Morquio disease and other lysosomal storage diseases: results from the 8-plex assay for 70,000 newborns. *Orphanet J Rare Dis*, 2020. **15**(1): p. 38.
2. Kuijper, E.C., et al., Opportunities and challenges for antisense oligonucleotide therapies. *J Inherit Metab Dis*, 2020.
3. Crooke, S.T., et al., RNA-Targeted Therapeutics. *Cell Metab*, 2018. **27**(4): p. 714-739.
4. Bennett, C.F., Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu Rev Med*, 2019. **70**: p. 307-321.
5. Palacino, J., et al., SMN2 splice modulators enhance U1-pre-mRNA association and rescue SMA mice. *Nat Chem Biol*, 2015. **11**(7): p. 511-7.
6. Kahn, M., Can we safely target the WNT pathway? *Nat Rev Drug Discov*, 2014. **13**(7): p. 513-32.
7. Klein, P.S. and D.A. Melton, A molecular mechanism for the effect of lithium on development. *Proc Natl Acad Sci U S A*, 1996. **93**(16): p. 8455-9.
8. Lee, H.J., et al., Sulindac inhibits canonical Wnt signaling by blocking the PDZ domain of the protein Dishevelled. *Angew Chem Int Ed Engl*, 2009. **48**(35): p. 6448-52.
9. Tai, W.P., et al., The inhibition of Wnt/beta-catenin signaling pathway in human colon cancer cells by sulindac. *Tumori*, 2014. **100**(1): p. 97-101.
10. Boon, E.M., et al., Sulindac targets nuclear beta-catenin accumulation and Wnt signalling in adenomas of patients with familial adenomatous polyposis and in human colorectal cancer cell lines. *Br J Cancer*, 2004. **90**(1): p. 224-9.
11. Egashira, I., et al., Celecoxib and 2,5-dimethylcelecoxib inhibit intestinal cancer growth by suppressing the Wnt/beta-catenin signaling pathway. *Cancer Sci*, 2017. **108**(1): p. 108-115.
12. Sareddy, G.R., et al., Nonsteroidal anti-inflammatory drugs diclofenac and celecoxib attenuates Wnt/beta-catenin/Tcf signaling pathway in human glioblastoma cells. *Neurochem Res*, 2013. **38**(11): p. 2313-22.
13. Xu, F., et al., Anthelmintic pyriminium pamoate blocks Wnt/beta-catenin and induces apoptosis in multiple myeloma cells. *Oncol Lett*, 2018. **15**(4): p. 5871-5878.
14. Saraswati, S., et al., Pyriminium, a potent small molecule Wnt inhibitor, promotes wound repair and post-MI cardiac remodeling. *PLoS One*, 2010. **5**(11): p. e15521.

