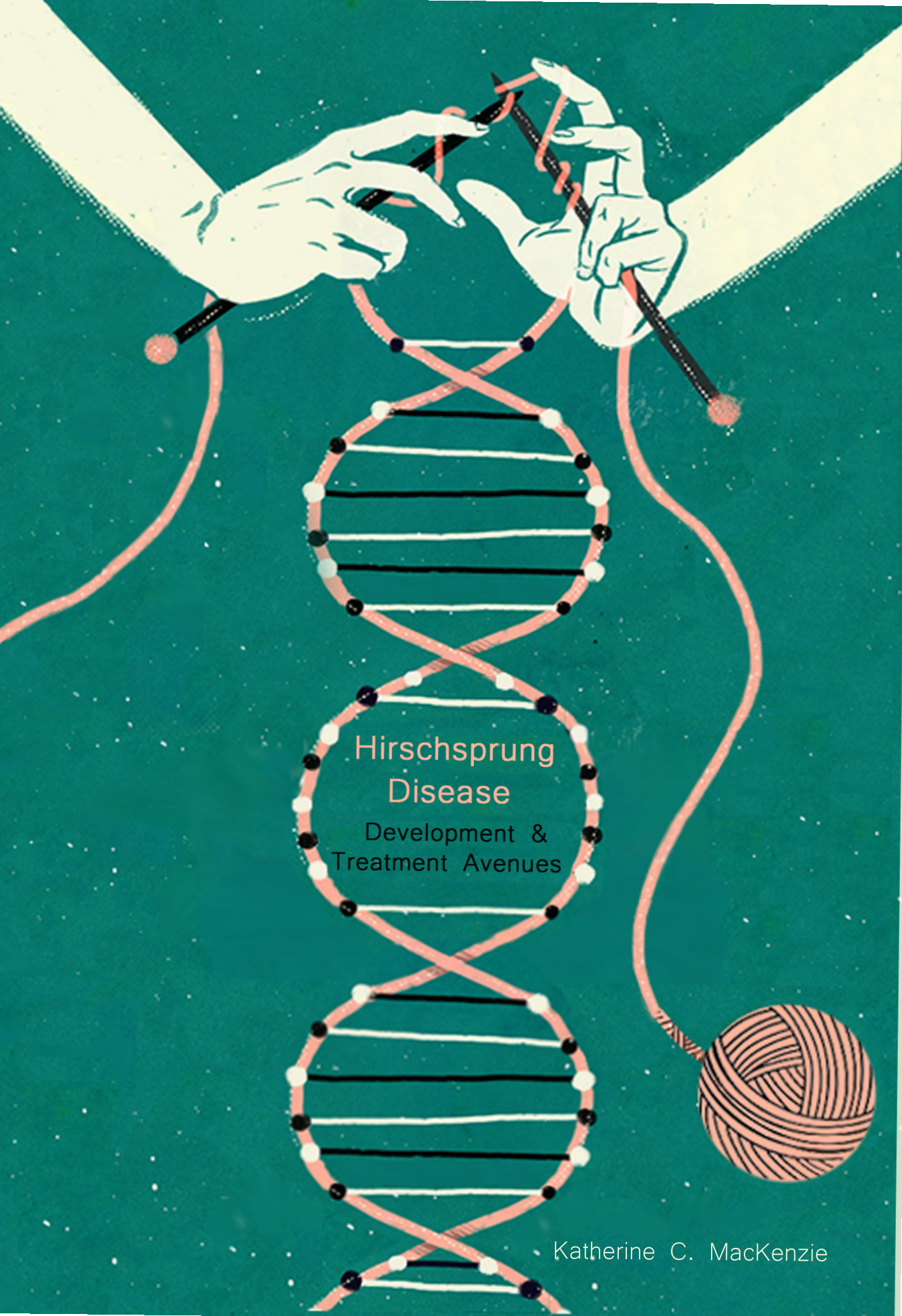

Hirschsprung
Disease
Development &
Treatment Avenues
Katherine C. MacKenzie

Hirschsprung Disease:

Development & Treatment Avenues

Katherine Christa MacKenzie

The work presented in this thesis was financially supported by Stichting Vrienden van het Sophia, Erasmus University Medical Centre, Erasmus Trustfonds, Simonsfonds. Thesis printing costs were sponsored by Erasmus University and the department of Clinical Genetics, Erasmus Medical Centre, Rotterdam.

The studies described in this thesis were performed in the departments of Clinical Genetics, Biomics and Paediatric Surgery in the Erasmus Medical Centre, Rotterdam, The Netherlands.

ISBN: 978-94-6323-646-1
Author: Katherine Christa MacKenzie

Cover Design: Matthew Forsythe and Katherine Christa MacKenzie

Layout: Katherine Christa MacKenzie

Printed By: Gildeprint, Enschede

Copyright © Katherine Christa MacKenzie, 2019. All rights reserved. No part of this thesis may be reproduced, stored in a retrieval system, or transmitted in any form by any means without the prior permission of the author.

Hirschsprung Disease:

development & treatment avenues

De ziekte van Hirschsprung:

ontwikkeling & behandelingsmogelijkheden

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam by command of the rector magnificus

Prof. Dr. R.C.M.E. Engels

and in accordance with the decision of the Doctoral Board.

The public defence will be held on

Tuesday 11th June 2019 at 11.30 hours

by

Katherine Christa MacKenzie
born in Alexandria, Scotland

Erasmus University Rotterdam

Doctoral Committee

Promoters	Prof. Dr. R.M.W. Hofstra
	Prof. Dr. R.M. Wijnen
Other Members	Prof. Dr. D. Tibboel
	Prof. Dr. I. de Blaauw
	Prof. Dr. N. Thapar
Co-promoters	Dr. A.J. Burns
	Dr. M.M. Alves
Paranymphs	Bianca M. de Graaf
	Martyna M. Grochowska

Table of Contents

List of Abbreviations

ARTN	Artemin
BP	Base Pair
CN	Copy Number
CNP	Copy Number Polymorphism
CNS	Central Nervous System
CNV	Copy Number Variation
DOF	Degree of Freedom
E	Embryonic Day
EAS	External Anal Sphincter
ECM	Extracellular Matrix
EDN3	Endothelin 3
EDNRB	Endothelin Receptor Type B
ENC	Enteric Neural Crest
ENCC	Enteric Neural Crest Cell
ENS	Enteric Nervous System
ESC	Embryonic Stem Cell
FCS	Foetal Calf Serum
FI	Faecal Incontinence
GABBR1	Gamma-aminobutyric acid B receptor 1
GDNF	Glial cell line-derived Neurotrophic Factor
GFRα1	Glial cell line-derived Neurotrophic Factor Receptor alpha 1
GI	Gastrointestinal
GNL1	Guanine nucleotide binding protein like 1
GOSHS	Goldberg-Shprintzen Syndrome
HEK	Human Embryonic Kidney
HSCR	Hirschsprung Disease
IAS	Internal Anal Sphincter
iPSC	Induced Pluripotent Stem Cell
KIF1BP	Kinesin Binding Protein/KIF1 Binding Protein
LOF	Loss of Function
MCAP	Megalencephaly-capillary malformation
MSC	Mesenchymal Stem Cell
MWS	Mowat-Wilson Syndrome
NCC	Neural Crest Cell
NGS	Next Generation Sequencing
NRG1	Neuregulin 1
NRTN	Neurturin
PI3K	Phosphatidylinositol-3-kinase
PGRS	Polygenic Risk Score

PSC	Pluripotent Stem Cell
PSPN	Persephin
RET	Rearranged During Transfection
RoH	Runs of Homozygosity
SCG10	Superior Cervical Ganglia 10
SEMA3A	Semaphorin 3A
SNP	Single Nucleotide Polymorphism
SNV	Single Nucleotide Variation
TCA	Total Colonic Aganglionosis
TEPT	Transanal Endorectal Pull-Through
USP32	Ubiquitin Specific Peptidase 32
WES	Whole exome sequencing
WS4	Waardenburg-Shah Syndrome
WT	Wild Type


1

Chapter 1

General Introduction

Hirschsprung disease (OMIM #142623) is a congenital malformation of the nervous system of the gastrointestinal tract, the Enteric Nervous System. The work in this thesis focuses on modes of development of this disease, *in vitro* disease modelling and possible therapeutic options.

The Enteric Nervous System

The gastrointestinal (GI) tract is the core of the digestive system and has many functions including the mixing and breaking down of ingested material, water and nutrient extraction and absorption, secretion of enzymes and propulsion of ingested material through the body for waste expulsion in defecation[1]. These functions require the input of various cell types communicating together in order to effectively digest food and avoid nutritional deficit or inflammatory responses from poor gut function. At birth, the GI tract, including the oesophagus, stomach, small intestine and colon, is approximately 3 meters in length, and increases to approximately 7 meters in adults[3]. The development of such an organ is complex and requires extensive elongation of the primitive gut tube, as well as migration of a wide variety of cell types along the gut, to ultimately result in co-ordinated gut activity[4]. Proper functioning of the GI system is regulated by the enteric nervous system (ENS).

The ENS is one of the three main divisions of the autonomic nervous system, along with the sympathetic and parasympathetic divisions. The ENS is also colloquially known as the second brain or the brain in the gut, due to its ability to function independently from the central nervous system (CNS). The basic functions of the ENS do not require input from the sympathetic and parasympathetic divisions, although it is influenced by both[5]. The ENS is made up of neuronal cells and supportive glial cells

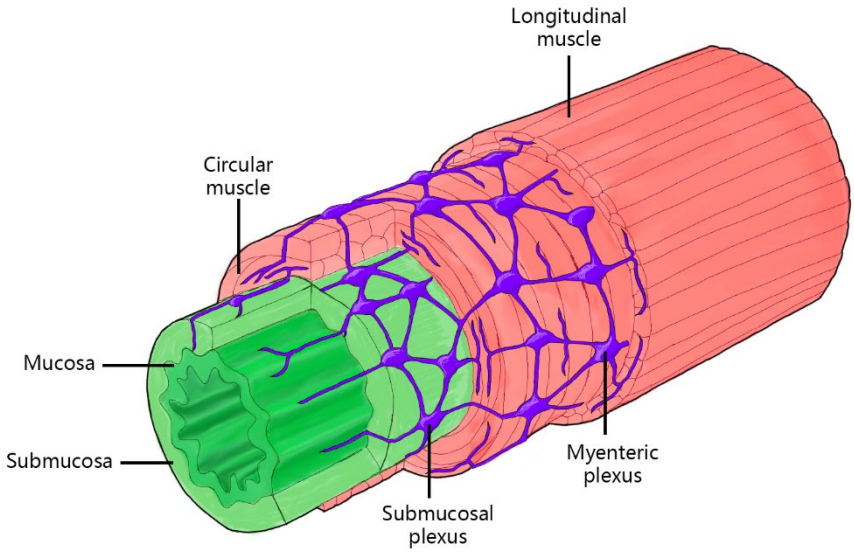


Figure 1. Schematic Cross-section of the adult GI tract showing the submucosal and myenteric plexuses either side of the circular muscle layer. Axons extend radially from the plexuses for communication between plexuses and into the gut mucosa.

located within ganglia, which form two distinct mesh-like plexuses in the wall of the GI tract (Figure 1).

The myenteric plexus spans the entire length of the GI tract and lies between the longitudinal and circular muscle layers. It controls muscle contraction and relaxation which generates the peristaltic movement of the gut. The submucosal plexus, located between the circular muscle and mucosa, is only prominently seen in the small and large intestines. It regulates fluid absorption and secretion, modulates blood flow, and responds to stimuli from the mucosa to support gut function and homeostasis[6].

Development of the ENS

The ENS is derived from the neural crest which arises at the border between the neural plate and the non-neural ectoderm. The majority of cells that are fated to become the ENS originate at the level of the vagal neural crest. These precursor cells, known as enteric neural crest cells (ENCCs), migrate to the cranial end of the gut tube at week 4 of human development (embryonic day [E]9.5 in mice)[7,8], and rapidly proliferate, migrate and differentiate to colonise the entire length of the gut by week 7 (E13.5 in mice) (Figure 2)[7,8]. A smaller population of neural cells also arises from the sacral region of the neural tube, but contributes to the ENS mainly in the distal hindgut[9,10]. In addition to the uniform directional (oral to anal) migration of vagal ENCCs along the gut, Nishiyama and colleagues showed that, in mice, ENCC migration halts

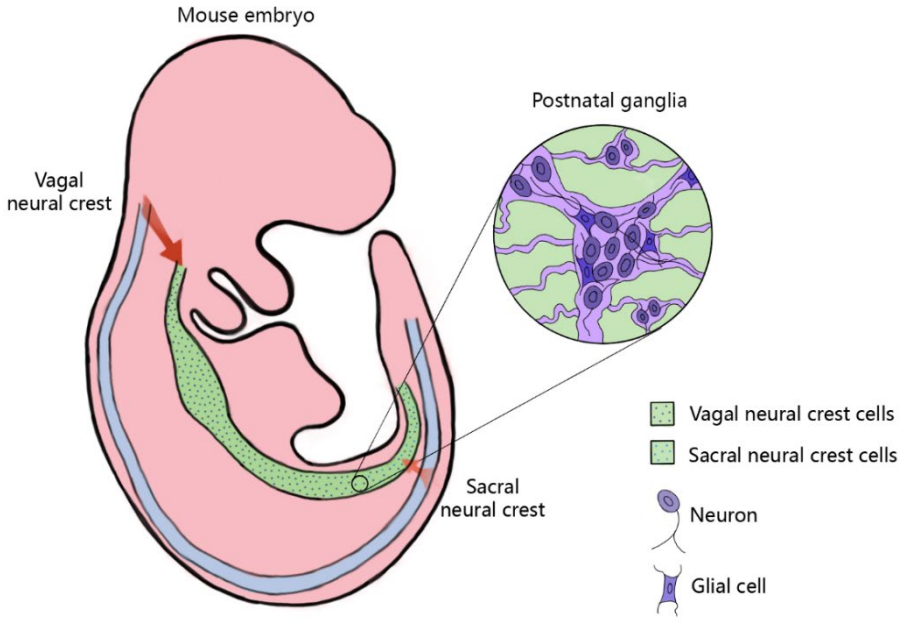


Figure 2. Schematic of vagal neural crest cell (NCC) migration into the embryonic gut tube and the contribution of sacral NCCs in the hindgut. The final ganglia contain glial and neuronal populations.

at the midgut to hindgut boundary to allow for trans-mesenteric migration to occur before the wave-front of ENCCs advances into the hindgut region[11]. It is unclear whether this trans-mesenteric migration occurs in other species, including humans, however given that Hirschsprung disease (HSCR) is mostly limited to the distal colon, this migratory path of ENCCs is interesting for further study.

As the gut is growing and elongating during embryogenesis and foetal development the ENCCs are highly migratory, with cells migrating in all directions rather than just towards the caudal end of the gut. The leading wavefront of migrating cells sets the tracks for the other cells to follow, determining the position of the ganglionic network[12,13]. Migratory pathways are dependent on the expression of surface receptors in order for the cells to recognise environmental signals, if present. For example, GDNF, expressed in the gut mesenchyme, is recognised by GFRα1, expressed on the migrating NCCs and acts as a chemoattractant for the migrating wavefront[7]. Studies in avian models have established that migration in the pre-umbilical stage, through the foregut and midgut, occurs before smooth muscle formation. As the cells migrate along this mesenchymal gut tube, the circular muscle layer begins to develop, creating a different migration environment for cells in the hindgut[10,13]. Additionally, cell adhesion molecules and appropriate extracellular matrix (ECM) components, such as collagen 18 and agrin proteins, secreted at the migrating wavefront, play a crucial role in these migratory pathways[12,14]. Intercellular and extracellular signals from surrounding cells and the ECM also influence cell fate decisions and components of the ECM have been studied *in vitro* to assess their effect on ENCCs and other gut cells. It has been shown that on tissue engineering matrices, nitrergic neuronal populations are

enriched with presence of collagen IV. Presence of laminin and/or heparan sulphate gives balanced relaxant and contractile motor neuron populations[15].

As cells stop migrating and find their final positions within the gut wall, they form connections and synapses with each other and the neural cell bodies group into ganglia[16,17]. The ENCCs differentiate into multiple neuronal subtypes as well as enteric glia. At least 20 enteric neuronal subtypes have been identified, varying in function, electrophysiology, neuro-transmitters and morphology[18].

Multiple trophic factors, morphogens, and transcriptional regulators control and influence enteric neural subtype specification within the ENS[19]. However, the mechanisms determining specific differentiation patterns are poorly understood and a fate map of ENCC differentiation does not yet exist. Factors influencing ENCC proliferation, migration or differentiation may alter the ratio of subtypes within the GI tract, depending on when and where a neuron is 'born' during ENS development, as evidenced by its exit from the cell cycle. The timing of cell cycle exit has been linked with lineage commitment and this could mean that a slowing of migration or proliferation of ENCCs would not similarly slow-down their differentiation and cell cycle exit[20]. It is possible that this could lead to insufficient naïve ENCCs reaching the distal colon. Neural activity also influences ENCC differentiation and migration, as treatment with neurotoxins that inhibit vesicular signalling has been shown to alter the ratio of neuronal subtypes and slow cell migration[21]. Thus, alterations in developmental signals can occur through many modes and pathways, which create opportunities for physiological and/or anatomical malformations of the ENS.

Hirschsprung Disease

Hirschsprung disease (HSCR) is a congenital neuropathy of the GI tract characterised by an absence of enteric ganglia in a variable length of the distal gut (Figure 3). This defect is usually confined to the colon but, although rare, total intestinal aganglionosis can occur. The length of the

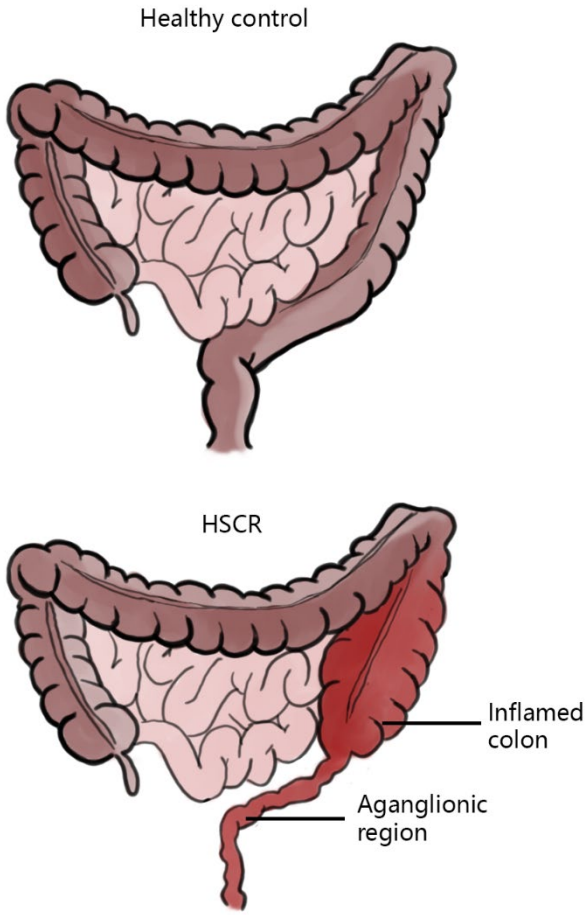


Figure 3. Schematic of the intestines of a healthy individual compared to a patient with HSCR. The tonic constriction of the distal colon in the patient leads to functional obstruction. The inability of faecal matter to pass causes inflammation and expansion of the proximal regions of the colon, observed by abdominal distension in the patient.

aganglionic segment varies from only the rectal regions, or rectal and sigmoid regions, (short segment), extended towards the splenic flexure or transverse colon (long segment), or further to the cecum (total colonic aganglionosis; TCA). The affected region of gut lacks intrinsic neural input with the result that the smooth muscles of the gut wall contract, causing life-threatening obstruction and preventing passage of stool. Complications from HSCR include infections, enterocolitis, abdominal swelling and potential rupture of the colon[6,22-24].

HSCR results from a failure of ENCCs to colonise the full length of the GI tract, which could be caused by functional deficit within the ENCCs themselves or in the local environment that the ENCCs encounter during their migration along and within the gut. The appropriate proliferation, migration, differentiation and survival of these cells is therefore essential for colonisation. The basic pathophysiological feature in HSCR is functional obstruction caused by the tonic contraction of smooth muscle of the aganglionic segment, and absence of the peristaltic motion of the gut. Despite extensive research, the pathophysiology of this is not fully understood. There is no clear explanation for the occurrence of tonic contraction of this smooth muscle other than the absence of signals from the ENS.

Genetics of HSCR

Genetics of Isolated HSCR

HSCR is a congenital disorder and can be present as an isolated feature or part of a syndrome. Isolated HSCR has been shown to have familial recurrence, but it most commonly occurs sporadically[2,25]. The incidence of HSCR is estimated at 1 in 5000 live births, although this varies between populations[6]. Developmentally, HSCR has a genetic

component and there is a sex-linked bias in classical and short segment cases with a male:female ratio of approximately 4:1[6]. To date, at least 17 genes have been found to play a role in isolated HSCR development, in patients and animal models (Table 1)[2,26,27]. Of these, the Rearranged during Transfection gene (*RET*) has proved to be the most important. This can be concluded from several genetic linkage analyses which have shown that, even in the absence of pathogenic coding variants in *RET*, over 90% of familial HSCR is linked to the gene[28,29]. Mutations affecting

Table 1. HSCR Associated genes

Gene	Location	Phenotype	Incidence	Pathway
RET	10q11.2	HSCR	50% familial; 20% sporadic	RET
GDNF	5p13	HSCR	Very rare	RET
GFRα1	10q26	HSCR	Very rare	RET
NTN	19p13	HSCR	Very rare	RET
PSPN	19p13	HSCR	Very rare	RET
EDNRB	13q22	HSCR; WS4	3-7%	Endothelin
EDN3	20q13	HSCR; WS4	<5%	Endothelin
SOX10	22q13	HSCR; WS4	>5%	Transcription factor
L1CAM	Xq28	HSCR x-linked hydrocephalus	Rare	-
NRG1	8p12	HSCR	6%	ERBB2; RET
NRG3	10q23.1	HSCR	-	ERBB2; RET
DENND3	8q24.3	HSCR (zebrafish)	-	-
NCLN	19p13.3	HSCR (zebrafish)	-	-
NUP98	11p15.4	HSCR (zebrafish)	-	-
TBATA	10q22.1	HSCR (zebrafish)	-	-
IHH	2q35	Hypoganglionosis (zebrafish)	-	Hedgehog; RET
GLI3	7p14.1	-	-	Hedgehog; RET

HSCR: Hirschsprung disease; **WS4**: Waardenburg-Shah syndrome.

its coding and non-coding regions have been described[30,31]. When screening for mutations, pathogenic coding variants in *RET* are identified in ~50% of familial and 15-35% of sporadic HSCR cases[25,32].

A number of loci in or around non-coding regions of *RET* have been identified, in several studies, to be linked with susceptibility to or protection from HSCR[33-41]. An enhancer variant in *RET* intron 1, a C>T SNP, was found to have a higher HSCR contribution risk than coding sequence variants. The frequency of the T allele in the European population is approximately 20% and is as high as 50% in the Chinese population, which could contribute to the higher incidence of HSCR in Asian populations[25,39]. Common variants in *RET* are well established as a susceptibility factor for HSCR[30,42]. All these studies point towards a central role for *RET* in the development of isolated HSCR and the ENS.

RET signalling balance

RET encodes for a receptor tyrosine kinase, RET, which is involved in several intracellular signalling cascades, regulating cell differentiation, migration, proliferation and survival. RET activation depends on binding with either glial cell line-derived neurotrophic factor (GDNF), neurturin (NRTN), artemin (ARTN) or persephin (PSPN) and with one of the four GDNF family receptor alphas (GFRα1-4) respectively to form a complex[43]. Disturbance of RET or any of these binding partners will in turn affect the action of downstream pathways, which can also disturb the development of other components of the peripheral and central nervous systems[44]. It has been hypothesized that, because *RET* proves to be the key player in isolated HSCR development, *RET* signalling is the fulcrum of ENS formation, with other pathways and the action of ENS-related proteins being dependent upon correct *RET* expression[2]. This

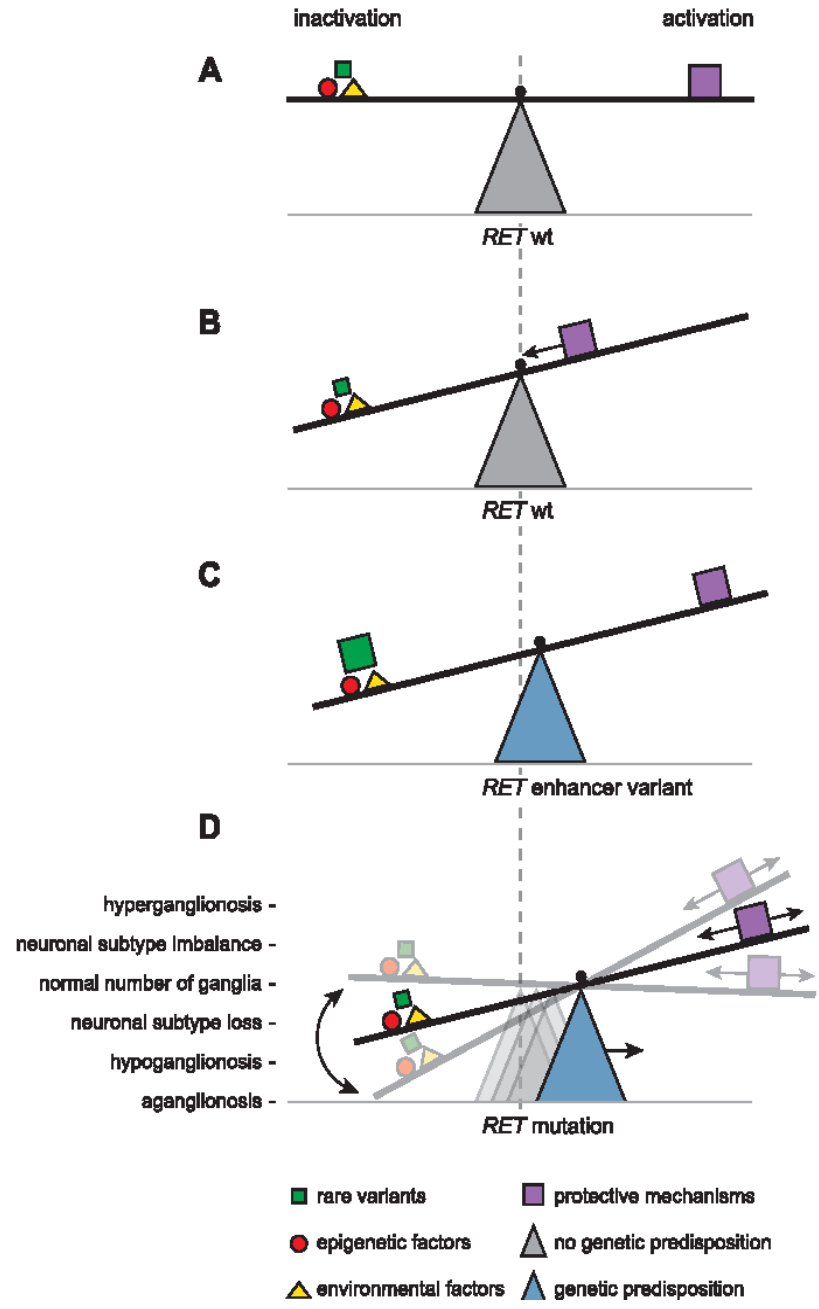


Figure 4. The *RET* signalling balance theory proposes *RET* as the fulcrum of ENS development. Other protective or predisposing factors can contribute in varying amounts towards the tipping of the normal balance towards hyperganglionosis or aganglionosis[2].

model proposes *RET* as the primary influencing factor in the multifactorial development of the ENS, the balance of which can be shifted by genetic or non-genetic factors to lead to a spectrum of ENS phenotypes, ranging from total aganglionosis to hyperganglionosis (Figure 4). It is likely to be a combination of (inherited) genetic and non-genetic factors that cause HSCR.

Genetics of Syndromic HSCR

HSCR occurs as an isolated phenotype with no associated anomalies in the majority of patients[45]. However, due to the genetic nature and the interaction of pathways in this disease it is inevitable that if ENS development is impaired there may be associated impairments in other cell types. Approximately 12% of HSCR patients have an associated chromosomal abnormality and 18% have additional congenital anomalies[2,25]. These associated anomalies most commonly affect other ectoderm or neural crest derivatives, but there are crossovers to other systems that may have links with cell migration or proliferation signals[46].

There are a number of defined syndromes that have HSCR as a feature and are generally explained by Mendelian inheritance (Table 2). In some of these it is a variable feature and its presence or absence may be influenced by modifying factors that either predispose a patient to HSCR or protect against its development[25,46]. Goldberg-Shprintzen syndrome (GOSHS) is caused by truncating variants in the KIF1 Binding Protein gene (*KIF1BP*)[47]. It is characterised by dysmorphic facial features, microcephaly, developmental delay, intellectual disability and short stature and has HSCR as a variable feature among other associated phenotypes. The presence or absence of HSCR in GOSHS patients is highly variable, even in members of the same family, with the same

Table 2. Characterised syndromes with HSCR as a mandatory or frequent feature.

Syndrome	Gene(s)	HSCR	Other features
WS4	*SOX10; EDNRB; EDN3*	100%	Pigmentary anomalies; sensorineural deafness
MWS	*ZEB2*		
GOSHS	*KIF1BP*	>70%	Craniofacial dysmorphia; microcephaly; polymicrogyria; developmental delay
DS	Tri21	~7%	Characteristic facial dysmorphism; intellectual disability; developmental delay
CCHS	*PHOX2B*	~20%	Autonomic respiratory failure
BBS	Several		Pigmentary anomalies; renal anomalies; intellectual disability; polydactyly
MKKS	*MKKS/BBS6*	~10%	Cardiac anomalies; polydactyly; hydrometrocolpos
SLOS	*DHCR7*	-	Developmental delay; intellectual disability; microcephaly; craniofacial dysmorphism; syndactyly
CHHS	*RMRP*	~10%	Metaphysial dysplasia; dwarfism; fine, sparse, blonde hair; anaemia; immunodeficiency
MEN2	*RET*	~2%	MTC; parathyroid tumours; pheochromocytoma

HSCR: Hirschsprung disease; **WS4**: Waardenburg-Shah syndrome; **CCHS**: Congenital Central Hypoventilation Syndrome; **MWS**: Mowat-Wilson Syndrome; **GOSHS**: Goldberg-Shprintzen Syndrome; **BBS**: Bardet-Biedl Syndrome; **MKKS**: McKusick-Kauffman Syndrome; **SLOS**: Smith-Lemli-Opitz Syndrome; **CHHS**: Cartilage-Hair Hypoplasia Syndrome; **MTC**: Medullary Thyroid Cancer.

causative variant[46,48,49]. *KIF1BP* is associated with microtubule dynamics, cargo trafficking and axonal outgrowth, but its precise functions in development are not well known[49-52].

In **chapter 3** we introduce new truncating variants in *KIF1BP* and add two patients with missense variants, one with classical GOSHS features and the other with an alternative phenotype. Functional studies of the missense variants indicate that a threshold of *KIF1BP* expression is necessary to avoid GOSHS development.

Copy Number Variation in HSCR

As well as defined monogenic syndromes, there are many patients with complex phenotypes and multiple associated developmental defects that are yet unexplained. Chromosomal abnormalities may explain part of these clinically complex patients. Changes in the number of copies of DNA present in the genome are termed Copy Number Variations (CNVs). These can contribute to phenotypes, diseases or syndromes that are influenced by gene dosage[53]. CNV is also known to contribute to HSCR disease aetiology. Chromosomal band deletions[54,55] and duplications[54,56-58] have been described in HSCR patients, most of these being syndromic patients. Deletions of chromosomes 10 and 13 were instrumental in the identification of *RET* and *EDNRB* as major HSCR genes[25,59]. Patients with Down Syndrome, trisomy 21, have a 100 times higher incidence of HSCR than the general population[25], implying that genes or regions on chromosome 21 are sensitive to dose increase and may increase susceptibility to ENS disorders. Large CNVs are more frequent in individuals with developmental anomalies compared to healthy controls[60]. In addition, more common CN polymorphisms (CNPs) are thought to be modifiers of the HSCR phenotype[61,62]. Therefore, we believe that rare CNVs could contribute significantly to syndromic patients with HSCR as a feature, where no pathogenic variant can be identified. In **chapter 4** this phenomenon is explored in HSCR patients with and without other congenital anomalies in order to find new genes or regions that may be causative for the HSCR phenotype.

Other factors in HSCR development

Known HSCR genes have only been implicated in ~30% of cases[25]. There has been no implicated high penetrant causative variant found in the remaining ~70% of cases and in sporadic, non-familial, HSCR genetic

counselling is challenging. There are a number of possible reasons for the missing heritability. There may be an epigenetic component affecting protein expression of the known, implicated pathways, or there may be other genes involved in pathways up or downstream of, or otherwise linked to, these pathways that have not yet been elucidated. Changes in expression of HSCR related genes, either due to variation in methylation regulating genes, or methylation changes to promotor or enhancer regions themselves, have been shown to influence ENS development[63-65]. Environmental factors including maternal diet and use of prescribed drugs can also play a role and may influence epigenetic patterns. This is most likely limited to being a small influencing factor rather than being significant enough alone to cause any damaging phenotype[66-69].

A further theory that warrants investigation is the existence of somatic cell variations affecting a subset of cells important in the ENS colonisation of the GI tract. During the accelerated cell division and growth of embryonic development there are many chances for mistakes to be made in DNA replication. If there is a failure of DNA repair mechanisms to identify and correct these mistakes this would lead to subsets of cells containing variants that are not present in the remaining embryonic cells. Dependent on the temporal and spatial identity of a cell in which this variation occurs this may lead to a whole system, organ or tissue containing a mosaic variant. However, as discussed in **chapter 2**, such defects would prove difficult to identify in patient material.

Diagnosis & Treatment

HSCR is suspected when a newborn infant fails to pass meconium within the first 48 hours of life, which is generally the case for >90% of HSCR

cases[22]. Patients are usually diagnosed before 3 months of age, although some within the first year or upon weaning from breastmilk. Rarely, an older patient will be diagnosed, but they generally have a history of chronic obstipation and the aganglionic segment is short[22]. When HSCR is suspected, the diagnosis is usually confirmed by taking a rectal suction biopsy. The length of the aganglionosis is established using histological examination for presence/absence of enteric ganglia.

The current standard of care for HSCR is surgical removal of the affected aganglionic region and anastomosis of the ganglionic bowel region to the anus. A colostomy may be fashioned prior to surgery. This allows the obstruction to be bypassed, and allows the gut to grow and inflammation to diminish. The most common surgical mode for resection is transanal endorectal pull-through (TEPT), which minimises the invasive nature of the surgery as the abdominal cavity does not need to be opened. TEPT can take place entirely transanally. However, it is commonly performed in combination with laparoscopy or laparotomy, to visualise and mobilise the colon abdominally[70]. Common surgical procedures are the Swenson, Soave, Duhamel and Rehbein procedures which are adaptations of similar pull-through approaches with differences in anastomosis[6,70,71].

Although surgical intervention is generally effective in preventing obstruction, it may not prevent further complications for the patient, including enterocolitis, faecal incontinence or chronic constipation, and an increased risk of infections[23,72,73]. TEPT may avoid some of the risks of open abdominal surgery, however the rates of faecal incontinence as well as the psychological and psychosexual side-effects of this surgery may negatively impact on the quality of life of HSCR patients[24,72,74,75]. It

has also been shown that TEPT can cause long-term damage to the anal sphincter, due to the prolonged and significant period of being stretched. This can be minimised if a combined laparotomy technique is implemented[6]. If complications arise following surgery, it may be necessary to have a redo surgery. A redo surgery is primarily for patients with persistent constipation and may involve removal of a further section of colonic tissue to ensure that the transition zone between the ganglionic and aganglionic regions is totally removed, as this is likely to have disturbed function and may have been inaccurately identified at the time of surgery[76]. Correction may also be necessary if there is a twisting of the bowel in the anastomosis which leads to discomfort or abdominal pain[77]. Patients who undergo corrective transanal surgery have a higher risk of damage to the anal sphincter, so this is only undertaken if the benefits outweigh this risk.

Additional, non-surgical, treatment options for HSCR include continued use of a colostomy, modified diet, laxatives and/or anti-diarrhoeal drugs, electrical nerve stimulation, hospital visits for rectal/colonic irrigation and injection of bulking agents to thicken the anal sphincter. These options, as well as the above-mentioned surgeries, leave a lot to be desired in terms of patient care. Quality of life can be negatively affected and, while paediatric patients are resilient to some psychological impacts, their frequent hospital visits and the social implications of faecal incontinence could complicate educational and social commitments in later life[24,72,75]. Due to normal life expectancy in HSCR patients the need for prolonged treatment can generate large healthcare costs[78-81]. New treatment avenues that are being explored in the field include nerve cell-replacement or transplantation therapies, discussed here and in **chapter 5**. Other, less well explored options can

be multidisciplinary and effective for a greater range of patients, as discussed in **chapter 6**.

Cell-based therapies for HSCR

One of the major drives in ENS translational research is to develop a cell replacement therapy for enteric neuropathies such as HSCR. HSCR is the primary focus for cell replacement therapy as cells could be transplanted into the aganglionic region in an attempt replenish and rescue the absent ENS. There are a number of cell types that are generally considered for any cell transplantation therapy: cells that are obtained from the same tissue as the cells that are to be replaced; ameliorative cells, such as mesenchymal stem cells (MSCs), that facilitate natural recovery within the body and decrease inflammation; or stem cells, either of embryonic, postnatal or reprogrammed origin, that are differentiated towards the desired cell lineage or type.

Sources of cells for ENS therapy

A number of options for a viable cell source for transplantation therapies have been explored[82-86]. The discovery of ENS stem cells that persist within the postnatal gut, and the assessment of their proliferative potential in mice[13,87], led to multiple attempts to isolate them from human colon and characterise them *in vitro*[88]. Although they provide an ideal cell source for functionally investigating the ENS, their proliferative capacity may be inadequate for generating required numbers of cells, especially when obtained from postnatal gut[89]. Current culture methods of ENCCs in neurospheres maintain a progenitor state in some cells, and these cells are able to integrate and form a functional network upon transplantation[85,90-92]. However, the area covered remains low in mouse

models and considering the scale-up to human gut, it is likely that many more cells would be required in a human therapy.

Pluripotent stem cells (PSCs) were initially thought to be ideal stem cell source for treating a variety of diseases. First studied from embryonic origin, embryonic stem cells (ESCs) offered a potentially unlimited source of cells to study differentiation pathways[93,94]. However, ethical considerations of using human embryos for this purpose, as well as the need for immunosuppression in transplantation, created significant barriers for their use[95]. The generation of induced PSCs (iPSCs) from human fibroblast cells by Takahashi and Yamanaka in 2006 led to a huge advance in developmental biology research. These cells have comparable differentiation potential to ESCs and can be created with somatic cells from the intended patient, circumventing source and immune rejection issues[96,97].