15. Zhang, C., et al., Targeting of Wnt/beta-Catenin by Anthelmintic Drug Pyrvinium Enhances Sensitivity of Ovarian Cancer Cells to Chemotherapy. *Med Sci Monit*, 2017. **23**: p. 266-275.
16. Pittenger, M.F., et al., Multilineage potential of adult human mesenchymal stem cells. *Science*, 1999. **284**(5411): p. 143-7.
17. Johnstone, B., et al., In vitro chondrogenesis of bone marrow-derived mesenchymal progenitor cells. *Exp Cell Res*, 1998. **238**(1): p. 265-72.
18. Yamashita, A., et al. Generation of scaffoldless hyaline cartilaginous tissue from human iPSCs. *Stem Cell Reports*, 2015. **4**(3):404-18.
19. Boreström, C., et al. Footprint-free human induced pluripotent stem cells from articular cartilage with redifferentiation capacity: a first step toward a clinical-grade cell source. *Stem Cells Transl Med*, 2014. **3**(4):433-47.
20. Lee, J., et al. Early induction of a prechondrogenic population allows efficient generation of stable chondrocytes from human induced pluripotent stem cells. *FASEB J*, 2015. **29**(8):3399-410
21. Zou, L., et al. A simple method for deriving functional MSCs and applied for osteogenesis in 3D scaffolds. *Sci Rep*, 2013. **3**: 2243. 2013.
22. Rosser, J., et al., Microfluidic nutrient gradient-based three-dimensional chondrocyte culture-on-a-chip as an in vitro equine arthritis model. *Materials Today Bio*, 2019. **4**: p. 100023.
23. Occhetta, P., et al., Hyperphysiological compression of articular cartilage induces an osteoarthritic phenotype in a cartilage-on-a-chip model. *Nat Biomed Eng*, 2019. **3**(7): p. 545-557.
24. Paggi, C., et al., Cartilage-on-chip: a multi-modal platform to study human chondrocyte's response to mechanical stimuli. *Osteoarthritis and Cartilage*, 2020. **28**: p. S176-S177.
25. Lin, Z., et al., Osteochondral Tissue Chip Derived From iPSCs: Modeling OA Pathologies and Testing Drugs. *Front Bioeng Biotechnol*, 2019. **7**: p. 411.
26. Ponder, K.P., et al., Neonatal gene therapy with a gamma retroviral vector in mucopolysaccharidosis VI cats. *Mol Ther*, 2012. **20**(5): p. 898-907.
27. Tessitore, A., et al., Biochemical, pathological, and skeletal improvement of mucopolysaccharidosis VI after gene transfer to liver but not to muscle. *Mol Ther*, 2008. **16**(1): p. 30-7.
28. Cotugno, G., et al., Long-term amelioration of feline Mucopolysaccharidosis VI after AAV-mediated liver gene transfer. *Mol Ther*, 2011. **19**(3): p. 461-9.
29. Godfrey, C., et al., Delivery is key: lessons learnt from developing splice-switching antisense therapies. *EMBO Mol Med*, 2017. **9**(5): p. 545-557.
30. Bergsma, A.J., et al., Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options. *Int Rev Cell Mol Biol*, 2018. **335**: p. 85-141.

31. Stein, C.A. and D. Castanotto, FDA-Approved Oligonucleotide Therapies in 2017. *Mol Ther*, 2017. **25**(5): p. 1069-1075.
32. Auclair, D., et al., Intra-articular enzyme administration for joint disease in feline mucopolysaccharidosis VI: enzyme dose and interval. *Pediatr Res*, 2006. **59**(4 Pt 1): p. 538-43.
33. Auclair, D., et al., Long-term intra-articular administration of recombinant human N-acetylgalactosamine-4-sulfatase in feline mucopolysaccharidosis VI. *Molecular Genetics and Metabolism*, 2007. **91**(4): p. 352-361.
34. Byers, S., et al., Lentiviral-mediated correction of MPS VI cells and gene transfer to joint tissues. *Mol Genet Metab*, 2009. **97**(2): p. 102-8.
35. Evans, C.H., S.C. Ghivizzani, and P.D. Robbins, Gene Delivery to Joints by Intra-Articular Injection. *Hum Gene Ther*, 2018. **29**(1): p. 2-14.
36. Kyostio-Moore, S., et al., Local gene delivery of heme oxygenase-1 by adeno-associated virus into osteoarthritic mouse joints exhibiting synovial oxidative stress. *Osteoarthritis and Cartilage*, 2013. **21**(2): p. 358-367.
37. Watanabe, S., et al., Adeno-associated virus mediates long-term gene transfer and delivery of chondroprotective IL-4 to murine synovium. *Mol Ther*, 2000. **2**(2): p. 147-52.
38. Watson, R.S., et al., scAAV-mediated gene transfer of interleukin-1-receptor antagonist to synovium and articular cartilage in large mammalian joints. *Gene Ther*, 2013. **20**(6): p. 670-7.
39. Neofytou, E., et al., Hurdles to clinical translation of human induced pluripotent stem cells. *J Clin Invest*, 2015. **125**(7): p. 2551-7.
40. Cheng, A., et al., Cartilage repair using human embryonic stem cell-derived chondroprogenitors. *Stem Cells Transl Med*, 2014. **3**(11): p. 1287-94.