However, persisting issues include the genomic stability of the iPSCs, which have been shown to accumulate chromosomal aberrations after a number of passages in culture, and the ability to differentiate the cells into the correct lineage whilst avoiding the introduction of tumorigenic cells to a patient[98-100]. An additional consideration when thinking of a transplantation therapy for a genetic disease, is the genetic background of the patient and whether the cells' function will be negatively affected by the pathogenic variant that initially caused the disease.

A number of protocols have been developed for the differentiation of PSCs towards an ENS lineage[101-104]. Most of them initially achieve a vagal neural crest expression pattern and the study from Fattahi, et al., shows promise in transplantation and rescue of a mouse model of HSCR[103]. However, if iPSC-derived neural crest cells (NCCs) are to be used for

disease modelling and transplantation, a standardised method for creation and characterisation should be developed to ensure replicability. Particularly in the case of using HSCR patient derived iPSC-NCCs for genetic characterisation, standardised and controlled methods would ensure that any differences measured are due to genetic background rather than protocol variation. The parameters for measuring cell function, to assess the need for genetic correction, as well as the necessary stage of differentiation for ideal integration efficiency following transplantation, remain to be determined[105].

Aims & Outline of this Thesis

The aims of the work described in this thesis were to investigate the missing heritability that is seen in HSCR in both isolated and syndromic cases, understand more about the development and differentiation of ENCCs, assess suitability of various cell sources for transplantation therapy, and explore possible treatment avenues for current and future HSCR patients.

The possibilities of somatic variants contributing to the development of, and accounting for the missing heritability in, HSCR are discussed in chapters 2.1 and 2.2. In **chapter 2.1** we outline the need for appropriate distinction between inherited parental mosaicism and true somatic mosaicism, proposing an appropriate experimental design to truly differentiate between the two in HSCR patient tissue. In **chapter 2.2** this experimental design is utilised to look for ENCC specific variation in patient tissue. We further discuss the mechanisms of somatic variation in ENCCs and why, due to the developmental patterning of the ENS, they are unlikely to be detected, even if they have contributed to the phenotype.

In syndromic cases, HSCR can be a variable feature, the presence of common *RET* variants as well as the type of causative variant may influence HSCR development. The presence of HSCR in cases of GOSHS is discussed in **chapter 3**. Pathogenic variants in *KIF1BP* lead to GOSHS, in which HSCR is a variable feature. Given that HSCR is not a mandatory feature, it is likely that predisposing factors can be found to have involvement in its development. A number of patients have been reported in the literature, and truncating variants in *KIF1BP* have been found to be causative in all sequenced cases. In this chapter, we report nine new patients with *KIF1BP* variants, and functionally investigate, for the first time, three new missense variants identified in two patients with differing phenotypes. We also look at the presence of *RET* common SNPs, as a determining factor for the presence or absence of HSCR.

Another possible genetic factor for HSCR in syndromic cases is the presence of CNVs that affect dosage sensitive HSCR loci. In **chapter 4** we compare the size and number of CNVs between syndromic and non-syndromic HSCR cases, to find new candidate genes/loci.

To further look into the underlying pathogenesis of HSCR development we created iPSC lines from four patients with different pathogenic variations. These cells offer possibilities for disease modelling, functional investigation of variants as well opening the door to future iPSC-enteric neuron transplantation options. In **chapter 5** we present the characterisation of these patient-derived iPSCs and explore variations in their function compared to iPSCs generated from healthy controls. We also discuss their genomic stability and viability with a view to their use in cell therapy.

In order to bring an expedient option for the treatment of HSCR patients who continue to suffer from gastrointestinal problems following the current surgical standard of care, we may have to think beyond biological interventions. The technological world has arguably been able to advance at a faster rate than the development of purely biological treatment options. With the advances in microelectronics and prosthetic technologies incorporating sensory input, the opportunities for developing transplantable devices may provide a more elegant solution than the more primitive prostheses currently available. These ideas are discussed in **chapter 6,** together with a patent proposal for an artificial prosthetic sphincter with an anatomically relevant mechanism and design which could provide therapeutic options for HSCR patients, but also to others suffering from faecal incontinence, or loss of anal sphincter control.

Chapter 7 summarises the work in this thesis and discusses future prospects in the field of ENS development, as well as possibilities for treating patients with HSCR and related disorders.

References

1. Furness JB, Costa M. *The enteric nervous system.* Churchill Livingstone; 1987.
2. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental Biology.* 2016;417(2):198-208.
3. Struijs M-C, Diamond IR, de Silva N, Wales PW. Establishing norms for intestinal length in children. *Journal of pediatric surgery.* 2009;44(5):933-938.
4. Roberts DJ. Molecular mechanisms of development of the gastrointestinal tract. *Developmental dynamics: an official publication of the American Association of Anatomists.* 2000;219(2):109-120.
5. Furness JB. The enteric nervous system and neurogastroenterology. *Nature Reviews Gastroenterology &Amp; Hepatology.* 2012;9:286.
6. Holschneider AM, Puri P. *Hirschsprung's disease and allied disorders.* Springer Science & Business Media; 2007.
7. Anderson RB, Newgreen DF, Young HM. Neural Crest and the Development of the Enteric Nervous System. In: Saint-Jeannet J-P, ed. *Neural Crest Induction and Differentiation.* Boston, MA: Springer US; 2006:181-196.
8. Young HM, Hearn CJ, Newgreen DF. Embryology and development of the enteric nervous system. *Gut.* 2000;47(suppl 4):iv12-iv14.
9. Burns AJ, Champeval D, Le Douarin NM. Sacral Neural Crest Cells Colonise Aganglionic Hindgut in Vivo but Fail to Compensate for Lack of Enteric Ganglia. *Developmental Biology.* 2000;219(1):30-43.
10. Burns AJ, Le Douarin NM. The sacral neural crest contributes neurons and glia to the post-umbilical gut: spatiotemporal analysis of the development of the enteric nervous system. *Development.* 1998;125(21):4335-4347.
11. Nishiyama C, Uesaka T, Manabe T, et al. Trans-mesenteric neural crest cells are the principal source of the colonic enteric nervous system. *Nature Neuroscience.* 2012;15:1211.
12. Nagy N, Barad C, Hotta R, et al. Collagen 18 and agrin are secreted by enteric neural crest cells to remodel their microenvironment and regulate their migration during ENS development. *Development.* 2018.
13. Burns AJ. Migration of neural crest-derived enteric nervous system precursor cells to and within the gastrointestinal tract. *Int J Dev Biol.* 2003;49(2-3):143-150.
14. Bondurand N, Southard-Smith EM. Mouse models of Hirschsprung disease and other developmental disorders of the enteric nervous system: Old and new players. *Developmental Biology.* 2016.
15. Raghavan S, Bitar KN. The influence of extracellular matrix composition on the differentiation of neuronal subtypes in tissue engineered innervated intestinal smooth muscle sheets. *Biomaterials.* 2014;35(26):7429-7440.
16. Le Douarin NM, Teillet M-A. The migration of neural crest cells to the wall of the digestive tract in avian embryo. *J Embryol Exp Morphol.* 1973;30(1):31-48.
17. Rollo BN, Zhang D, Simkin JE, Menheniott TR, Newgreen DF. Why are enteric ganglia so small? Role of differential adhesion of enteric neurons and enteric neural crest cells. *F1000Research.* 2015;4.

18. Furness JB. Types of neurons in the enteric nervous system. *Journal of the Autonomic Nervous System.* 2000;81(1-3):87-96.
19. Avetisyan M, Schill EM, Heuckeroth RO. Building a second brain in the bowel. *The Journal of clinical investigation.* 2015;125(3):1-9.
20. Bergner AJ, Stamp LA, Gonsalvez DG, et al. Birthdating of myenteric neuron subtypes in the small intestine of the mouse. *J Comp Neurol.* 2014;522(3):514-527.
21. Hao MM, Moore RE, Roberts RR, et al. The role of neural activity in the migration and differentiation of enteric neuron precursors. *Neurogastroenterology & Motility.* 2010;22(5):e127-e137.
22. Dasgupta R, Langer JC. Hirschsprung disease. *Curr Probl Surg.* 2004;41(12):949-988.
23. Frykman PK, Short SS. Hirschsprung-associated enterocolitis: prevention and therapy. *Semin Pediatr Surg.* 2012;21(4):328-335.
24. Onishi S, Nakame K, Kaji T, et al. The bowel function and quality of life of Hirschsprung disease patients who have reached 18 years of age or older – the long-term outcomes after undergoing the transabdominal soave procedure. *Journal of Pediatric Surgery.* 2017;52(12):2001-2005.
25. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.
26. Gui H, Schriemer D, Cheng WW, et al. Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. *Genome Biol.* 2017;18(1):48.
27. Sribudiani Y, Chauhan RK, Alves MM, et al. Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease. *Gastroenterology.* 2018.
28. Bolk S, Pelet A, Hofstra RMW, et al. A human model for multigenic inheritance: Phenotypic expression in Hirschsprung disease requires both the RET gene and a new 9q31 locus. *Proceedings of the National Academy of Sciences.* 2000;97(1):268-273.
29. Gabriel SB, Salomon R, Pelet A, et al. Segregation at three loci explains familial and population risk in Hirschsprung disease. *Nature Genetics.* 2002;31:89.
30. Emison ES, Garcia-Barcelo M, Grice EA, et al. Differential Contributions of Rare and Common, Coding and Noncoding Ret Mutations to Multifactorial Hirschsprung Disease Liability. *The American Journal of Human Genetics.* 2010;87(1):60-74.
31. Hofstra RM, Wu Y, Stulp RP, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. *Human mutation.* 2000;15(5):418-429.
32. Attié T, Pelet A, Edery P, et al. Diversity of RET proto-oncogene mutations in familial and sporadic Hirschsprung disease. *Human molecular genetics.* 1995;4(8):1381-1386.
33. Burzynski GM, Nolte IM, Osinga J, et al. Localizing a putative mutation as the major contributor to the development of sporadic Hirschsprung disease to the RET genomic sequence between the promoter region and exon 2. *European journal of human genetics.* 2004;12(8):604.

34. Fitze G, Appelt H, König IR, et al. Functional haplotypes of the RET proto-oncogene promoter are associated with Hirschsprung disease (HSCR). *Human molecular genetics.* 2003;12(24):3207-3214.
35. Pelet A, De Pontual L, Clement-Ziza M, et al. Homozygosity for a frequent and weakly penetrant predisposing allele at the RET locus in sporadic Hirschsprung disease. *Journal of medical genetics.* 2005;42(3):e18-e18.
36. Sancandi M, Griseri P, Pesce B, et al. Single nucleotide polymorphic alleles in the 5′ region of the RET proto-oncogene define a risk haplotype in Hirschsprung's disease. *Journal of Medical Genetics.* 2003;40(9):714-718.
37. Borrego S, Wright FA, Fernández RM, et al. A founding locus within the RET proto-oncogene may account for a large proportion of apparently sporadic Hirschsprung disease and a subset of cases of sporadic medullary thyroid carcinoma. *The American Journal of Human Genetics.* 2003;72(1):88-100.
38. Griseri P, Lantieri F, Puppo F, et al. A common variant located in the 3′ UTR of the RET gene is associated with protection from Hirschsprung disease. *Hum Mutat.* 2007;28(2):168-176.
39. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature.* 2005;434:857.
40. Sribudiani Y, Metzger M, Osinga J, et al. Variants in RET Associated With Hirschsprung's Disease Affect Binding of Transcription Factors and Gene Expression. *Gastroenterology.* 2011;140(2):572-582. e572.
41. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer variants synergistically drive dysfunction of a gene regulatory network in Hirschsprung disease. *Cell.* 2016;167(2):355-368.e310.
42. Burzynski GM, Nolte IM, Bronda A, et al. Identifying Candidate Hirschsprung Disease–Associated RET Variants. *The American Journal of Human Genetics.* 2005;76(5):850-858.
43. Takahashi M. The GDNF/RET signaling pathway and human diseases. *Cytokine Growth Factor Rev.* 2001;12(4):361-373.
44. Enomoto H, Crawford PA, Gorodinsky A, Heuckeroth RO, Johnson EM, Milbrandt J. RET signaling is essential for migration, axonal growth and axon guidance of developing sympathetic neurons. *Development.* 2001;128(20):3963-3974.
45. Alves MM, Sribudiani Y, Brouwer RWW, et al. Contribution of rare and common variants determine complex diseases—Hirschsprung disease as a model. *Developmental Biology.* 2013;382(1):320-329.
46. Brooks A. *Genetics of syndromic and non-syndromic Hirschsprung disease*. Enschede: Clinical Genetics, Erasmus University Rotterdam; 2005.
47. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous Nonsense Mutations in *KIAA1279* Are Associated with Malformations of the Central and Enteric Nervous Systems. *The American Journal of Human Genetics.* 2005;77(1):120-126.
48. Brooks AS, Breuning MH, Osinga J, et al. A consanguineous family with Hirschsprung disease, microcephaly, and mental retardation (Goldberg-Shprintzen syndrome). *Journal of Medical Genetics.* 1999;36(6):485-489.
49. Drévillon L, Megarbane A, Demeer B, et al. KBP–cytoskeleton interactions underlie developmental anomalies in Goldberg–Shprintzen syndrome. *Human molecular genetics.* 2013;22(12):2387-2399.

50. Hirst CS, Stamp LA, Bergner AJ, et al. Kif1bp loss in mice leads to defects in the peripheral and central nervous system and perinatal death. *Scientific Reports.* 2017;7(1):16676.
51. Kevenaar Josta T, Bianchi S, van Spronsen M, et al. Kinesin-Binding Protein Controls Microtubule Dynamics and Cargo Trafficking by Regulating Kinesin Motor Activity. *Current Biology.* 2016;26(7):849-861.
52. Wozniak MJ, Melzer M, Dorner C, Haring H-U, Lammers R. The novel protein KBP regulates mitochondria localization by interaction with a kinesin-like protein. *BMC Cell Biology.* 2005;6(1):1-15.
53. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med.* 2010;61:437-455.
54. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. *J Med Genet.* 2001;38(11):729-739.
55. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal deletion map of human malformations. *Am J Hum Genet.* 1998;63(4):1153-1159.
56. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality--and tolerance of segmental aneuploidy--in humans. *Am J Hum Genet.* 1999;64(6):1702-1708.
57. Mahboubi S, Templeton JM, Jr. Association of Hirschsprung's disease and imperforate anus in a patient with "cat-eye" syndrome. A report of one case and review of the literature. *Pediatr Radiol.* 1984;14(6):441-442.
58. Kerstjens-Frederikse WS, Hofstra RM, van Essen AJ, Meijers JH, Buys CH. A Hirschsprung disease locus at 22q11? *J Med Genet.* 1999;36(3):221-224.
59. Brooks A, Oostra B, Hofstra R. Studying the genetics of Hirschsprung's disease: unraveling an oligogenic disorder. *Clinical genetics.* 2005;67(1):6-14.
60. Cooper GM, Coe BP, Girirajan S, et al. A copy number variation morbidity map of developmental delay. *Nat Genet.* 2011;43(9):838-846.
61. Jiang Q, Ho YY, Hao L, Nichols Berrios C, Chakravarti A. Copy number variants in candidate genes are genetic modifiers of Hirschsprung disease. *PLoS One.* 2011;6(6):e21219.
62. Tang CS, Cheng G, So MT, et al. Genome-wide copy number analysis uncovers a new HSCR gene: NRG3. *PLoS Genet.* 2012;8(5):e1002687.
63. Torroglosa A, Alves MM, Fernández RM, Antiñolo G, Hofstra RM, Borrego S. Epigenetics in ENS development and Hirschsprung disease. *Developmental Biology.* 2016;417(2):209-216.
64. Torroglosa A, Enguix-Riego MV, Fernández RM, et al. Involvement of DNMT3B in the pathogenesis of Hirschsprung disease and its possible role as a regulator of neurogenesis in the human enteric nervous system. *Genet Med.* 2014;16(9):703.
65. Tang W, Li B, Tang J, et al. Methylation analysis of EDNRB in human colon tissues of Hirschsprung's disease. *Pediatric Surgery International.* 2013;29(7):683-688.
66. Heuckeroth RO, Schäfer K-H. Gene-environment interactions and the enteric nervous system: Neural plasticity and Hirschsprung disease prevention. *Developmental biology.* 2016.

67. Lake JI, Tusheva OA, Graham BL, Heuckeroth RO. Hirschsprung-like disease is exacerbated by reduced de novo GMP synthesis. *The Journal of Clinical Investigation.* 2013;123(11):4875-4887.
68. Fu M, Sato Y, Lyons-Warren A, et al. Vitamin A facilitates enteric nervous system precursor migration by reducing Pten accumulation. *Development.* 2010;137(4):631-640.
69. Schill EM, Lake JI, Tusheva OA, et al. Ibuprofen slows migration and inhibits bowel colonization by enteric nervous system precursors in zebrafish, chick and mouse. *Developmental Biology.* 2016;409(2):473-488.
70. Tannuri ACA, Tannuri U, Romão RLP. Transanal endorectal pull-through in children with Hirschsprung's disease—technical refinements and comparison of results with the Duhamel procedure. *Journal of Pediatric Surgery.* 2009;44(4):767-772.
71. De la Torre L, Ortega A. Transanal versus open endorectal pull-through for Hirschsprung's disease. *Journal of Pediatric Surgery.* 2000;35(11):1630-1632.
72. Neuvonen MI, Kyrklund K, Rintala RJ, Pakarinen MP. Bowel Function and Quality of Life After Transanal Endorectal Pull-through for Hirschsprung Disease. *Ann Surg.* 2017;265(3):622-629.
73. Westfal ML, Goldstein AM. Pediatric enteric neuropathies: diagnosis and current management. *Curr Opin Pediatr.* 2017;29(3):347-353.
74. van den Hondel D, Sloots CEJ, Bolt JM, Wijnen RMH, de Blaauw I, Ijsselstijn H. Psychosexual Well-Being after Childhood Surgery for Anorectal Malformation or Hirschsprung's Disease. *The Journal of Sexual Medicine.* 2015;12(7):1616-1625.
75. Neuvonen M, Kyrklund K, Taskinen S, Koivusalo A, Rintala RJ, Pakarinen MP. Lower urinary tract symptoms and sexual functions after endorectal pull-through for Hirschsprung disease: controlled long-term outcomes. *Journal of Pediatric Surgery.* 2017;52(8):1296-1301.
76. Maia DM. The reliability of frozen-section diagnosis in the pathologic evaluation of Hirschsprung's disease. *The American journal of surgical pathology.* 2000;24(12):1675-1677.
77. Pini-Prato A, Mattioli G, Giunta C, et al. Redo surgery in Hirschsprung disease: what did we learn? Unicentric experience on 70 patients. *Journal of pediatric surgery.* 2010;45(4):747-754.
78. Huang E, Tolley E, Blakely M, Langham M. Changes in hospital utilization and management of Hirschsprung disease: analysis using the kids' inpatient database. *Ann Surg.* 2013;257(2):371-375.
79. Bradnock TJ, Knight M, Kenny S, et al. Hirschsprung's disease in the UK and Ireland: incidence and anomalies. *Archives of Disease in Childhood.* 2017.
80. Arth AC, Tinker SC, Simeone RM, Ailes EC, Cragan JD, Grosse SD. Inpatient hospitalization costs associated with birth defects among persons of all ages—United States, 2013. *MMWR Morbidity and mortality weekly report.* 2017;66(2):41.
81. Hackam D, Reblock K, Redlinger R, Barksdale E. Diagnosis and outcome of Hirschsprung's disease: does age really matter? *Pediatric surgery international.* 2004;20(5):319-322.
82. Hotta R, Natarajan D, Thapar N. Potential of cell therapy to treat pediatric motility disorders. Paper presented at: Seminars in pediatric surgery2009.

83. Sasselli V, Micci M-A, Kahrig KM, Pasricha PJ. Evaluation of ES-derived neural progenitors as a potential source for cell replacement therapy in the gut. *BMC gastroenterology.* 2012;12(1):1.
84. Wagner JP, Sullins VF, Dunn JCY. Skin-derived precursors generate enteric-type neurons in aganglionic jejunum. *Journal of pediatric surgery.* 2014;49(12):1809-1814.
85. Cooper JE, McCann CJ, Natarajan D, et al. In Vivo Transplantation of Enteric Neural Crest Cells into Mouse Gut; Engraftment, Functional Integration and Long-Term Safety. *PLoS ONE.* 2016;11(1):e0147989.
86. Burns AJ, Goldstein AM, Newgreen DF, et al. White paper on guidelines concerning enteric nervous system stem cell therapy for enteric neuropathies. *Developmental biology.* 2016.
87. Burns AJ, Pasricha PJ, Young HM. Enteric neural crest-derived cells and neural stem cells: biology and therapeutic potential. *Neurogastroenterology & Motility.* 2004;16(s1):3-7.
88. Natarajan D, Grigoriou M, Marcos-Gutierrez CV, Atkins C, Pachnis V. Multipotential progenitors of the mammalian enteric nervous system capable of colonising aganglionic bowel in organ culture. *Development.* 1999;126(1):157-168.
89. Zhang D, Rollo BN, Nagy N, Stamp L, Newgreen DF. The enteric neural crest progressively loses capacity to form enteric nervous system. *Developmental Biology.* 2018.
90. Natarajan D, Cooper J, Choudhury S, et al. Lentiviral labeling of mouse and human enteric nervous system stem cells for regenerative medicine studies. *Neurogastroenterology & Motility.* 2014;26(10):1513-1518.
91. Binder E, Natarajan D, Cooper J, et al. Enteric Neurospheres Are Not Specific to Neural Crest Cultures: Implications for Neural Stem Cell Therapies. *PLoS ONE.* 2015;10(3):e0119467.
92. Nishikawa R, Hotta R, Shimojima N, et al. Migration and differentiation of transplanted enteric neural crest-derived cells in murine model of Hirschsprung 's disease. *Cytotechnology.* 2014:1-10.
93. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *science.* 1998;282(5391):1145-1147.
94. Itskovitz-Eldor J, Schuldiner M, Karsenti D, et al. Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Mol Med.* 2000;6(2):88.
95. Wert Gd, Mummery C. Human embryonic stem cells: research, ethics and policy. *Hum Reprod.* 2003;18(4):672-682.
96. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *cell.* 2006;126(4):663-676.
97. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *cell.* 2007;131(5):861-872.
98. Taapken SM, Nisler BS, Newton MA, et al. Karyotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat Biotechnol.* 2011;29(4):313.
99. Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nature Reviews Cancer.* 2011;11(4):268.

100. Pera MF. Stem cells: The dark side of induced pluripotency. *Nature.* 2011;471(7336):46.
101. Lai FP, Lau ST, Wong JK, et al. Correction of Hirschsprung-Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function. *Gastroenterology.* 2017;153(1):139-153 e138.
102. Hackland JOS, Frith TJR, Thompson O, et al. Top-Down Inhibition of BMP Signaling Enables Robust Induction of hPSCs Into Neural Crest in Fully Defined, Xeno-free Conditions. *Stem Cell Reports.* 2017;9(4):1043-1052.
103. Fattahi F, Steinbeck JA, Kriks S, et al. Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. *Nature.* 2016;531(7592):105-109.
104. Fernandes TG, Duarte ST, Ghazvini M, et al. Neural commitment of human pluripotent stem cells under defined conditions recapitulates neural development and generates patient-specific neural cells. *Biotechnology Journal.* 2015;10(10):1578-1588.
105. Zhu Q, Lu Q, Gao R, Cao T. Prospect of Human Pluripotent Stem Cell-derived Neural Crest Stem Cells in Clinical Application.

2

Chapter 2.1

Do *RET* somatic mutations play a role in Hirschsprung disease?

Erwin Brosens*, **Katherine C. MacKenzie***, Maria M. Alves Robert M.W. Hofstra

Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands

* These authors contributed equally to this letter.

Refers to:
Jiang Q, Liu F, Miao C, Li Q, Zhang Z, Xiao P, Su L, Yu K, Chen X, Zhang F, Chakravarti A, Li L. "RET somatic mutations are underrecognized in Hirschsprung disease" Genet Med. 2017. doi:10.1038/gim.2017.178.

Genetics in Medicine, 2018

Dear Editor,

We have read the manuscript from Jiang et al.,[1] recently published in your journal, with great interest. In this, a contribution of *RET* somatic mutations for Hirschsprung disease (HSCR) is hypothesized.

HSCR is a complex inherited disorder characterized by the absence of enteric ganglia in the distal part of the colon. Several genes and loci have been described to underlie disease pathogenesis. However, variants in these genes explain no more than 20% of all cases[2]. This missing heritability seen in HSCR is a common feature of many complex disorders and explaining it remains challenging. Considering that HSCR develops during embryogenesis as a result of either impaired migration, proliferation or differentiation of enteric neural crest cells (ENCCs), it is tempting to consider that somatic mutations occurring during the development of the enteric nervous system, and specifically affecting ENCCs, can also contribute to HSCR genetics. Jiang and colleagues think that this is possible, and we share the same opinion, as we believe that somatic variants could be underrecognized in HSCR, and thus, possibly account for some of the missing heritability. However, the results described by the authors do not, in our opinion, fully support the conclusions of the manuscript. This is mainly because routine genetic testing on DNA derived from blood or saliva would not find these ENCC specific mutations, nor would easily detect low mosaic variants. With this letter, we intend to further discuss our concerns and highlight the difference between causative

somatic mutations and germline mosaicism resulting in seemingly *de novo* mutations in the next of kin.

In their study, the authors screened 152 HSCR patients by targeted exome sequencing and direct gene screening. In eight patients they identified putative deleterious *de novo* variants in *RET*. Since *RET* is the major HSCR causing gene,[3] they assumed that these variants were responsible for the disease phenotype. However, only six patients carried a truly *de novo RET* mutation, as in the other two the variants identified were also found in one of the parents. The authors continued the genetic analysis of the six remaining patients, and described based on their findings, that the *RET* mutations identified were somatic. This result led to the conclusion that *RET* somatic mosaicism is present in 75% of the HSCR cases and is, therefore, underrecognized. However, in four of the six patients studied, germline mosaicism was identified in the parents. As these variants are transmitted to the affected next of kin as heterozygous variants, these cannot represent somatic mosaicism. In addition, the parents are not affected, leading us to speculate that their ENCCs are either not affected by the mutation or that the mutational load in their ENCCs does not cross a threshold for abnormal ENS development. Therefore, although these four families are likely to represent germline mosaicism resulting in seemingly *de novo* mutations in the next of kin, they do not support the conclusion that *RET* somatic mutations are underrecognized. These heterozygous mutations would be detected (and are detected in these patients) in routine genetic screening and as such, do not resolve missing heritability due to "recognizing" somatic mosaicism.

For the remaining two patients, the RET variants identified were present in blood, saliva, and colon, in family 1 (39/39/44%) and family 2 (44/35/39%). The authors validated the sequencing data with TA cloning, but the results are variable even within the same tissue, and the differences seen in the amplicon-based sequencing test are well within the normal range for detection of a heterozygous variant. Moreover, the samples tested—blood, saliva, and colon—derive from tissues originating from different germ layers. Blood, mostly lymphocytes, is derived from mesoderm; saliva, lymphocytes, and epithelial lining of the mouth are derived from mesoderm and ectoderm; and the colon has contributions from all three germ layers. The high allele frequencies identified in all three samples for both patients (>30%) does not allow for discrimination between very early developmental stage somatic mutations or de novo variants present as a germ-line mosaicism in one of the parents. Although the authors acknowledge this fact in the discussion, it is more fitting to conclude that the variants identified in these two remaining patients are likely to be de novo heterozygous variants.

In conclusion, we agree with the hypothesis of Jiang et al. and think that somatic variants might well play a role in HSCR development. However, based on the results presented we think that it is not possible to make such a conclusion, as no true somatic RET mutations were identified in any of the presented patients. We believe that to draw the conclusions stated, the allele frequency of the mutations in ENCCs, surface ectoderm, and, for instance, blood or fibroblasts would need to be compared because these represent different germ layers and include the cells involved in HSCR. An experimental setup that would isolate cells specifically from each germ layer, as well as ENCCs,

or that would separate the colon into various cell types, would be necessary to determine whether true somatic mutations impact HSCR development. Only if differences were found under such conditions could the authors show that RET somatic mutations are indeed underrecognized in HSCR.

References

1. Jiang Q, Liu F, Miao C, et al. RET somatic mutations are underrecognized in Hirschsprung disease. *Genet. Med.* 2018;20(7):770-777.
2. Alves MM, Sribudiani Y, Brouwer RWW, et al. Contribution of rare and common variants determine complex diseases—Hirschsprung disease as a model. *Developmental Biology.* 2013;382(1):320-329.
3. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.


2

Chapter 2.2

The somatic mutation paradigm in Hirschsprung Disease

Katherine C. MacKenzie[1*], Erwin Brosens[1*], Rhiana Garritsen[1,2*], Rajendra K. Chauhan[1], Yunia Sribudiani[1,3], Bianca M. de Graaf[1], Tim Rugenbrink[1], Rutger Brouwer[4], Wilfred van Ijcken[4], Ivo de Blaauw[2,5], Alice S. Brooks[1], Cornelius E.J. Sloots[2], Conny J.H.M. Meeuwsen[2], René M. Wijnen[2], Donald F. Newgreen[6], Alan J. Burns[1,7,8], Maria M. Alves[1], Robert M.W. Hofstra[1,6].

[1] Department of Clinical Genetics, Erasmus University Medical Centre, Rotterdam, Netherlands
[2] Department of Paediatric Surgery, Sophia Children's Hospital, Erasmus University Medical Centre, Rotterdam, Netherlands
[3] Department of Medical Biochemistry and Molecular Biology, Faculty of Medicine, Universitas of Padjadjaran, Bandung, Indonesia
[4] Centre for Biomics, Erasmus University Medical Centre, Rotterdam, Netherlands
[5] Department of Paediatric Surgery, Amalia Children's Hospital, Radboud University Medical Center, Nijmegen, Netherlands
[6] Department of Cell Biology, Murdoch Children's Research Institute, Royal Children's Hospital, Parkville, Australia
[7] Department of Stem Cells and Regenerative Medicine, University College London, London, UK
[8] Takeda Pharmaceuticals, Cambridge, Massachusetts, USA

*these authors contributed equally to this project

Corresponding author: Prof. Dr. Robert M.W. Hofstra
Department of Clinical Genetics, Erasmus University Medical Centre, PO BOX 2040, 3000CA Rotterdam, The Netherlands.
Tel: +31-10-7030683. E-mail: r.hofstra@erasmusmc.nl

Submitted

Abstract

Not all patients with a suspected genetic condition receive a diagnosis after routine screening in clinical practice. One of the reasons could be that the causal mutation is not present in the cell types usually tested - whole blood, dermal fibroblasts or saliva - but in the affected tissue only. The last, somatic mutations, can occur in a given cell at any stage of development after conception and will be present in all subsequent daughter cells. If somatic mutations were to play a role in such cases then it would be unlikely to find the causative mutation in unaffected tissue, but in the daughter cells of the originally mutated cell, likely in the affected tissue. Using Hirschsprung disease (HSCR) as a model of complex genetics in a disease with an unequivocal phenotype we investigated multiple tissues from five patients in order to find somatic differences and investigate possible modes of disease development. The patients showed many somatic variations between the tissues, however causative mutations for HSCR were not specifically identified in the enteric neural crest cells of these patients. Larger copy number variations were also found, however these were not specific to enteric neural crest cells. Here we present these results, postulate various modes of development following somatic mutation, describe the challenges in detecting somatic mutations and hypothesize how this may contribute to 'missing heritability' in developmental defects.

Introduction

Congenital disorders can be caused by germline variants that, depending on the mode of inheritance, can be present in one or both parents of an affected individual. In the case of a dominant congenital disorder the evolutionary pressure against disease-causing genetic variations is high, and for that reason mutations often occur *de novo* in a patient. These changes appear new in the patient and originate from a *de novo* mutation in the germ cells of one of the parents. An alternative route for the development of what seems to be an inherited dominant disease is the occurrence of a somatic mutation(s) affecting a specific tissue(s). This process, somatic mosaicism, is a well characterized phenomenon known to contribute to a number of diseases, most notably skin disorders such as McCune-Albright syndrome and Darier-White disease [7,8]. Recently increasing attention has been focussed on somatic mosaicism in a range of neurological functional disorders particularly in children[9,10]. Somatic mutations can be evenly distributed throughout an organism, be segmental or tissue-specific and can affect somatic tissues, the germline or both. They can arise at any stage of development or adult life, and they may accumulate with age[11] and can be caused by alterations of a normal to a mutant genotype and vice-versa[12]. However, for a somatic mutation to play a significant role in tissue development and contribute to a congenital disorder, it is important that a threshold proportion of cells are affected. This situation most likely eventuates when the initiating mutation occurs at an extremely early stage of development[13], or when there is some form of clonal dominance such as when the mutation brings selective advantage/survival to the mutated cells[14,15]. Cancers can be seen as examples of the latter[16].

In this study we investigated the involvement of somatic mutations in a congenital neuropathy of the gastrointestinal (GI) tract called Hirschsprung disease (HSCR). HSCR is a developmental disease characterized by an inability of enteric neural crest cells (ENCCs) to colonize a variable length segment of the distal gut, which therefore lacks neurons and glial cells of the enteric nervous system (ENS)[17]. It is often inherited and at least 14 genes have been identified that play a role in HSCR development[18]. *RET*, which codes for an ENCC growth factor receptor, is the major gene for HSCR and mutations affecting its coding and non-coding regions have been described[19,20]. However, these 14 genes only explain 30% of all HSCR cases[21]. This missing heritability seen in HSCR is a common feature of complex disorders and can partly be explained by low penetrant non-coding variants[22-24], as well as by combinations of both rare coding and modifying variants[25,26]. The occurrence of somatic mosaicism could also account for, or at least contribute to, some of the cases. However, in this case pathogenic variants should preferably occur early in development to have an effect on the ENS. If they occur later during ENS development, the variants should result in a beneficial growth or survival advantage over the non-affected cells, in order to colonise a substantial portion of the colon and result in a phenotype.

The theory that somatic mutations may play a role in HSCR is not new, it has already been postulated in two and tested in three independent studies. The first study[27] investigated the presence of *RET* intronic variants [SNP1 (rs2506004) and SNP2 (rs2435357)] in a series of tissues (aganglionic, transition zone and ganglionic gut in distal to proximal order) collected from several HSCR patients. The authors observed that the aganglionic segments of HSCR patients tend to be

homozygous (hemizygous) for the disease-associated variants, particularly in patients with long segment HSCR, whereas the ganglionic intestinal tissues of the same patient were found to be heterozygous. It was postulated that this was caused by a deletion of the wildtype allele [27]. A similar study was performed by a different group, but they were not able to find allele frequency differences for three RET intronic polymorphisms (rs2506030, rs2506004 and rs2435357) in DNA isolated from blood and colon of the same patient[28]. The third study described the existence of low frequency *RET* somatic variants in different tissues from the same patient. In two patients, mutations were found in several tissues (blood, saliva and colon) at low frequencies and were absent in blood-derived DNA from the parents[29]. Although these studies claimed the involvement of *RET* somatic mutations in HSCR development, there are a number of caveats concerning the validity of the results. The first study stated that in the aganglionic segment of the HSCR patients 100% of the alleles were found to be mutated[27]. This implies that all examined cells had lost the wildtype allele. The enteric ganglia are derived from ENCCs, which only constitute a minority of the cells of the gut. Considering that no selection method was performed to enrich the neuronal population, it is likely that the authors not only analysed enteric neural crest (ENC)-derived cells but also mucosal cells, connective tissue cells and smooth muscle cells. Since gut mucosa, smooth muscle and ENS cells derive from different germ layers, i.e. endoderm, mesoderm and ectoderm respectively (Figure 1), the mutation identified at the level detected must have occurred extremely early in development and should therefore be present in all or most cells in all tissues of the body. Therefore, the identification of

RET intronic mutations only in the aganglionic region is difficult to explain. The third study used a different approach to study somatic mosaicism, but the questions raised are comparable, since *RET* mutations were identified in tissues derived from all three germ layers, making it difficult to discriminate between very early developmental stage somatic mutations, and *de novo* variants present as a germline mosaicism in one of the parents[30]. In light of these results, we believe that the involvement of somatic mutations in HSCR is still unclear. Here we further investigate its existence by comparing exome sequencing derived genotypes and single nucleotide polymorphism (SNP)-array derived Copy Number profiles of DNA isolated from ENC-derived cells with either blood or fibroblast samples collected from HSCR patients.

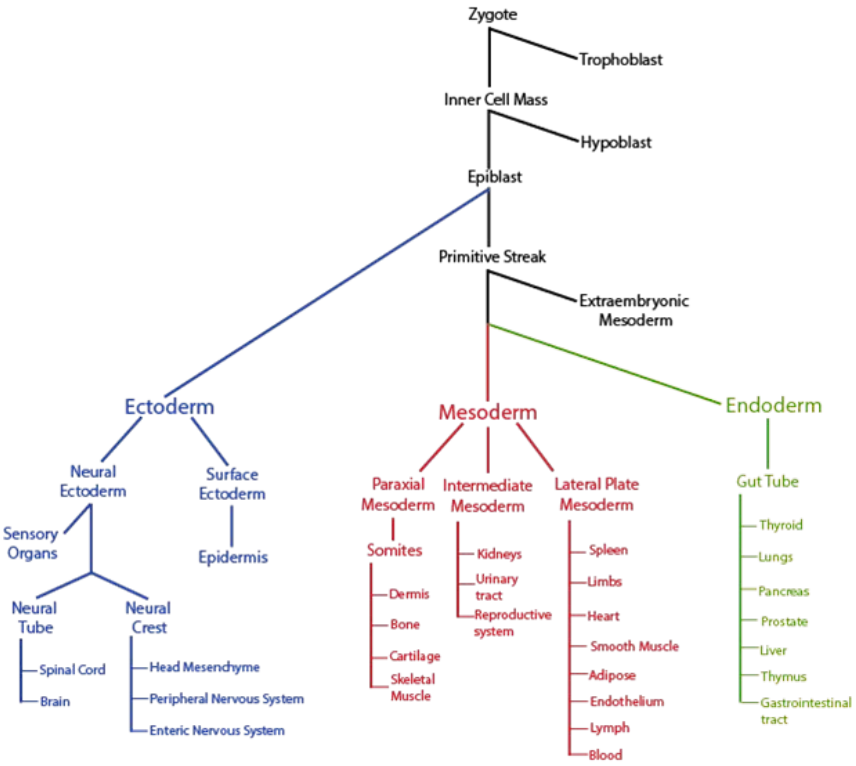


Figure 1. Developmental lineage tree highlighting where each of the three germ layers branch from and what structures and organs they form or contribute to. Blood samples would be mesodermal, saliva samples would contain leukocytes of mesodermal origin and epithelial cells of ectodermal origin from the surface ectoderm [5] and unsorted gut samples would contain derivatives of all three germ layers with ectodermal neural crest making up the minority of cells. Adapted from LifeMap [6].