ADDENDUM

SUMMARY

SAMENVATTING

CURRICULUM VITAE

PHD PORTFOLIO

LIST OF PUBLICATIONS

ACKNOWLEDGEMENTS

Summary

Mucopolysaccharidosis type VI (MPS VI), also known as Maroteaux-Lamy disease, is an autosomal-recessive disorder caused by disease-associated variants in the arylsulfatase B (*ARSB*) gene. This leads to *ARSB* enzyme deficiency, which impairs degradation of the glycosaminoglycans (GAGs) dermatan sulfate and chondroitin sulfate. The progressive accumulation of these GAGs in lysosomes ultimately leads to cellular and organ deterioration.

The identification and characterization of disease-associated variants in genetic diseases is an important aspect of diagnosis, prediction of disease severity and genetic counseling. However, the effects of variants on pre-mRNA splicing and mRNA degradation are difficult to predict and often missed in current molecular diagnostics. Functional analysis of potential splicing variants is scarce which likely results in underestimation of the percentage of splicing variants involved in genetic diseases. In **chapter 2** we provide an overview of alternative splicing in genetic diseases, the current state of diagnostics to detect splicing variants and novel treatment options based on splicing correction. In **chapter 3**, we identified and characterized alternative splicing in MPS VI. We developed a generic assay for the unbiased identification and quantification of alternative splicing of *ARSB*. We found inefficient *ARSB* splicing due to natural skipping of exon 5 and the inclusion of two pseudoexons in introns 5 and 6 in healthy individuals. We analyzed 12 MPS VI patients with 10 different genotypes, which resulted in the identification of a 151-bp intron inclusion caused by the c.1142+2T>C variant, detection of low *ARSB* expression from alleles with the c.629A>G variant and a special case showed skipping of exon 4 and low *ARSB* expression. The characterization of a patient with an unknown disease-associated variant led to a molecular diagnosis at the RNA level. The inefficient natural splicing of *ARSB* may be a therapeutic target by promoting canonical splicing. These results highlight the importance of functional characterization of *ARSB* variants and their effect on pre-mRNA splicing in MPS VI patients.

In recent years, various gene editing platforms have been developed to allow alteration of the genome at specific loci to generate targeted genomic modifications. Gene editing is now widely applied in biomedical science and, due to the permanent nature of the modifications, holds great promise for the treatment of genetic diseases. **Chapter 4** provides an overview of the most recent technological advances in gene editing and the challenges to their clinical implementation. Gene editing strategies, off-target effects, emerging Cas9 alternatives and engineered variants with increased accuracy, the delivery of gene-editing components both ex

vivo and *in vivo*, and the remaining obstacles for large scale clinical implementation are discussed. In **chapter 5**, we provide an overview of clinical trials involving CRISPR-Cas9, TALENs, or ZFNs and discuss the gene editing strategies involved in these clinical trials. In **chapter 6**, we developed a gene editing strategy for the correction of genetic disorders such as lysosomal storage disorders. As proof of principle, we inserted the acid α -glucosidase cDNA into the *AAVS1* locus of human induced pluripotent stem cells (hiPSCs) using CRISPR/Cas9 and differentiated these cells into myogenic progenitors to model Pompe disease, a different lysosomal storage disease. This gene editing protocol was further refined in **chapter 7**; we provide protocols for CRISPR/Cas9-mediated gene editing in hiPSCs for the generation of gene knock outs, large deletions, and the introduction of a donor template in a safe harbor.

In **chapter 8**, we generated an *in vitro* disease model to study the cartilage pathology in MPS VI. Four patient-derived were reprogrammed into hiPSCs and isogenic controls were generated using CRISPR-Cas9 by insertion of wild type *ARSB* cDNA or an empty vector control. Differentiation of these hiPSCs into chondrogenic cells showed high expression of chondrogenic markers. Gene correction normalized the number of Periodic acid–Schiff (PAS) positive cells in chondrogenic cells compared to MPS VI chondrogenic cells. Genome-wide mRNA sequence analysis showed differentially expressed genes between isogenic pairs which were linked to processes and pathways including bone and cartilage development, Wnt signaling, metabolic processes, ion transport, cell growth and apoptosis. These results suggest a MPS VI-related pathology already at an early stage of chondrogenic development, and are in agreement with an early disease phenotype in cartilage. The hiPSC-derived model will provide a novel tool to test the response to treatment options and to study MPS VI disease pathology in chondrogenic cells.

Samenvatting

Mucopolysaccharidose type VI (MPS VI), ook bekend als de ziekte van Maroteaux-Lamy, is een autosomaal-recessieve ziekte en wordt veroorzaakt door ziektegerelateerde varianten in het arylsulfatase B (ARSB) gen. Dit leidt tot een ARSB enzymdeficiëntie die de afbraak van de glycosaminoglycanen (GAG's) dermatansulfaat en chondroitinesulfaat schaadt. De progressieve accumulatie van deze GAG's in het lysosoom leidt uiteindelijk tot cellulaire en orgaanafbraak.

De identificatie en karakterisering van ziektegerelateerde varianten is een belangrijk aspect van de diagnose, het voorspellen van de ernst van de ziekte en counseling bij genetische ziekten. De effecten van varianten op pre-mRNA-splicing en mRNA-afbraak zijn echter moeilijk te voorspellen en worden vaak over het hoofd gezien in de huidige moleculaire diagnostiek. Functionele analyse van mogelijke splicingsvarianten is schaars, wat waarschijnlijk resulteert in een onderschatting van het percentage splicingsvarianten dat betrokken is bij genetische ziekten. In **hoofdstuk 2** geven we een overzicht van alternatieve splicing bij genetische ziekten, de huidige staat van diagnostiek om splicingsvarianten te detecteren en nieuwe behandelingsopties op basis van splicingscorrectie. In **hoofdstuk 3** hebben we alternatieve splicing in MPS VI geïdentificeerd en gekarakteriseerd. We hebben een generieke test ontwikkeld voor de identificatie en kwantificering van alternatieve splicing van ARSB. We vonden inefficiënte ARSB-splicing als gevolg van het natuurlijk overslaan van exon 5 en de inclusie van twee pseudoexons in introns 5 en 6 bij gezonde controlepersonen. We hebben 12 MPS VI-patiënten met 10 verschillende genotypen geanalyseerd, wat resulteerde in de identificatie van een intron-inclusie van 151 bp veroorzaakt door de c.1142 + 2T> C-variant, detectie van lage ARSB-expressie van allelen met de c.629A> G-variant en een speciaal geval toonde het overslaan van exon 4 en lage ARSB-expressie. De karakterisatie van een patiënt met een onbekende ziektegerelateerde variant leidde tot een moleculaire diagnose op RNA-niveau. De inefficiënte natuurlijke splicing van ARSB kan een therapeutisch doelwit zijn door normale splicing te bevorderen. Deze resultaten benadrukken het belang van een functionele karakterisering van ARSB-varianten en de effecten op pre-mRNA splicing in MPS VI-patiënten.

In de afgelopen jaren zijn verschillende platformen voor het bewerken van genen ontwikkeld om gerichte genomische modificaties te genereren op specifieke loci van het genoom. Het bewerken van genen wordt nu op grote schaal toegepast in biomedische wetenschappen en houdt vanwege het permanente karakter van de modificaties een grote belofte in voor de behandeling van genetische ziekten.