Materials & methods

Patients & METC

Five HSCR patients undergoing routine colonic pull-through surgery in the Paediatric Surgery department of the Erasmus Medical Centre, Sophia Children's Hospital, Rotterdam and the Radboud Medical Centre, Amalia Children's Hospital, Nijmegen, were included in this study. All patients were operated in their first year of life (ranging from 88 to 194 days after birth). An overview of the clinical characteristics of the patients is given in Supplementary Table 1. Written informed consent for inclusion was obtained from the parents of the patients, and this project was approved by the medical ethics committee of the Erasmus Medical Centre (METC-2012-582).

Sample collection

Full-thickness colon biopsies from patients were obtained from the most proximal region of removed colon from the pull-through surgeries, and confirmed as ganglionic. Colon biopsies were washed with sterile PBS, and excess connective tissue as well as most of the mucosal layer, was removed. The remaining tissue was dissected and dissociated in 200U/ml Collagenase IV (Gibco, Thermo Fisher Scientific, USA) for 1 hour at 37°C. EDTA blood and a small skin biopsy (2mm) from the incision site in the abdomen, were also taken during surgeries. Fibroblasts were used as source of DNA when insufficient DNA was isolated from blood.

Cell culture & fluorescence activated cell sorting

Skin biopsies were dissected and plated in Ham's F-10 nutrient mix (Gibco, Thermo Fisher Scientific, USA) supplemented with 15% foetal calf serum (FCS) and 1% penicillin/streptomycin. Medium was

refreshed every 2-3 days and once confluent, cells were split at a ratio of 1:3, using TrypLE Express (Gibco, Thermo Fisher Scientific, USA) according to manufacturer's instructions. Dissociated colon cells were strained with a 100µm cell strainer (Falcon, Corning, USA) to yield a single cell suspension. Cells were plated on fibronectin-coated plates (Invitrogen, Thermo Fisher Scientific, USA) to form neurosphere-like bodies, as previously described[31]. Medium was refreshed every 2-3 days and cells were expanded *in vitro* for 1-4 weeks. The culture was split using Accutase (Sigma Aldrich, USA). In order to isolate ENCCs from the mixed cell population the culture was dissociated with Accutase, strained with a 100µm cell strainer, washed in PBS containing 10% FCS and stained with an antibody against $p75^{NTR}$, a neural crest and early neural marker, conjugated with phycoerythrin (1:100, ab157333, Abcam, USA). Cells were sorted using a BD FACSAria™ III (BD Biosciences, USA), and snap-frozen in liquid nitrogen for DNA isolation.

Amplicon based exome sequencing

Genomic DNA was isolated from peripheral blood cells using standard methods. Genomic DNA was isolated from fibroblasts and ENCCs using a QIAamp DNA Micro kit (Qiagen, NL) according to manufacturer's instructions. DNA libraries for exome sequencing were constructed using 250ng of germline DNA (blood or fibroblast derived) and 250 ng of DNA isolated from cultured ENCCs captured with the Haloplex exome target enrichment kit (Agilent Technologies, USA). Captured fragments were sequenced [paired-end 101 base pair (bp) read length] on the Illumina HiSeq2500sequencers (Illumina, San Diego, USA). The raw sequence data was processed using the Nimbus Suite, which is specifically designed to process, align and call variants

from amplicon-based sequence data[32]. Reads were aligned to the hg19 reference sequence, and alternative as well as reference alleles were counted per genomic position.

Data analysis and selection of somatic variants

Exome sequence data was first analysed to exclude (likely-) pathogenic variants[33,34] in known HSCR disease genes (see supplementary table 2 for the list of these genes). Next, somatic mutations were determined based on sequencing quality. When determining cell-type specific somatic mutations (ENCC or germline), the alternate allele had to be present at least 5 times in two different amplicons, be present in at least 10% of reads, and be absent in the other cell type. The minimal coverage per base used in the analysis was 20X in both cell types. Furthermore, we only considered protein-altering variants with a minor allele frequency below 0.001 in GnomAD exome and/or GnomAD genome (http://gnomad.broadinstitute.org/). The following criteria were subsequently used to prioritize the putative somatic mutations on deleteriousness and involvement of the gene in ENS development: changes in putative loss of function or predicted to affect splicing[35] in intolerant genes[36]; changes with a CADD[37] v1.4 score of 15 or higher (http://cadd.gs.washington.edu/home); and/or changes with a predicted deleterious in the best performing prediction tools from the first three clusters previously described[37-45]. mRNA expression in the developing ENS was inspected using publically available mouse data sets (Gene expression omnibus: GSE34208 and GSE111307). Prioritized genes were the ones with a mouse orthologue differentially expressed between (a) E14.5 intestine or ENS cells[46], or (b) E11.5 and E15.5 ENS, progenitors or intestine[47]. We also used the RNA sequencing data

available in-house for human embryonic intestine at embryonic week 12,14 and 16 (McCann et al. unpublished data).

Validation of putative mosaic differences

Top ranking ENCC-specific variants (based on either quality or deleteriousness) were validated using Sanger sequencing as previously described[48]. The primers are available on request.

Analysis of somatic copy number changes

Germline DNA of all five patients was inspected for the presence of rare deleterious copy number changes. Moreover, ENCC derived DNA was also inspected for patient 1, 2 and 3. An insufficient amount of ENCC DNA was available for patients 4 and 5. For this purpose, we used either the HumanOmni5-4_v1.1 beadchip (patient 1 and 2) or the Infinium Global Screening Array-24 v1.0 (patient 3, 4 and 5) (Illumina Inc., San Diego, CA, USA). All protocols and procedures were performed as previously described[49]. CNV profiles were inspected visually in Biodiscovery Nexus CN8.0 (Biodiscovery Inc., Hawthorne, CA, USA), with special focus on allele frequency differences between copy number changes present in germline and ENCC-derived DNA.

Results

Exome sequence to identify germline mutations

For this study, five HSCR patients were selected to undergo exome sequencing (WES). The quality of the data generated can be seen in Table 1. No obvious deleterious germline mutation was identified in any of the known HSCR associated genes or in genes highly expressed in mouse ENCCs[25,50]. However, we did find three protein altering and ten synonymous variants in genes previously associated with HSCR (Supplementary table 2). Twelve of the thirteen variants identified were believed to be benign/mild and only one was predicted to be deleterious. This variant is located in the glial cell line-derived neurotrophic factor receptor alpha 1 (*GFRα1*) and leads to loss of its starting codon. *GFRα1* encodes for an extracellular protein that works as receptor for the glial cell line-derived neurotrophic factor (GDNF), and is required for RET activation. Although *GFRα1* is considered to be a HSCR candidate gene, no pathogenic mutations have been found in HSCR patients, making it difficult for us to assess the contribution of the variant identified to the overall phenotype. Due to the fact that such variant in a diagnostic setting would not be considered causative, we concluded that none of the five patients analysed had an obvious pathogenic germline mutation in a known HSCR gene, which alone could explain the phenotype.

Identification of somatic mutations in HSCR

In order to detect the presence of somatic mosaicism in our cohort of patients, genetic variants identified in DNA from blood or abdominal fibroblasts were compared to the ones found in DNA from ENCCs. Cell-type specific alternative alleles were identified (Table 2). Putative somatic variants were also found in four of the five patients analysed

Table 1. Exome sequencing quality information

Patient	Cell type	reads	Mapped reads	Mapped reads (%)	Mean coverage	20X coverage (n bases)	Overlapping 20X coverage	Variants (n)#	Overlapping variants (n)#	Concordance rate 20 X (%)
1	blood	43692456	42653058	97.6	-	1009526	891280	77521	65576	98.4
1	ENCC	-	-	-	-	936993		87810		-
2	blood	46663496	45464802	97.4	-	1052613	1005681	82750	73769	99.0
2	ENCC	51612528	50248904	97.4	-	1056654		85342		-
3	blood	73084700	72360584	99.0	-	2951928	2765133	92565	64458	98.8
3	ENCC	71473344	70723194	99.0	-	2887416		95400		-
4	fibroblasts	71515650	70770498	99.0	-	2896756	1988647	95491	65780	98.6
4	ENCC	68160814	58726936	86.2	-	2068315		109987		-
5	blood	72927952	72048920	98.8	-	2917188	2661452	96952	84900	99.6
5	ENCC	69489918	68471026	98.5	-	2768779		97171		-

Table 1. Depicted are the number of reads sequenced and mapped to the hg19 reference genome. Coverage statistics and ENCC-germline concordance rate at 20X coverage represent the starting amount of bases at which somatic mosaicism was evaluated. #variants counted with a read depth of five or higher at 20X coverage.

Table 2. Presence of alternative alleles in ENCCs

Patient	All variation alleles in ENCC only	All variation alleles in blood only	protein altering alleles in ENCC only[#]	protein altering alleles in blood only[#]	Variants in ENCC meeting the prioritization criteria[©]	Number of variant sequenced and validated
1[$]	50	43	8	11	5	0
2[$]	16	28	2	4	1	0
3[$]	25	33	0	1	0	0
4[#$]	96	178	17	11	15	0[$$]
5[$]	29	35	2	1	2	0

Table 2. Depicted are the number of variants passing subsequent quality and prioritization criteria. [#] see filtering steps describing the selection criteria for the best somatic candidate variants.; [©] see variant prioritization critera, [$$] No ENCC DNA available for Sanger sequencing validation, only exclusion of the variant in fibroblast derived DNA.

(Table 3). However, validation of any of these variants by Sanger sequencing failed to confirm their presence in ENCCs, leading us to conclude that no real somatic variants were present in our cohort of patients.

Analysis of somatic copy number changes

In total, 12 rare germline Copy Number Variants (CNVs) were detected in our cohort of patients. Each patient holds at least one rare CNV (see supplementary table 3). However, only the 1q25.3 gain (patient 1), the 14q24.1 gain on 2q24.1 (patient 2), the 10p25.1 gain (patient 3) and the 6p22.3 gain (patient 5) contain genes and can be classified as variant of unknown significance. None of the CNVs impacted the HSCR disease genes described in Table 2. New copy number changes were not detected in the ENCC population of the patients tested (1, 2 and 3). Inspection of the allele frequencies of germline and ENCC-derived profiles in patient 1, 2 and 3 did not reveal any significant differences.

Table 3. Putative somatic variants passing quality filtering and variant prioritization

Patient	Gene	cDNA	Type	dbSNP	Class	gnomAD exome	gnomAD genome		pLI	Human	Mouse
1	*FMN2*	c.162delC	Frameshift deletion	.	Likely Deleterious	0.000000	0.000000	1.42	0.99	no	yes
1	*YWHAE*	c.G142A	Missense	.	Likely Deleterious	0.000000	0.000000	3.25	0.96	yes	.
1	*YWHAE*	c.T116C	Missense	.	Likely Deleterious	0.000000	0.000000	3.25	0.96	yes	.
1	*PHAX*	c.C379T	premature stop codon	.	Variant of Unknown Significance	0.000000	0.000000	-0.51	0.00	yes	.
1	*POR*	c.T1231C	Missense	.	Variant of Unknown Significance	0.000000	0.000000	-0.54	0.00	yes	.
2	*DEPDC1*	c.T1459A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	-0.32	0.00	yes	.
4	*F5*	c.A1867G	Missense	.	Variant of Unknown Significance	0.000000	0.000000	-1.30	0.00	no	.
4	*PHRF1*	c.G1075A	Missense	rs551874512	Variant of Unknown Significance	0.000033	0.000032	-1.36	0.95	yes	.
4	*MYBPC3*	c.C482A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	0.69	0.00	no	.
4	*PACS1*	c.G1069A	Missense	rs750459659	Likely Deleterious	0.000041	0.000032	4.32	1.00	yes	.
4	*OAS3*	c.C1390T	Missense	rs750291946	Variant of Unknown Significance	0.000012	0.000000	-0.60	0.00	yes	.
4	*MAN2A2*	c.G478A	Missense	rs374688808	Variant of Unknown Significance	0.000012	0.000032	1.28	0.00	yes	yes
4	*SNF8*	c.G578A	Missense	rs775611332	Variant of Unknown Significance	0.000025	0.000000	0.97	0.29	yes	yes
4	*MED15*	c.C730A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	2.50	0.96	yes	.
4	*IQCF5*	c.C283T	Missense	rs772101978	Variant of Unknown Significance	0.000100	0.000000	-1.59	0.43	no	.
4	*TMEM165*	c.C782A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	1.83	0.94	yes	.
4	*NOTCH4*	c.G1118A	Missense	rs745883985	Variant of Unknown Significance	0.000033	0.000032	2.45	0.00	yes	.
4	DPPA5	c.G214A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	1.64	0.00	no	.
4	SLC22A1	c.C523T	Missense	rs768905186	Variant of Unknown Significance	0.000004	0.000000	-0.28	0.00	no	.
4	MGAM2	c.G3015T	Missense	.	Variant of Unknown Significance	0.000000	0.000000	.	.	unknown	.
4	IKBKB	c.G809A	Missense	rs200841053	Variant of Unknown Significance	0.000024	0.000032	2.90	1.00	yes	.
5	PCDH15	c.G139A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	-3.27	0.00	no	yes
5	ZNF592	c.C3433A	Missense	.	Variant of Unknown Significance	0.000000	0.000000	1.10	0.95	yes	.

Table 3. Depicted are all variants passing quality criteria and prioritization based on predicted deleteriousness, expression pattern and the sensitivity for a gene to rare variation. More detailed information is available in supplementary table 3a and 3b. No variants passed the quality criteria for putative somatic mutations in patient 3. Human expression based on logCPM in EW12-16. Mouse based on expression in the ENS in E11-15.5

Table 4. Rare germline variants in HSCR disease genes

Patient	Gene	cDNA	Type	dbSNP	gnomAD exome	gnomAD genome	Cadd	Class
1	*TBATA*	c.T666C	Synonymous	rs2254433	0.26	0.30	-	LB
1	*VCL*	c.G2388A	Synonymous	rs767809	0.48	0.43	-	LB
1	*ZEB2*	c.G3480A	Synonymous	-	.	.		LB
2	*GFRA1**	c.A1T	Start site loss	-	.	.	23	VUS
2	*ZEB2*	c.G3480A	Synonymous	-	.	.	-	LB
2	*DENND3*	c.C1110T	Synonymous	rs2289001	0.34	0.27	-	LB
4	*NRG3**	c.C59G	Missense	rs1884282	0.12	0.13	3.3	LB
4	*NRG3*	c.A1770G	Synonymous	rs17101196	0.08	0.08	-	LB
4	*NRG3*	c.C1986T	Synonymous	rs2295933	0.38	0.38	-	LB
4	*EDNRB**	c.G1392A	Missense	-			25.7	VUS
4	*ZEB2*	c.G3480A	Synonymous	-	.	.	-	LB
4	*DENND3*	c.G3090A	Synonymous	rs1045303	0.34	0.28	-	LB
5	*NUP98*	c.G2688A	Synonymous	rs35803045	0.05	0.05	3.8	LB

Table 4. Variants in known HSCR disease genes. All variants are heterozygous and detected in both cell types. Variants marked with * are tested and confirmed with Sanger sequencing. Abbreviations: LB; Likely Benign, VUS; Variant of uncertain significance, Combined Annotation Dependent Depletion (CADD) phred based value can be used to evaluate the likelihood a variant is deleterious (http://cadd.gs.washington.edu/), Population frequency depicted is the using GnomAD genome sequence value (http://gnomad.broadinstitute.org/).

Discussion

Establishing the involvement of somatic mosaicism in congenital malformations can explain the disease occurrence in the absence of inherited or *de novo* coding mutations, but it is also important for counselling, as the recurrence risk for a tissue-specific somatic mutation is null, while a germline somatic mutation can still be present in the remaining germ cells and thus can be transmitted. In this study, we investigated if somatic variants substantially contribute to the development of HSCR, a congenital malformation of the ENS. We excluded pathogenic germline variants in known HSCR disease genes and specifically searched for the presence of somatic mutations in ENCCs isolated from gastrointestinal (GI) biopsies obtained from a set of five HSCR patients. As somatic variants can originate at any stage of life, a fraction of all human cells are likely to carry a variant. Therefore, we reasoned that in order to be missed in a diagnostic setting, these possible somatic mutations should be present in ENC-derived cells, but absent in blood cells or fibroblasts. Exome sequencing on DNA isolated from purified ENCCs and blood or fibroblasts of these patients resulted in a set of putative somatic mutations. However, none of these could be confirmed by traditional methods, suggesting that the variants identified are technical artefacts or technical noise. This result led us to conclude that in our patients, somatic variants do not play a role in HSCR development. As this is a very small number of patients we cannot generalize and exclude the involvement of somatic variants in HSCR. However, there are a substantial number of reasons to assume that if somatic variants were to be the cause of HSCR they would likely be extremely difficult to detect. In this discussion we elaborate further on this, as well as on the detection and possible

contribution of somatic variants for the development of congenital disorders.

How to detect somatic mosaicism?

In order to detect somatic mosaicism, analysis of multiple tissues within an individual is required. In some cases, the choice of tissue is suggested by recognition of a phenotype, for example in HSCR the relevant cells to be investigated are ENCCs. However, somatic mosaicism can also be searched for in the affected tissue only using sensitive genotyping techniques such as, single-nucleotide polymorphism (SNP) microarrays or next-generation sequencing (NGS), which can detect low-level mosaicism in a more routine fashion[12,51,52]. The downside of using such methods in a single tissue only is that they will not prove the somatic nature of a presumed somatic variant. This is due to the fact that discriminating technical artefacts from real somatic variants is far from easy. Therefore, analysis of multiple tissues within an individual is a better option. It should be noted though that somatic variants might be present in low frequencies in multiple, if not all, tissues, including the affected tissue. However, if this is the case they would have arisen at or before the epiblast stage of embryogenesis, and be virtually indistinguishable from a mutation arisen in the germ cells of one of the parents. Technical distinction of these two types of mosaicisms in such cases is challenging. For instance, if alternative alleles are present in the DNA of both blood and ENCCs, the allele frequency will likely be high and fall within the normal range (between 20-70%), meaning that they will appear as *de novo* heterozygous variants. To circumvent these issues, we opted to search for somatic variants (alternative alleles) present in the affected tissue only, by comparing WES data from ENCCs and

blood or fibroblasts of the same patient. Even though we were unable to identify real somatic variants in the patients analysed, we believe that this is not dependent on our experimental design, but simply due to the fact that somatic variants do not contribute to HSCR development in these patients.

However, one could argue that this is not the case and is simply due to the sensitivity of Sanger sequencing. In our hands the sensitivity of this technique for a known variant is approximately 10%, and we set the minimal variant quality for validation of somatic changes accordingly. Therefore, we should have been able to detect them if they were present. Of course, we cannot exclude that some variants might be missed if the real allele frequency is lower than 10%, but this is unlikely to be true for all validated variants. A more likely explanation for the lack of validation is that we are looking at sequencing artefacts. This 'sequencing noise' is inherent to WES. Although the number of such variants looks high, it is in fact extremely low compared to the total number of true variants identified, as the concordance rate between cell types at 20X coverage is between 98.6 and 99.6%. Inclusion of parental information would help reducing the noise substantially, as it would allow us to pinpoint *de novo* variant(s) and better discriminate artefacts from true somatic mosaicism. However, this was only possible for one patient (patient 5).

ENS-specific somatic changes in DNA Copy Number.