In **hoofdstuk 4** wordt een overzicht gegeven van de meest recente technologische vooruitgangen in genbewerking en de uitdagingen voor hun klinische implementatie. Genbewerkingsstrategieën, off-target effecten, opkomende Cas9-alternatieven en doorontwikkelde varianten met verhoogde nauwkeurigheid, de aanlevering van genbewerking componenten zowel *ex vivo* als *in vivo*, en de resterende obstakels voor grootschalige klinische implementatie worden besproken. In **hoofdstuk 5** geven we een overzicht van klinische onderzoeken met CRISPR/Cas9, TALEN's of ZFN's en bespreken we de gebruikte strategieën voor het bewerken van genen in deze klinische onderzoeken. In **hoofdstuk 6** hebben we een strategie voor het bewerken van genen ontwikkeld voor de correctie van genetische aandoeningen zoals lysosomale stapelingsstoornissen. Als proof of principle, hebben we een gezonde kopie van zure alpha-glucosidase ingebracht in het AAVS1 locus van humane geïnduceerde pluripotente stamcellen (hiPSC's) met behulp van CRISPR/Cas9. Vervolgens hebben we deze cellen gedifferentieerd tot myogene voorlopers om de ziekte van Pompe (een andere lysosomale stapelingsziekte) te modelleren. Het protocol voor het bewerken van genen is verder verfijnd in **hoofdstuk 7**; waar we protocollen presenteren voor CRISPR/Cas9-gemedieerde genbewerking in hiPSC's voor het genereren van gen knock-outs, grote deleties en de introductie van donor DNA in een safe harbour.

In **hoofdstuk 8** hebben we een *in vitro* ziektemodel gegenereerd om de kraakbeenpathologie in MPS VI te bestuderen. Vier patiënt afgeleide fibroblasten cel lijnen werden gereprogrammeerd in hiPSC's en isogene controles werden gegenereerd met behulp van CRISPR/Cas9 door insertie van een gezonde versie van het ARSB cDNA of een lege vectorcontrole. Differentiatie van deze hiPSC's tot chondrogene cellen toonde een hoge expressie van chondrogene markers. Gencorrectie normaliseerde het aantal Periodic acid-Schiff (PAS)-positieve cellen in chondrogene cellen. Genoomwijde mRNA-sequentieanalyse toonde differentieel tot expressie gebrachte processen en signaleringsroutes tussen isogene paren, waaronder bot- en kraakbeenontwikkeling, Wnt-signaleringsroute, metabole processen, ionentransport, celgroei en apoptose. Deze resultaten suggereren een MPS VI-gerelateerde pathologie in een vroeg stadium van chondrogene ontwikkeling en ondersteunen de theorie van een vroeg ziektefenotype in kraakbeen. Het hiPSC-afgeleide model biedt een nieuw hulpmiddel om de respons op behandelingsopties te testen en om de pathologie van de MPS VI-ziekte in chondrogene cellen te bestuderen.

Curriculum vitae

Mike Broeders was born on September 24th 1988 in Tilburg. He started his educational career on the Prisma College Graaf Engelbrecht in Breda and finished his VWO in 2007. He studied Biomedical laboratory technology at the Avans Hogeschool in Breda. His bachelor internship was performed at Nickerson-Zwaan in Melle under supervision of Dr. Wessel van Leeuwen. After graduating in 2011 he worked as a research technician at Nickerson-Zwaan before starting his Master's in Life Science & Technology at the Leiden University in 2012. He performed his master internship in the lab of Prof. J.N.J. Philipsen in the Cell Biology department of Erasmus MC where he worked on CRISPR/Cas9-mediated genome engineering and their application to globin gene regulation. In 2014 he finished his master's degree and started his PhD under supervision of Prof. Dr. A.T. Van der Ploeg (promotor) and Dr. W.W.M. Pim Pijnappel (co-promotor). The research performed during his PhD led to multiple scientific publications and the thesis you are reading now. He started his postdoctoral research in October 2020 under supervision of Dr. Roberto Narsici in the orthopedics department in Erasmus MC.