Somatic mosaic CNVs have been described previously[53]. We know that the human brain is especially sensitive for such events[54,55], and hypothesized that this is also the case for ENCCs. The detection limit for new changes is close to 10% and we can detect allele specific

differences between cell types close to 20%[56]. In our study, we could identify several germline changes in DNA copy number, most of them were known polymorphisms and not related to HSCR. A few rare germline CNVs were found, but they were not likely related to the development of HSCR. Moreover, none of the germline changes differed in allele frequency in the ENCCs, nor did we find new ENCC specific alterations. In conclusion, we could not find or confirm any somatic change in ENCCs, neither small variants nor larger changes in DNA Copy Number.

Would ENC-derived cells with a somatic variant remain to be sampled?

To answer this question, we have to consider two parameters: the development of the ENS and the different models that could represent the effect of a somatic variant on the subsequent distribution of the mutated cells. The ENS is mainly formed from the vagal neural crest with a small contribution from the sacral neural crest. Cells at the neural crest bud off from the folding neural tube to migrate towards the cranial end of the developing gut (Figure 2)[57,58]. These ENCCs migrate further in the gut mesoderm layer following signals to proliferate along the gut tube. As the gut is also elongating, the developing ENS cells are highly migratory, with individual cells migrating locally in all directions although the net direction as a population is towards the caudal end. The leading wave front sets the tracks for the other cells to follow, determining the position of the later ganglionic network[59]. A previous study in mice has shown the ENCC migration to the proximal midgut allow for a trans-mesenteric 'short-cut' migration of a relatively small number of cells direct to the hindgut, because the midgut and hindgut form a U-shaped loop. Once in the hindgut the trans-mesenteric ENCCs dominate the colonization

of this region as they proliferate and spread not only distally to the rectum but also back proximally[60]. This may also occur in humans since the U-shaped intestinal loop is also present at equivalent embryonic stages.

This complex development of the ENS can be disrupted by germline defects which affect all ENS precursor cells. A somatic mutation would only affect the cells containing the mutation and thus, the effect would be dependent on when and in which lineage the variant arose, as well as which proteins and functions the variant affects directly and indirectly. There are, of course, countless possible models which could represent the effect of a somatic variant in HSCR, but we believe that they can be grouped into three broad categories: no selective effect, selective advantage, and selective disadvantage (Figure 3).

The first model, no selective effect, can be simply explained by thinking that the variant will always be passed on to the daughter cells without any selection effect against the mutated cell. This model is somewhat comparable to a PCR reaction (Figure 3A). When a variant is introduced in the first PCR cycle, its detection after 30 cycles depends on the number of copies of that specific allele on the starting DNA fraction. Detecting such variants can only be done by sequencing single products, for instance via cloning of the PCR product. In a no selective effect model the phenotype could only be seen if the variant was to occur very early in development when there are few cells in total[13], or if there is some other mechanism for restricting clonal variance. Schemes for the later have been proposed specifically for the ENS and the previously mentioned trans-mesenteric migration could also restrict clonal variety in the colon[60,61]. Variants that occur later in

life will only affect the progeny of that specific cell and might be difficult to identify. This model is the simplest one, and is likely not representative for progenitors of the ENS, as we believe that without a selective effect the population of mutated cells would be able to colonize the entire gut and would not result in HSCR.

Mutations that result in a selective advantage for the cell are also not likely to be present in HSCR, as the disease is characterized by a loss of cells in the ENS (Figure 3B). Selective variants are a hallmark of cancer, and are also seen in overgrowth syndromes[12]. However, if selective variants were to occur, they should be detectable with a high allele frequency[62], and this is not the case for HSCR patients. Interestingly, for normal ENS development a stochastic model involving superstar cells have been proposed and proven using both computer-modelling and grafting experiments. It was shown that the eventual ENS is formed by just a few cells which therefore were named superstars[61].

We believe that the opposite is more likely to occur in HSCR, namely that the variant causes a selective disadvantage for affected cells. In this model, we hypothesize that a variant could result in problems in migration, proliferation or ENS differentiation. Migration defects would render the affected cells incapable of following migration signals or physically migrating, leaving the unaffected cells to grow but affecting the total number of cells able to reach the distal gut. A variant affecting the cell cycle may cause a decrease or arrest in division. Variants leading to decreased cell viability would result in apoptotic pathway activation and death of the affected cells (Figure 3C). All affecting the total number of cells able to colonize the gut.

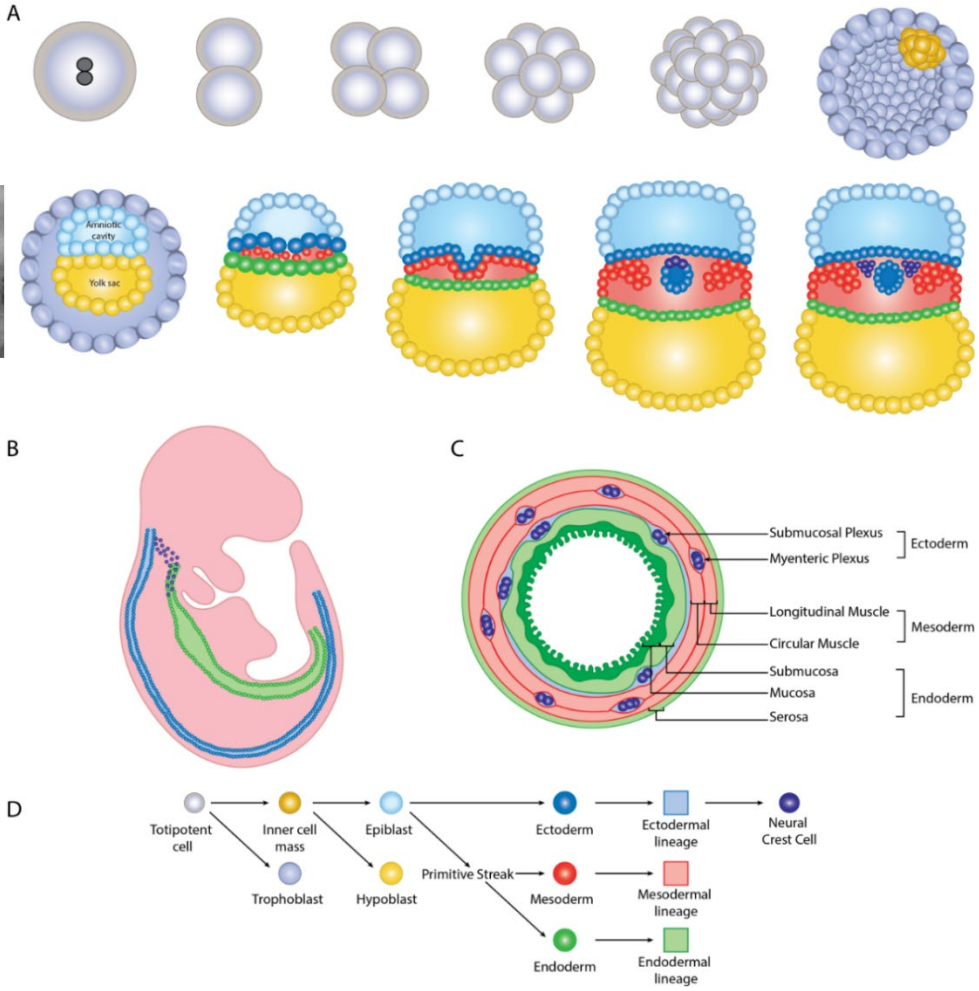


Figure 2. Schematic representation of early embryonic development. A) The first divisions and differentiations of the developing embryo depicting the formation of the three germ layers and the neural crest cells. B) The highly proliferative and migratory ectodermal neural crest begins to enter the endodermal gut tube at week four of human gestation. Some of these cells will also contribute to the neural cells of the lungs and the pancreas [1-4]. C) Cross section of colon to highlight contribution from all three germ layers with the lowest contribution being from ectoderm/neural crest. D) Legend of cell type and their origins.

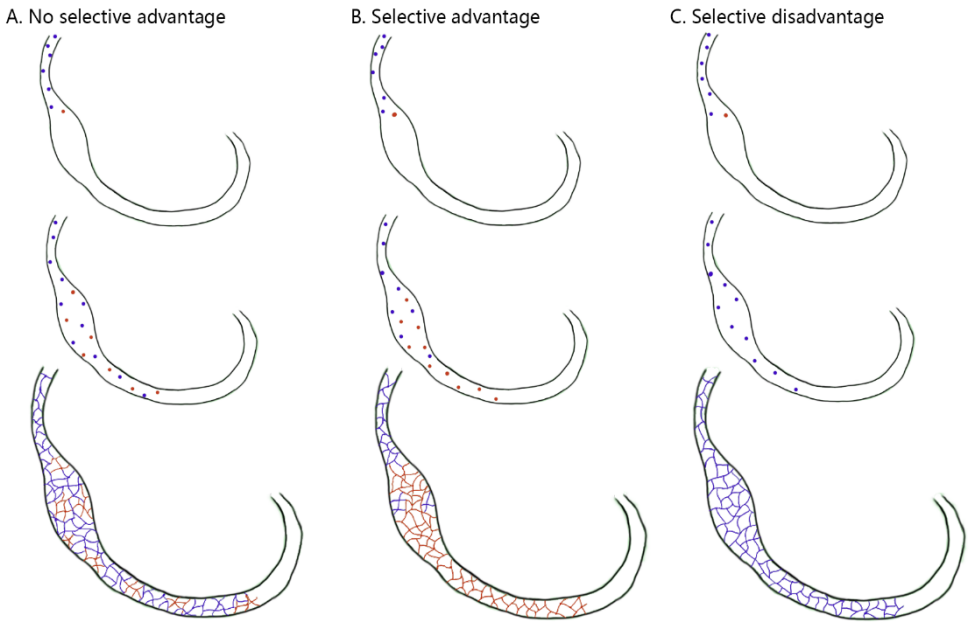


Figure 3. Proposed models of somatic mutations leading to HSCR. A) Non-selective clonal growth and division, the variant gives no competitive advantage or disadvantage during migration or proliferation. However, this may lead to decreased differentiation or maturation capacity, for example. B) Selective advantage, the variant confers a proliferation or migratory advantage to the cells allowing them to out-compete the wild-type cells. However due to early differentiation, altered subtype specification or inability to follow extracellular signals, the cells are unable to reach the distal portion of gut. C) Selective disadvantage, the variant affects the viability of the cells leading to apoptosis at some stage of migration or differentiation. This leads to a decreased total cell count leaving too few cells to fully colonise the length of the gut tube. This is the most likely model that would result in a HSCR phenotype.

However, the effect of a variant would also depend on which cell type expresses the affected gene or pathway, as it may only be after a certain point in differentiation that an apoptotic pathway would be activated. If using this model we expect an aganglionosis to occur due to somatic changes of the vagal neural crest cells, we know that the sacral neural crest will also be unable to colonize the distal portion of colon[63]. Based on this evidence, we hypothesize that if an unfavourable variant were to have occurred very early in development or occurred in a superstar cell, the chances of detecting such variant in the ganglionic biopsies sampled as part of this study, are incredibly low, as the affected cells are unlikely to have reached the end of the GI tract. In this case, it is likely that the total number of precursor cells migrating into the GI tract is too low to exceed the critical number of cells required for normal ENS formation[64]. Similarly, if the variant prohibited ENCC specification to ganglionic cell lineages, these cells would not exist. These outcomes might, of course, also result in low allele frequencies due to the low number of cells that would reach the distal colon. More proximal regions of the GI tract would have therefore had to be examined in order to find higher allele frequencies. Yet another possibility is that a small subset of cells acquire a genetic change that could then functionally affect the cells that they interact or synapse with, giving a larger functional effect than could be predicted by sequencing. Single-cell sequencing techniques might provide a powerful alternative to mass sequencing modalities in order to give a true picture of the extent of mosaicism present in various tissues[65].

Conclusions

In this study we investigated the involvement of somatic mosaicism for a congenital disorder of the gut, HSCR. Although we were unable to confirm this involvement in our small patient cohort, we still consider that somatic variants could possibly play a role in HSCR. As new technologies emerge the use of smaller concentrations of DNA for sequencing will be possible, allowing for the inclusion of suction biopsies from multiple regions along the GI tract. This, together with the appropriate separation of cell lineages, selection of sequencing modality and filtering of variants, will make it easier to determine whether somatic variants play a role in congenital GI tract disorders and if their occurrence rate can explain at least part of the missing-heritability seen for most of these disorders.

References

1. Burns AJ, Delalande J-M. Neural crest cell origin for intrinsic ganglia of the developing chicken lung. *Developmental Biology.* 2005;277(1):63-79.
2. Kirchgessner AL, Adlersberg MA, Gershon MD. Colonization of the developing pancreas by neural precursors from the bowel. *Developmental Dynamics.* 1992;194(2):142-154.
3. Jiang Y, Liu M-t, Gershon MD. Netrins and DCC in the guidance of migrating neural crest-derived cells in the developing bowel and pancreas. *Developmental Biology.* 2003;258(2):364-384.
4. Neural Crest and the Development of the Enteric Nervous System. Landes Bioscience; 2013. https://www.ncbi.nlm.nih.gov/books/NBK6273/.
5. Theda C, Hwang SH, Czajko A, Loke YJ, Leong P, Craig JM. Quantitation of the cellular content of saliva and buccal swab samples. *Scientific Reports.* 2018;8(1):6944.
6. Sciences L. Organ/Tissue Ontology Tree. LifeMap Sciences Inc. https://discovery.lifemapsc.com/in-vivo-development/organ-tissues. Published 2012. Accessed 15-08-2018, 2018.
7. Dumitrescu CE, Collins MT. McCune-Albright syndrome. *Orphanet J Rare Dis.* 2008;3(1):12.
8. O'Malley MP, Haake A, Goldsmith L, Berg D. Localized darier disease: Implications for genetic studies. *Arch Dermatol.* 1997;133(9):1134-1138.
9. Rohrback S, Siddoway B, Liu CS, Chun J. Genomic mosaicism in the developing and adult brain. *Dev Neurobiol.* 2018.
10. D'Gama AM, Walsh CA. Somatic mosaicism and neurodevelopmental disease. *Nature Neuroscience.* 2018:1.
11. Vijg J. Somatic mutations, genome mosaicism, cancer and aging. *Current opinion in genetics & development.* 2014;26:141-149.
12. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nature Reviews Genetics.* 2013;14:307.
13. Erickson RP. Somatic gene mutation and human disease other than cancer: an update. *Mutation Research/Reviews in Mutation Research.* 2010;705(2):96-106.
14. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. *Nature Reviews Genetics.* 2002;3:748.
15. Newgreen DF, Zhang D, Cheeseman BL, Binder BJ, Landman KA. Differential Clonal Expansion in an Invading Cell Population: Clonal Advantage or Dumb Luck? *Cells Tissues Organs.* 2017;203(2):105-113.
16. Fernández LC, Torres M, Real FX. Somatic mosaicism: on the road to cancer. *Nature Reviews Cancer.* 2016;16(1):43.
17. Holschneider AM, Puri P. *Hirschsprung's disease and allied disorders.* Springer Science & Business Media; 2007.
18. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental Biology.* 2016;417(2):198-208.
19. Emison ES, Garcia-Barcelo M, Grice EA, et al. Differential Contributions of Rare and Common, Coding and Noncoding Ret Mutations to Multifactorial Hirschsprung Disease Liability. *The American Journal of Human Genetics.* 2010;87(1):60-74.

20. Hofstra RM, Wu Y, Stulp RP, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. *Human mutation.* 2000;15(5):418-429.
21. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.
22. Tang CS, Gui H, Kapoor A, et al. Trans-ethnic meta-analysis of genome-wide association studies for Hirschsprung disease. *Human molecular genetics.* 2016;25(23):5265-5275.
23. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease. *Cell.* 2016;167(2):355-368.e310.
24. Kapoor A, Jiang Q, Chatterjee S, et al. Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms. *Human molecular genetics.* 2015;24(10):2997-3003.
25. Alves MM, Sribudiani Y, Brouwer RWW, et al. Contribution of rare and common variants determine complex diseases—Hirschsprung disease as a model. *Developmental Biology.* 2013;382(1):320-329.
26. Sribudiani Y, Chauhan RK, Alves MM, et al. Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease. *Gastroenterology.* 2018;155(1):118-129.e116.
27. Moore SW, Zaahl MG. Tissue specific somatic mutations and aganglionosis in Hirschsprung's disease. *Journal of Pediatric Surgery.* 2014;49(2):258-261.
28. Zhang Z, Jiang Q, Li Q, et al. Genotyping analysis of 3 RET polymorphisms demonstrates low somatic mutation rate in Chinese Hirschsprung disease patients. *Int J Clin Exp Pathol.* 2015;8(5):5528.
29. Jiang Q, Liu F, Miao C, et al. RET somatic mutations are underrecognized in Hirschsprung disease. *Genet Med.* 2018;20(7):770-777.
30. Brosens E, MacKenzie KC, Alves MM, Hofstra RMW. Do RET somatic mutations play a role in Hirschsprung disease? *Genet Med.* 2018.
31. Metzger M, Caldwell C, Barlow AJ, Burns AJ, Thapar N. Enteric Nervous System Stem Cells Derived From Human Gut Mucosa for the Treatment of Aganglionic Gut Disorders. *Gastroenterology.* 2009;136(7):2214-2225.e2213.
32. Brouwer RWW, van den Hout M, Kockx CEM, et al. Nimbus: A design-driven analyses suite for amplicon based NGS data. *Bioinformatics (Oxford, England).* 2018.
33. MacArthur DG, Manolio TA, Dimmock DP, et al. Guidelines for investigating causality of sequence variants in human disease. *Nature.* 2014;508(7497):469-476.
34. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2015;17(5):405-424.
35. Jian X, Liu X. In Silico Prediction of Deleteriousness for Nonsynonymous and Splice-Altering Single Nucleotide Variants in the Human Genome. *Methods in molecular biology (Clifton, NJ).* 2017;1498:191-197.

36. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291.
37. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature genetics.* 2014;46(3):310-315.
38. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. *Human mutation.* 2016;37(3):235-241.
39. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. *BMC genomics.* 2013;14 Suppl 3:S3.
40. Douville C, Masica DL, Stenson PD, et al. Assessing the Pathogenicity of Insertion and Deletion Variants with the Variant Effect Scoring Tool (VEST-Indel). *Human mutation.* 2016;37(1):28-35.
41. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics (Oxford, England).* 2015;31(5):761-763.
42. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nature methods.* 2014;11(4):361-362.
43. Shihab HA, Rogers MF, Gough J, et al. An integrative approach to predicting the functional effects of non-coding and coding sequence variation. *Bioinformatics (Oxford, England).* 2015;31(10):1536-1543.
44. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nature methods.* 2010;7(4):248-249.
45. Dong C, Wei P, Jian X, et al. Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human molecular genetics.* 2015;24(8):2125-2137.
46. Schriemer D, Sribudiani Y, A IJ, et al. Regulators of gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes. *Developmental biology.* 2016;416(1):255-265.
47. Memic F, Knoflach V, Morarach K, et al. Transcription and Signaling Regulators in Developing Neuronal Subtypes of Mouse and Human Enteric Nervous System. *Gastroenterology.* 2018;154(3):624-636.
48. Widowati T, Melhem S, Patria SY, et al. RET and EDNRB mutation screening in patients with Hirschsprung disease: Functional studies and its implications for genetic counseling. *European Journal Of Human Genetics.* 2015;24:823.
49. Brosens E, Marsch F, de Jong EM, et al. Copy number variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula. *Eur J Hum Genet.* 2016;24(12):1715-1723.
50. Schriemer D, Sribudiani Y, Ijpma A, et al. Regulators of gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes. *Developmental biology.* 2016.
51. Conlin LK, Thiel BD, Bonnemann CG, et al. Mechanisms of mosaicism, chimerism and uniparental disomy identified by single nucleotide polymorphism array analysis. *Human Molecular Genetics.* 2010;19(7):1263-1275.