PhD Portfolio

| | Year | ECTS |
|---|------|------|
| Courses: | | |
| Genetics course | 2015 | 3 |
| Laboratory animal science | 2017 | 3 |
| Biomedical English Writing and Communication | 2017 | 2 |
| Safely working in the laboratory | 2017 | 0,5 |
| Home office animal course (Manchester, UK) | 2017 | 2 |
| Research integrity | 2018 | 0,5 |
| Practical radiation protection level 5B | 2018 | 1 |
| Workshops: | | |
| 22th MGC PhD workshop, Maastricht | 2015 | 1 |
| 23th MGC PhD workshop, Dortmund Germany | 2016 | 1 |
| (Inter)National Meetings: | | |
| 5th Winterschool of the Collaborative Research Centre | 2015 | 2 |
| (TRR 81), Kleinwalsertal Austria | | |
| 8th Annual Meeting Dutch Society for Stem Cell Research (DSSCR), Utrecht | 2015 | 0,5 |
| MPS Masterclass, Istanbul | 2015 | 1 |
| Genome on demand? Exploring the implications of human genome editing, Amsterdam | 2015 | 0,5 |
| 9th Annual Meeting Dutch Society for Stem Cell Research (DSSCR), Utrecht | 2016 | 0,5 |
| 14 th international Symposium on Mucopolysaccharide and related diseases, Bonn Germany | 2016 | 1 |
| 10th Annual Meeting Dutch Society for Stem Cell Research (DSSCR), Utrecht | 2017 | 0,5 |
| WORLDSymposium, San Diego USA | 2018 | 1,5 |
| Erfelijke Stofwisselingsziekten Nederland (ESN) Voorjaarssymposium, Valkenburg aan de Geul | 2018 | 0,5 |
| 11th Annual Meeting Dutch Society for Stem Cell Research (DSSCR), Utrecht | 2018 | 0,5 |
| Sophia Research Day | 2018 | 0,5 |
| Erfelijke Stofwisselingsziekten Nederland (ESN) Najaarsymposium, Utrecht | 2018 | 0,5 |
| 22nd Annual meeting ASGCT, Washington USA | 2019 | 1,5 |
| 12th Annual Meeting Dutch Society for Stem Cell Research (DSSCR), Utrecht | 2019 | 0,5 |

| Presentations: | | |
|--|------|-------------|
| Oral presentation 5th Winterschool of the Collaborative Research Centre (TRR 81), Kleinwalsertal Austria | 2015 | 0,5 |
| Poster presentation WORLDSymposium, San Diego USA | 2018 | 0,25 |
| Oral presentation Erfelijke Stofwisselingsziekten Nederland (ESN) Voorjaarsymposium | 2018 | 0,5 |
| Oral presentation Sophia Research Day | 2018 | 0,5 |
| Poster presentation 22nd Annual meeting ASGCT, Washington USA | 2019 | 0,25 |
| Teaching Activities: | | |
| Supervisor of Bachelor student | 2017 | 2 |
| Supervisor of Bachelor student | 2018 | 2 |
| Other: | | |
| Organizing committee for labday out, Rotterdam | | 1 |
| Total | | 32,5 |

List of Publications

Bergsma AJ, van der Wal E, **Broeders M**, van der Ploeg AT, Pijnappel WWMP. (2018) *Alternative Splicing in Genetic Diseases: Improved Diagnosis and Novel Treatment Options*. *Int Rev Cell Mol Biol*. 2018;335:85-141. doi: 10.1016/bs.ircmb.2017.07.008

van der Wal E, Herrero-Hernandez P, Wan R, **Broeders M**, In 't Groen SLM, van Gestel TJM, van IJcken WFJ, Cheung TH, van der Ploeg AT, Schaaf GJ, Pijnappel WWMP. (2018) *Large-Scale Expansion of Human iPSC-Derived Skeletal Muscle Cells for Disease Modeling and Cell-Based Therapeutic Strategies*. *Stem Cell Reports*. 2018 Jun 5;10(6):1975-1990. doi: 10.1016/j.stemcr.2018.04.002.

Broeders M*, Herrero-Hernandez P*, Ernst MPT, van der Ploeg AT, Pijnappel WWMP. (2019) *Sharpening the Molecular Scissors: Advances in Gene-Editing Technology*. *iScience*. 2020 Jan 24;23(1):100789. doi: 10.1016/j.isci.2019.100789

Ernst MPT, **Broeders M***, Herrero-Hernandez P*, van der Ploeg AT, Pijnappel WWMP. (2020) *Ready for repair? Gene editing for the treatment of human disease*. *Mol Ther Methods Clin Dev*. 2020 Jul 3;18:532-557. doi: 10.1016/j.omtm.2020.06.022.