52. Gottlieb B, Beitel LK, Alvarado C, Trifiro MA. Selection and mutation in the "new" genetics: an emerging hypothesis. *Human Genetics.* 2010;127(5):491-501.
53. Campbell IM, Yuan B, Robberecht C, et al. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am J Hum Genet.* 2014;95(2):173-182.
54. McConnell MJ, Lindberg MR, Brennand KJ, et al. Mosaic copy number variation in human neurons. *Science (New York, NY).* 2013;342(6158):632-637.
55. Villela D, Suemoto CK, Leite R, et al. Increased DNA Copy Number Variation Mosaicism in Elderly Human Brain. *Neural plasticity.* 2018;2018:2406170.
56. Veenma D, Beurskens N, Douben H, et al. Comparable low-level mosaicism in affected and non affected tissue of a complex CDH patient. *PLoS One.* 2010;5(12):e15348.
57. Young HM, Hearn CJ, Newgreen DF. Embryology and development of the enteric nervous system. *Gut.* 2000;47(suppl 4):iv12-iv14.
58. Le Douarin NM, Teillet M-A. The migration of neural crest cells to the wall of the digestive tract in avian embryo. *J Embryol Exp Morphol.* 1973;30(1):31-48.
59. Lake JI, Heuckeroth RO. Enteric nervous system development: migration, differentiation, and disease. *American Journal of Physiology-Gastrointestinal and Liver Physiology.* 2013;305(1):G1-G24.
60. Nishiyama C, Uesaka T, Manabe T, et al. Trans-mesenteric neural crest cells are the principal source of the colonic enteric nervous system. *Nature Neuroscience.* 2012;15:1211.
61. Cheeseman BL, Zhang D, Binder BJ, Newgreen DF, Landman KA. Cell lineage tracing in the developing enteric nervous system: superstars revealed by experiment and simulation. *Journal of The Royal Society Interface.* 2014;11(93).
62. Lindhurst MJ, Parker VER, Payne F, et al. Mosaic overgrowth with fibroadipose hyperplasia is caused by somatic activating mutations in PIK3CA. *Nature Genetics.* 2012;44:928.
63. Burns AJ, Champeval D, Le Douarin NM. Sacral Neural Crest Cells Colonise Aganglionic Hindgut in Vivo but Fail to Compensate for Lack of Enteric Ganglia. *Developmental Biology.* 2000;219(1):30-43.
64. Barlow AJ, Wallace AS, Thapar N, Burns AJ. Critical numbers of neural crest cells are required in the pathways from the neural tube to the foregut to ensure complete enteric nervous system formation. *Development.* 2008;135(9):1681-1691.
65. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics.* 2013;14:618.

Supplementary methods

Data analysis and selection of somatic variants

WES data was analysed to exclude (likely-)pathogenic variants[1, 2] in known HSCR disease genes (see Supplementary table 1). Next, somatic variations were determined based on sequencing quality. When determining cell-type specific somatic variations (ENCC or germline), the alternate allele had to be present at least 5 times in 2 amplicons, be present in at least 10% of reads, and be absent in the other cell-type. The minimal coverage per base used in the analysis was 20X in both cell-types. We only considered protein-altering variants with a minor allele frequency below 0.001 in GnomAD exome and/or GnomAD genome (http://gnomad.broadinstitute.org/). The following criteria were subsequently used to prioritize the putative somatic variations on deleteriousness and involvement of the gene in ENS development: changes in putative loss of function or predicted to affect splicing[3] in intolerant genes[4]; changes with a CADD[5] v1.4 score of 15 or higher (http://cadd.gs.washington.edu/home); and/or changes with a predicted deleterious in the best performing prediction tools from the first three clusters previously described[5-13]. mRNA expression in the developing ENS was inspected using publicly available mouse data sets (Gene expression omnibus: GSE34208 and GSE111307). Prioritized genes were the ones with a mouse orthologue differentially expressed between (a) E14.5 intestine or ENS cells[14], or (b) E11.5 and E15.5 ENS, progenitors or intestine[15]. We also used the RNA sequencing data available in-house for human embryonic

intestine at embryonic week 12,14 and 16 (McCann et al. unpublished data).

References

1. MacArthur, D.G., et al., *Guidelines for investigating causality of sequence variants in human disease.* Nature, 2014. **508**(7497): p. 469-76.
2. Richards, S., et al., *Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.* Genet Med, 2015. **17**(5): p. 405-24.
3. Jian, X. and X. Liu, *In Silico Prediction of Deleteriousness for Nonsynonymous and Splice-Altering Single Nucleotide Variants in the Human Genome.* Methods Mol Biol, 2017. **1498**: p. 191-197.
4. Lek, M., et al., *Analysis of protein-coding genetic variation in 60,706 humans.* Nature, 2016. **536**(7616): p. 285-91.
5. Kircher, M., et al., *A general framework for estimating the relative pathogenicity of human genetic variants.* Nat Genet, 2014. **46**(3): p. 310-5.
6. Liu, X., et al., *dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs.* Hum Mutat, 2016. **37**(3): p. 235-41.
7. Carter, H., et al., *Identifying Mendelian disease genes with the variant effect scoring tool.* BMC Genomics, 2013. **14 Suppl 3**: p. S3.
8. Douville, C., et al., *Assessing the Pathogenicity of Insertion and Deletion Variants with the Variant Effect Scoring Tool (VEST-Indel).* Hum Mutat, 2016. **37**(1): p. 28-35.
9. Quang, D., Y. Chen, and X. Xie, *DANN: a deep learning approach for annotating the pathogenicity of genetic variants.* Bioinformatics, 2015. **31**(5): p. 761-3.
10. Schwarz, J.M., et al., *MutationTaster2: mutation prediction for the deep-sequencing age.* Nat Methods, 2014. **11**(4): p. 361-2.
11. Shihab, H.A., et al., *An integrative approach to predicting the functional effects of non-coding and coding sequence variation.* Bioinformatics, 2015. **31**(10): p. 1536-43.
12. Adzhubei, I.A., et al., *A method and server for predicting damaging missense mutations.* Nat Methods, 2010. **7**(4): p. 248-9.
13. Dong, C., et al., *Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies.* Hum Mol Genet, 2015. **24**(8): p. 2125-37.
14. Schriemer, D., et al., *Regulators of gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes.* Dev Biol, 2016. **416**(1): p. 255-265.

15. Memic, F., et al., *Transcription and Signaling Regulators in Developing Neuronal Subtypes of Mouse and Human Enteric Nervous System.* Gastroenterology, 2018. **154**(3): p. 624-636.

Supplementary Table 1. Patients included in the study and their clinical features and complications

Patient	Age at surgery (days)	Length HSCR	Additional features	Complications post-surgery
1	88	15cm; rectosigmoid	None	No complications
2	132	13cm; rectosigmoid	None	Hirschsprung Associated Enterocolitis
3	194	15cm; rectosigmoid	None	Obstructed defecation requiring rectal flushings
4	105	15cm; rectosigmoid	None	Hirschsprung Associated Enterocolitis
5	114	10cm; rectosigmoid	None	No complications

Supplementary Table 2. Rare germline variants in HSCR disease genes

Patient	Gene	cDNA	Type	dbSNP	gnomAD exome	gnomAD genome	Cadd	Class
1	*TBATA*	c.T666C	Synonymous	rs2254433	0.26	0.30	-	LB
1	*VCL*	c.G2388A	Synonymous	rs767809	0.48	0.43	-	LB
1	*ZEB2*	c.G3480A	Synonymous	-	.	.		LB
2	*GFRA1**	c.A1T	Start site loss	-	.	.	23	VUS
2	*ZEB2*	c.G3480A	Synonymous	-	.	.	-	LB
2	*DENND3*	c.C1110T	Synonymous	rs2289001	0.34	0.27	-	LB
4	*NRG3**	c.C59G	Missense	rs1884282	0.12	0.13	3.3	LB
4	*NRG3*	c.A1770G	Synonymous	rs17101196	0.08	0.08	-	LB
4	*NRG3*	c.C1986T	Synonymous	rs2295933	0.38	0.38	-	LB
4	*EDNRB**	c.G1392A	Missense	-			25.7	VUS
4	*ZEB2*	c.G3480A	Synonymous	-	.	.	-	LB
4	*DENND3*	c.G3090A	Synonymous	rs1045303	0.34	0.28	-	LB
5	*NUP98*	c.G2688A	Synonymous	rs35803045	0.05	0.05	3.8	LB

Supplementary Table 2. Variants in known HSCR disease genes. All variants are heterozygous and detected in both cell types. Variants marked with * are tested and confirmed with Sanger sequencing. Abbreviations: LB; Likely Benign, VUS; Variant of uncertain significance, Combined Annotation Dependent Depletion (CADD) phred based value can be used to evaluate the likelihood a variant is deleterious (http://cadd.gs.washington.edu/), Population frequency depicted is the using GnomAD genome sequence value (http://gnomad.broadinstitute.org/).

Supplementary Table 3. Rare CNVs are found in exome sequence.

Sample	Chromosome Region	Event	Length	Cytoband	Probes	Gene Symbols	Min Size	Min Region	Max Size	Max Region	Notes	Classification
1	chr1:185,109,784-185,132,629	CN Gain	22846	q25.3	41	TRMT1L, SWT1	22050	chr1:185,110,367-185,132,416	23641	chr1:185,109,202-185,132,842	rare CNV	VUS
1	chr1:41,346,254-41,380,652	CN Gain	34399	p34.2	62	.	34116	chr1:41,346,428-41,380,543	34683	chr1:41,346,080-41,380,762	rare CNV	Likely Benign
1	chr11:54,820,982-54,849,549	CN Gain	28568	q11	25	.	28070	chr11:54,821,340-54,849,409	29066	chr11:54,820,624-54,849,689	rare CNV	Likely Benign
1	chr5:98,757,280-98,811,517	CN Loss	54238	q21.1	45	.	47863	chr5:98,761,171-98,809,033	60614	chr5:98,753,389-98,814,002	rare CNV	Likely Benign
2	chr14:69,924,025-69,983,696	CN Gain	59672	q24.1	84	SLC39A9, BC062762, PLEKHD1	59055	chr14:69,924,349-69,983,403	60289	chr14:69,923,702-69,983,990	rare CNV	VUS
2	chr17:30,106,771-30,107,696	CN Loss	926	q11.2	5	.	554	chr17:30,106,956-30,107,509	1298	chr17:30,106,587-30,107,884	rare CNV	Likely Benign
2	chr8:43,727,152-43,831,881	CN Gain	104730	p11.1	23	.	10306	chr8:43,727,412-43,737,717	3212050	chr8:43,726,893-46,938,942	rare CNV	Likely Benign
2	chrX:5,159,410-5,163,874	Homozygous Copy Loss	4465	p22.32	5	.	1871	chrX:5,159,684-5,161,554	7059	chrX:5,159,137-5,166,195	rare CNV	Likely Benign
3	chr2:10,664,398-10,914,786	CN Gain	250389	p25.1	70	NOL10, RN7SL832P, ATP6V1C2	246862	chr2:10,667,749-10,914,610	253917	chr2:10,661,047-10,914,963	rare CNV	VUS
4	chr1:37,629,746-37,692,299	CN Loss	62554	p34.3	7	.	48357	chr1:37,637,422-37,685,778	76751	chr1:37,622,070-37,698,820	rare CNV	Likely Benign
5	chr6:22,008,230-22,093,109	CN Loss	84880	p22.3	32	CASC15, LINC00340	73287	chr6:22,012,566-22,085,852	96473	chr6:22,003,895-22,100,367	rare CNV	VUS
5	chrX:33,886,710-34,031,196	Homozygous Copy Loss	144487	p21.1	15	.	127713	chrX:33,895,605-34,023,317	161261	chrX:33,877,815-34,039,075	rare CNV	Likely Benign

Supplementary Table 4. Exome sequencing quality information

Patient	Cell type	reads	Mapped reads	Mapped reads (%)	Mean coverage	20X coverage (n bases)	Overlapping 20X coverage	Variants (n)[#]	Overlapping variants (n)[#]	Concordance rate 20 X (%)
1	blood	43692456	42653058	97.6	-	1009526	891280	77521	65576	98.4
1	ENCC	-	-	-	-	936993		87810		-
2	blood	46663496	45464802	97.4	-	1052613	1005681	82750	73769	99.0
2	ENCC	51612528	50248904	97.4	-	1056654		85342		-
3	blood	73084700	72360584	99.0	-	2951928	2765133	92565	64458	98.8
3	ENCC	71473344	70723194	99.0	-	2887416		95400		-
4	fibroblasts	71515650	70770498	99.0	-	2896756	1988647	95491	65780	98.6
4	ENCC	68160814	58726936	86.2	-	2068315		109987		-
5	blood	72927952	72048920	98.8	-	2917188	2661452	96952	84900	99.6
5	ENCC	69489918	68471026	98.5	-	2768779		97171		-

Supplementary Table 4. Depicted are the number of reads sequenced and mapped to the hg19 reference genome. Coverage statistics and ENCC-germline concordance rate at 20X coverage represent the starting amount of bases at which somatic mosaicism was evaluated. [#]variants counted with a read depth of five or higher at 20X coverage.


3

Chapter 3

The development of Goldberg-Shprintzen syndrome is determined by the absence, or reduced expression levels, of *KIF1BP*

Katherine C. MacKenzie[1], Bianca M. de Graaf[1], Andreas Syrimis[2], Yuying Zhao[1], Erwin Brosens[1], Grazia M. S. Mancini[1], Arve Vøllo[3], Frances Flinter[4], Andrew Green[5], Sahar Mansour[6], Jacek Pilch[7], Zornitza Stark[8], Eleni Zamba-Papanicolaou[9], Violetta Christophidou-Anastasiadou[2], Robert M. W. Hofstra[1,10], Jan D.H. Jongbloed[11], Nayia Nicolaou[2], George A. Tanteles[2], Alice S. Brooks[1], Maria M. Alves[1]

[1] Department of Clinical Genetics, Erasmus University Medical Centre, Rotterdam, Netherlands
[2] Department of Clinical Genetics, The Cyprus Institute of Neurology & Genetics and Archbishop Makarios III Medical Centre, Nicosia, Cyprus
[3] Department of Paediatrics, Sykehuset Østfold HF, Fredrikstad, Norway
[4] South East Thames Regional Genetics Centre, Guy's and St Thomas' NHS Foundation Trust, Guy's Hospital, London, UK
[5] National Centre for Medical Genetics, Our Lady's Hospital, Dublin, Ireland
[6] South West Thames Regional Genetic Service, St George's Hospital Medical School, London, UK
[7] Department of Child Neurology, Medical University of Silesia, Katowice, Poland
[8] Genetic Health Services Victoria, Murdoch Children's Research Institute, Melbourne, Australia
[9] Neurology Clinic D, The Cyprus Institute of Neurology & Genetics, Nicosia, Cyprus
[10] Stem Cells and Regenerative Medicine, UCL Great Ormond Street Institute of Child Health, London, UK
[11] Department of Genetics, University of Groningen Medical Centre, Groningen, Netherlands

Corresponding author: Maria M. Alves, Department of Clinical Genetics, Erasmus University Medical Centre, PO BOX 2040, 3000CA Rotterdam, The Netherlands. Tel: +31-10-7030683. E-mail: m.alves@erasmusmc.nl

Submitted

Abstract

Goldberg-Shprintzen syndrome (GOSHS) is a rare recessive disorder with a wide phenotypic range known to be caused by loss of function variants in the kinesin 1 binding protein gene (*KIF1BP*). To date, 37 GOSHS patients have been reported in literature. Here, we document nine new patients with variants in *KIF1BP*: seven with nonsense variants, and two with missense variants. To our knowledge, this is the first time that missense variants have been reported in GOSHS patients. We functionally investigated the effect of these missense variants, in an attempt to find a genotype-phenotype correlation. We also determined whether known HSCR associated single nucleotide polymorphisms (SNPs), could explain the presence of HSCR in GOSHS. Our results showed that the missense variants identified led to reduced expression levels of KIF1BP, and are thus, pathogenic. However, there seems to be no correlation between the severity of GOSHS and the location of the *KIF1BP* variants. We were also unable to find a correlation between known HSCR associated SNPs and the variable expression of HSCR in GOSHS. In conclusion, we show that reduced KIF1BP expression can lead to GOSHS. However, our results suggest that a threshold expression of KIF1BP may determine the classification of the syndrome.

Introduction

Goldberg-Shprintzen syndrome (GOSHS) (OMIM# 609460) is a rare and severe autosomal recessive disorder, characterised by moderate intellectual disability, dysmorphic facial features, microcephaly and axonal neuropathy. GOSHS was first described by Goldberg and Shprintzen in 1981, and to date, 37 cases have been reported through clinical diagnosis, with variable severity and additional features[1-20]. Homozygosity mapping followed by Sanger sequencing, identified homozygous or compound heterozygous loss of function (LOF) variants in the kinesin 1 binding protein gene (*KIF1BP*, previously known as *KIAA1279*), as causative for GOSHS[13,18,19]. Twenty-five out of the 37 reported cases have been shown to have truncating variants in this gene[1,3,11,13-20]. KIF1BP is expressed throughout the developing embryo at early stages of development, and becomes overexpressed in the central and peripheral nervous systems at later developmental stages[21]. KIF1BP is 621 amino acids long and contains two tetratricopeptide repeats. It is involved in axonal structure and outgrowth, microtubule dynamics and cargo trafficking, functioning by binding with various microtubule associated proteins, such as kinesins and the superior cervical ganglia 10 (SCG10)[21-24]. However, the location of the binding domains of KIF1BP remains unmapped, and its precise function is unknown. In mice inactivation of *Kif1bp* is lethal, leading to central and peripheral nervous system defects and delayed enteric nervous system development[25]. Similarly, knocking out the *KIF1BP* ortholog in zebrafish led to disruption of axonal structure and outgrowth, including axonal defects in the enteric nervous system[23].

Hirschsprung disease (HSCR) is reported in ~70% of GOSHS patients, but is considered to be a variable feature. HSCR is characterised by the

absence of enteric ganglia in the distal colon, and occurs in multiple defined syndromes[26]. The link between GOSHS and HSCR is poorly understood, especially considering the variability of its presence or absence, even between family members sharing the same pathogenic variant. It is suspected that a balance of protective and/or predisposing factors in the (epi)genome influence HSCR development[27-30]. Common variants in the Rearranged during transfection gene (*RET*), the Semaphorin 3A gene (*SEMA3A*) and the Neuregulin 1 gene (*NRG1*) were previously found to be associated with HSCR risk. However, no pattern has been determined in GOSHS patients[31-33]. With the addition of new patients we confirm the findings of de Pontual et al.[32], that the occurrence of HSCR in GOSHS cannot be explained by associations with common modifier alleles.

Here, we give an update of all *KIF1BP* reported cases and add nine unpublished cases with six new *KIF1BP* variants, three of which are missense. This is the first time that missense variants have been reported to play a role in GOSHS. Whether these missense variants also result in LOF is unknown. Therefore, we functionally tested the effect of the missense variants identified, on KIF1BP expression levels and cellular localization. In this manuscript we describe our findings.

Materials & methods

Patient inclusion

In this study, nine patients were included (Table 1). These patients were seen routinely in hospitals in the UK, Ireland, Norway, Poland, Australia, Netherlands and Cyprus. Seven of these patients were screened diagnostically for *KIF1BP* variants in the University Medical Centre Groningen, Groningen, NL, to confirm diagnosis of GOSHS. The two patients carrying missense variants were screened at the department of Clinical Genetics in the Erasmus Medical Centre, Rotterdam, NL, and the department of Clinical Genetics at the Cyprus Institute of Neurology & Genetics, in Nicosia, Cyprus. Permission to use diagnostic findings for publication was obtained from all parents.