Broeders M, Smits K, Goynuk B, Oussoren E, van den Hout HJ, Bergsma AJ, van der Ploeg AT, Pijnappel WWMP. (2020) *A generic assay to detect aberrant ARSB splicing and mRNA degradation for the molecular diagnosis of MPS VI*. *Molecular Therapy: Methods & Clinical Development*, doi: <https://doi.org/10.1016/j.omtm.2020.09.004>.

Broeders M*, in 't Groen SLM*, Pijnappel WWMP. (2020) *CRISPR/Cas9-mediated gene editing in human induced pluripotent stem cells*. *CRISPR-Cas Methods: Volume II*. Nature Springer. In press

Broeders M, van Rooij JGJ, Oussoren E, van Gestel TJM, Smith CA, Kimber SJ, Verdijk RM, Wagenmakers MAEM, van den Hout JMP, van der Ploeg AT, Narcisi R, Pijnappel WWMP. (2021) *Generation of a disease model for cartilage pathology in MPS VI using patient-derived and isogenic gene-corrected hiPSCs*. Submitted

Niño Martinez NMY, Wijgerde M*, Oliveira Soares de Faria D*, Bergsma AJ, **Broeders M**, van der Beek A.M.E, van den Hout JMP, van der Ploeg AT, Verheijen FW, Pijnappel WWMP (2020) *Enzymatic Diagnosis of Pompe Disease: lessons from 28 years of experience*. *Eur J Hum Genet*. 2020 Nov 8. doi: 10.1038/s41431-020-00752-2.

Acknowledgements

First and foremost, I would like to thank my promoter and co-promoter Prof. dr. Ans van de Ploeg and Dr. Pim Pijnappel for their continuous support during these past years. Thanks for letting me develop as a scientist and as a person these last years.

I would like to thank Prof. dr. Joost Gribnau, Prof. dr. Marianna Tryfonidou Prof. dr. Edmond H.H.M. Rings for your time and effort to be in my reading committee, I hope you found it an interesting read! I would also like to express my gratitude to the other members of my committee: Prof. dr. Gerjo G.J.V.M. van Osch, Prof. dr. Sue. Kimber, Dr. Roberto Narcisi, Dr. Joyce.B.J. van Meurs and Drs. Esmee Oussoren, thank you all for taking the time to be in my committee.

To the current and former members of our lab: Atze, Douglas, Tom, Joon, Merel, Stijn, Gerben, Erik, Qiushi (Susan!), Monica, and the kids and babies Pablo, Rodrigo, Eva, Fabio & Ale (you really do grow up so quickly!) a massive thanks for all the ideas, activities, laughs and for making my PhD experience a really great one. A special thanks goes out to my own two students, Kasper and Büsra, I have learned a lot by supervising you both and I hope I repaid the favour and taught you a thing or two as well! Also thanks to all the other students we had in the lab, I am not going to mention you all, but some of you were actually quite decent 😊.

All my colleagues at the department of Clinical genetics, Pediatrics, Center for Lysosomal and Metabolic Diseases, our neighbours at Developmental biology, and other colleagues and collaborators: thanks for the input during our meetings, help with experiments, fun activities and all the chit chat. As I am a man of few words, I will use someone else's: "I don't know half of you half as well as I should like; and I like less than half of you half as well as you deserve".

To my friends, sisters, brothers in law, nieces and nephews, parents, and parents in law, thank you for all the support and love all these years and for keeping me (sort of) sane. Koda and Wendor, you two are the best dogs in the world. Hélène, I cannot express in words how grateful I am to have you by my side. I really could not have done this without your support, input, proofreading, patience, and unconditional love. I am looking forward to all the future adventures we will face together, I cannot think of a better person to face them with than you. Je t'aime.