Sequencing

Sanger sequencing of *KIF1BP* was performed for all nine patients as previously described[13]. A list of primers is available on request. All patients were also Sanger sequenced for the presence of common HSCR associated polymorphisms in *RET*[27], *NRG1* and *SEMA3A*[28]. Primers used are listed in Supplementary Table 1. Exome sequencing was performed on patient NL1 and their parents as previously described[34].

Expression vectors

The pcDNA-HA-hKIF1BP vector was described before[21]. The three missense variants identified were generated by site-directed mutagenesis on pCDNA-HA-hKIF1BP, according to the manufacturer's instructions (QuickChange II site-directed mutagenesis kit, Agilent technologies). Sanger sequencing confirmed the presence of the variants in *KIF1BP*. No extra variants were inserted. Primers used are listed in Supplementary Table 1.

Table 1. Overview of published and unpublished patients with *KIF1B* variants and their clinical features.

Code	US1	US2	Case1	Case2	BP1	BP2	V-4	V-6	V-9
Sex	M	F	M	M	F	F	M	F	M
KIF1BP Mutation	c.1551-1552insA	c.1551-1552insA	c.250G>T	c.250G>T	c.250G>T	c.250G>T	c.268C>T	c.268C>T	c.268C>T
Protein	p.Gln518Thr	p.Gln518Thr	p.Glu84X	p.Glu84X	p.Glu84X	p.Glu84X	p.Arg90X	p.Arg90X	p.Arg90X
HSCR	+	+	+	+	+	-	+	+	-
Facial dysmorphism	+	+	+	+	+	+	+	+	+
Microcephaly	+	+	+	+	+	+	-	+	+
Brain malformation	+	+	+	+	+	+	+	+	+
Developmental delay	+	+	+	+	+	+	+	+	+
Seizures	+	-	-	-	-	-	-	-	-
Neuropathy	?	?	?	?	?	?	?	?	?
Short stature	+	+	+	+	+	+	?	+	+
Hypotonia	+	+	?	?	?	?	?	?	+
Eye anomalies	+	+	+	+	+	+	?	+	+
Cardiac anomalies	-	-	-	-	-	-	?	-	-
Skeletal anomalies	+	+	+	-	-	-	+	+	-
OFC centile	<2nd centile	<2nd centile	<3rd centile	<3rd centile	3rd centile	?	3rd centile	3rd centile	<3rd centile
Ref	1	1	3	3	3	14	11	11	11

Table 1. Overview of published and unpublished patients with *KIF1B* variants and their clinical features.

Code	VI-1	VI-3	CYP1	CYP2	UK1	UK2	AU1	IV-2	IV-1
Sex	F	F	M	F	M	M	F	M	M
KIF1BP Mutation	c.268C>T	c.268C>T	c.718G>T	c.718G>T	c.1117-1118insA	c.1117-1118insA	c.1397dupA	c.268C>T	c.599C>A
Protein	p.Arg90X	p.Arg90X	p.Glu240X	p.Glu240X	p.Ala373AsnfX17	p.Ala373AsnfX17	p.Tyr466X	p.Arg90X	p.Ser200X
HSCR	+	+	+	-	+	+	+	+	+
Facial dysmorphism	+	+	+	+	+	+	+	+	+
Microcephaly	+	+	+		+	+	+	+	+
Brain malformation	+	+	+	+	+	+	+	+	+
Developmental delay	+	+	?	?	+	+	+	+	+
Seizures	-	-	-	-	+	-	?	?	?
Neuropathy	?	?	?	?	?	?	?	?	?
Short stature	+	?	?	?	?	?	+	?	?
Hypotonia	+	?	+	+	+	+	?	?	?
Eye anomalies	?	?	?	?	-	-	-	+	+
Cardiac anomalies	-	?	?	?	-	-	-	-	+
Skeletal anomalies	+	?	-	+	+	+	-	-	+
OFC centile	3rd centile	?	<3rd centile	?	<3rd centile	<3rd centile	?	?	?
Ref	11	13	14	14	15	15	16	17	17

Table 1. Overview of published and unpublished patients with *KIF1B* variants and their clinical features.

Code	IV-3	IV-1.2	V-2	IV.5	IV.8	UK7	IRN1	IE1	UK3
Sex	M	F	F	F	F	F	M	F	M
KIF1BP Mutation	c.599C>A	c.604-605delAG	c.604-605delAG	Deletion exon 2 & 3	Deletion exon 2 & 3	Deletion exon 5 & 6	c.976C>T	c.1694_1695 delAG	Deletion exon 6
Protein	p.Ser200X	p.Arg202 IlefsX2	p.Arg202 IlefsX2	p.Asn143 fsX1	p.Asn143 fsX1	-	p.Gln326X	p.Glu565 AsnfX15	-
HSCR	+	-	+	-	?	+	+	+	+
Facial dysmorphism	+	+	+	+	-	+	+	+	+
Microcephaly	+	+	+	+	+	+	+	+	+
Brain malformation	+	+	+	+	+	+	+	+	+
Developmental delay	+	+	+	+	+	+	+	+	+
Seizures	?	?	?	?	?	?	+	?	-
Neuropathy	?	?	?	?	?	?	?	?	+
Short stature	?	?	?	?	?	+	+	?	+
Hypotonia	?	?	?	?	?	+	+	+	?
Eye anomalies	+	-	-	?	?	+	+	+	+
Cardiac anomalies	-	-	+	?	?	+	-	+	-
Skeletal anomalies	+	-	-	?	?	+	-	-	+
OFC centile	?	?	?	3rd centile	3rd centile	<0.4th centile	normal	?	<0.4th centile
Ref	17	17	17	18	18	19	20	This paper	This paper

Table 1. Overview of published and unpublished patients with *KIF1B* variants and their clinical features.

Code	UK4	PL1	PL2	NO1	NO2	CYP3	NL1
Sex	F	M	M	F	M	F	M
KIF1BP Mutation	Deletion exon 6	c.1516_1517 insA	c.1516_1517 insA	Deletion exon 5&6	Deletion exon 5&6	c. 565C>T	c.68A>G; c.1279A>G
Protein	-	p.Ile506Asnfs X3	p.Ile506Asnfs X3	-	-	p.Pro189Ser	p.Glu23Gly; p.Ser427Gly
HSCR	-	+	-	+	-	+	-
Facial dysmorphism	+	+	+	+	+	+	-
Microcephaly	+	+	+	+	+	+	+
Brain malformation	-	-	+	+	+	+	+
Developmental delay	+	+	+	+	+	+	+
Seizures	-	-	-	-	-	-	-
Neuropathy	?	-	-	-	-	?	?
Short stature	+	+	+	+	+	+	+
Hypotonia	?	-	-	+	+		?
Eye anomalies	-	+	+	+	+	+	?
Cardiac anomalies	-	-	-	-	+	-	?
Skeletal anomalies	+	+	+	-	+	+	?
OFC centile	<0.4th centile	<3rd centile	<3rd centile	?	?	?	?
Ref	This paper	This paper	This paper	This paper	This paper	This paper	This paper

Cell culture and transfection

Human embryonic kidney cells (HEK293) were cultured in DMEM high medium containing 4.5g/l glucose, L-glutamine and pyruvate (Gibco), supplemented with 10% fetal calf serum (Gibco) and 1% penicillin/streptomycin (Gibco). Cells were incubated at 37°C and 5% CO_2. For transient transfection, $3x10^5$ cells were seeded per well in 6-well plates. After 24 hours, transfection was performed using GeneJuice® transfection reagent (Millipore), according to the manufacturer's instructions.

RNA isolation and Q-PCR

RNA was isolated from HEK293 cells transfected with KIF1BP wild-type (WT) and mutant constructs, using the RNAeasy kit (Qiagen), according to the manufacturer's instructions. cDNA preparation and Q-PCR were performed as previously described[35]. A list of primers used can be found in Supplementary Table 1.

Immunofluorescence and confocal microscopy

Following KIF1BP overexpression in HEK293, cells were fixed with 4% paraformaldehyde for 15 minutes, and made permeable with 1% BSA and 0.1% Triton X-100 in PBS. Cells were stained for HA using the HA-Tag antibody (C29F4, Cell Signaling Technology, USA) at 1:1500 dilution, and the Cy3 AffiniPure donkey anti-rabbit IgG, 1:200 dilution (Jackson Immunoresearch, UK). Cells were imaged on a Leica SP5 confocal microscope.

Cell lysates and western blot analysis

Twenty-four to 48 hours after transfection, cells were washed with PBS and lysed as described before[21]. Protein quantification was determined using the Pierce BCA Protein Assay kit (Thermo Fisher Scientific), and

40ug of cell lysates were stored in loading buffer at -80°C before they were processed further. SDS–PAGE followed by western blot analysis was performed using an in-house anti-HA antibody, and a GAPDH antibody (Millipore), both at 1:5000 dilution. Secondary antibodies used were the IRDye 680RD Goat anti-rabbit and the IRDye 800CW Goat anti-mouse (Li-Cor), at 1:10.000 dilution.

Results

Three novel truncating variants in *KIF1BP* were identified in five GOSHS patients

Seven previously unreported GOSHS patients, three sibling pairs and an isolated patient, were sequenced for *KIF1BP* (Table 1). Large exon deletions, as well as new frameshift variants, were identified in these patients (Figure 1, Table 1). Patient IE1 is a female with a classical GOSHS phenotype and HSCR. She was found to have a two base-pair deletion at c.1694_1695, causing a premature stop in exon 7. Siblings PL1 and PL2, are both male with classical GOSHS facial dysmorphism. PL1 has HSCR, where PL2 does not. An insertion of an A at position c.1516_1517 of *KIF1BP* was identified in these patients, leading to the appearance of a premature stop in exon 7. Siblings UK3 and UK4 were found to have a deletion of exon 6 of *KIF1BP*. Patient UK3 is a male, with classical GOSHS phenotype and HSCR. Patient UK4 is a female and does not have HSCR. Siblings NO1 and NO2 have a deletion of exons 5 and 6 of *KIF1BP*. Both have a classical GOSHS phenotype, but NO1, female, has HSCR where it is absent in her brother, NO2. This deletion has been previously reported[19].

Three novel missense *KIF1BP* variants were identified in two patients

The first patient (CYP3), is a 28-year-old female of Cypriot ancestry, born to reportedly non-consanguineous parents. She had a history of microcephaly, mild intellectual disability and developmental delay. She presented with short stature, typical dysmorphic facial features with bilateral blepharoptosis, and corneal ulcers. She was diagnosed with HSCR at the age of 3 years. Additionally, she had scoliosis, lordosis, pes cavus as well as mild sensory motor neuropathy with both axonal and demyelinating features. Brain MRI did not reveal a CNS abnormality. No

copy number variations (CNVs) were detected with array-CGH. Review of the family history revealed that the patient's younger brother was diagnosed with HSCR and died in the neonatal period from sepsis, following surgery for meconium ileus. There was also a report of a maternal relative who apparently died in infancy and had features suggestive of GOSHS. Sequencing of *KIF1BP* showed a homozygous missense variant in exon 3, c.565C>T (p.Pro189Ser; Mut3; Figure 1; Table 1). Both parents of the CYP3 patient were found to be heterozygous. This variant has not been previously reported. Prediction tools, Polyphen-2 and SIFT, predict this variant as benign. However, the CADD score of this variant is 17.95, indicating a potentially pathogenic effect.

The second patient (NL1) is a 10-year-old male of Moroccan ancestry, born to consanguineous parents. He had a history of microcephaly and presented with short stature. Brain imaging showed pachygyria and he was affected by demyelinating peripheral neuropathy and perceptive deafness. However, in contrast with CYP3, he lacked the typical facial features of GOSHS, had no skeletal symptoms and had no reported gastrointestinal or enteric nervous system abnormalities. Whole exome sequencing was conducted on DNA from blood for diagnostic purposes, and two heterozygous missense variants in *KIF1BP* were identified, one in exon 1, c.68A>G (p.Glu23Gly; Mut1), and one in exon 7, c.1279A>G (p.Ser427Gly; Mut2; Figure 1; Table 1). No other likely pathogenic variants were identified in this patient. None of these variants were previously reported, and both of them are predicted to have a pathogenic effect based on prediction tools such as, Polyphen-2 and SIFT.

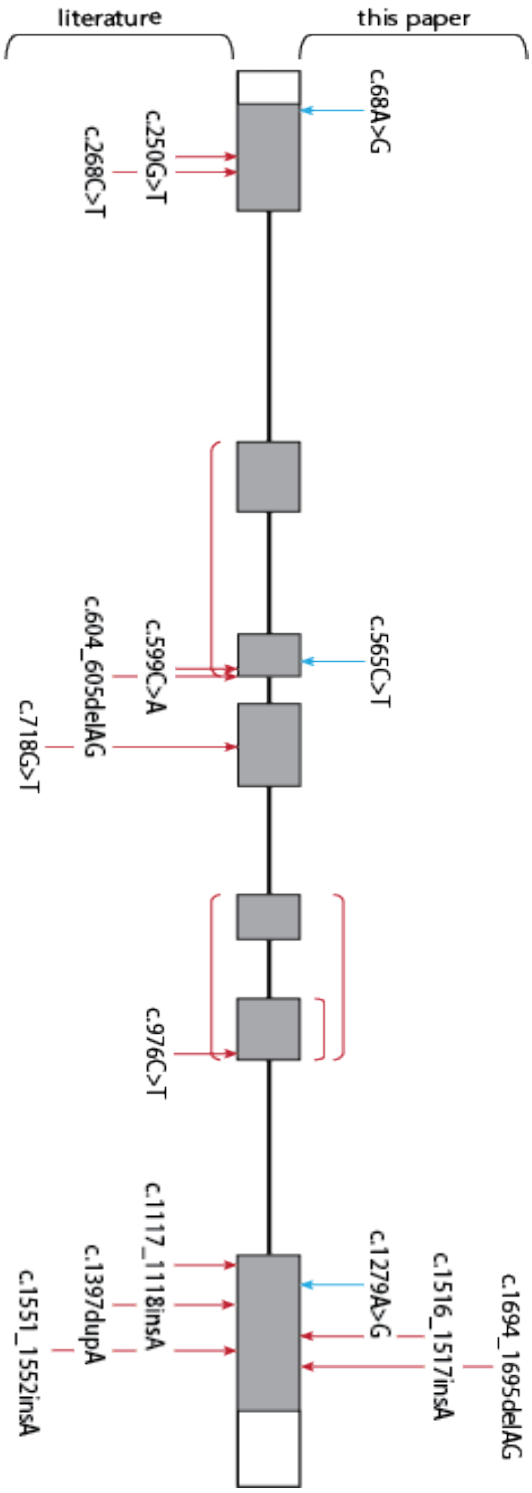


Figure 1. Schematic representation of *KIF1BP* highlighting positions of all reported variants. Missense variants are indicated with blue, and nonsense variants with red. Square brackets show exon deletions.

KIF1BP expression levels were reduced by the missense variants

To evaluate the effect of the missense variants, expression levels of KIF1BP were determine after transfection of the HEK293 cells with constructs expressing the WT, and mutant KIF1BP cDNA. Q-PCR showed decreased, RNA levels for all three missense variants (Figure 2A). Western blot analysis showed decreased expression level of KIF1BP for all three mutants (50% decrease), when compared to WT (Figure 2B, 2C). The effects of Mut3, of patient CYP3, seem to be the most pronounced.

Cellular localisation of KIF1BP is unaffected by the missense variants

Tagged WT and mutant KIF1BP constructs were overexpressed in HEK293 cells, to determine any effect of the variants in the organisation or localisation of KIF1BP within the cell. The WT protein is seen to have high cytoplasmic expression, as previously described[21]. For the mutant proteins, while the intensity of KIF1BP expression seems decreased, no effect on KIF1BP localisation was observed (Figure 3).

Common SNPs in *RET*, *NRG1* and *SEMA3A* do not affect HSCR development in GOSHS

It is known that phenotypic variability exists in GOSHS patients, and that HSCR is a variable feature, even within families with the same *KIF1BP* truncating variant (Table 1). Here, we investigated whether the presence of common SNPs associated to HSCR, would be the determinant factor for the presence of this disorder in GOSHS. The SNPs we decided to investigate are located in intron 1 of *RET*, *SEMA3A* and *NRG1*. Although, all these SNPS have been described to increase the risk for HSCR, we were unable to find a polygenic risk effect for their presence and the occurrence of HSCR in our subset of GOSHS patients and unaffected family members (p=0.526, Supplementary Table 2).

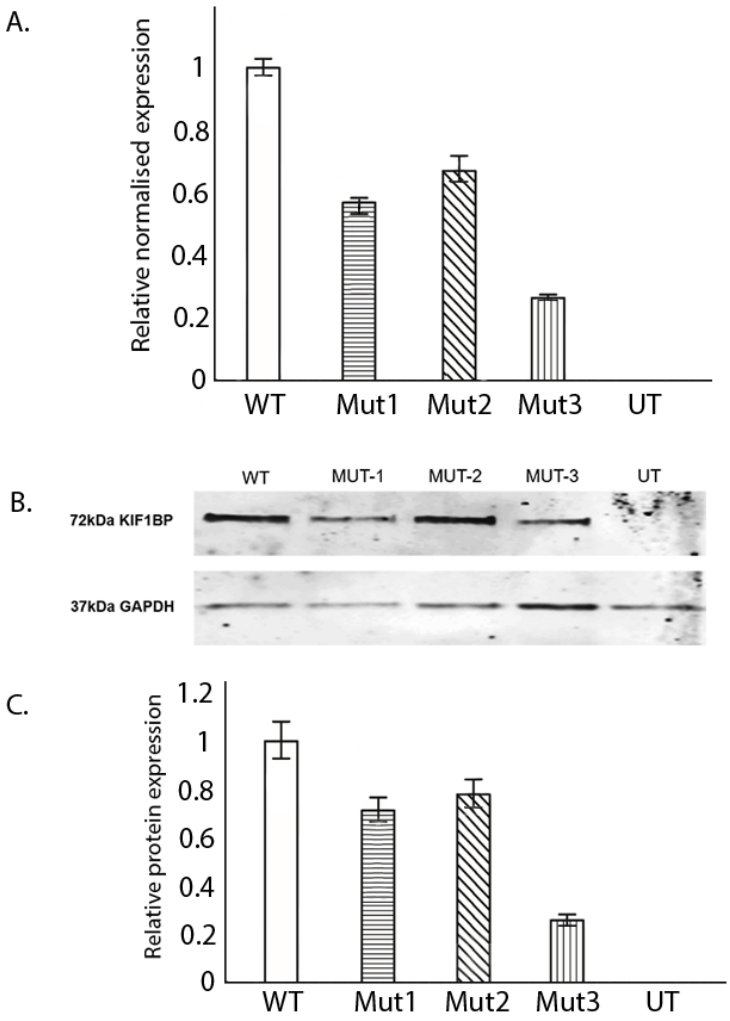


Figure 2. Expression of KIF1BP is altered in the presence of the missense variants. A) Q-PCR results showing relative normalised expression of *KIF1BP* following transfection with wild type (WT) or mutant constructs. All mutant constructs show a decrease in KIF1BP expression compared to WT levels. B) Western blot of KIF1BP expression following transfection of either WT or mutant constructs. C) Quantification of protein expression after normalisation for GAPDH expression. Mut1, A68G, shows ~70% expression, Mut2, A1279G, shows ~80% expression and Mut3, C565T, shows ~25% expression, when compared to the WT. Error bars show SEM.

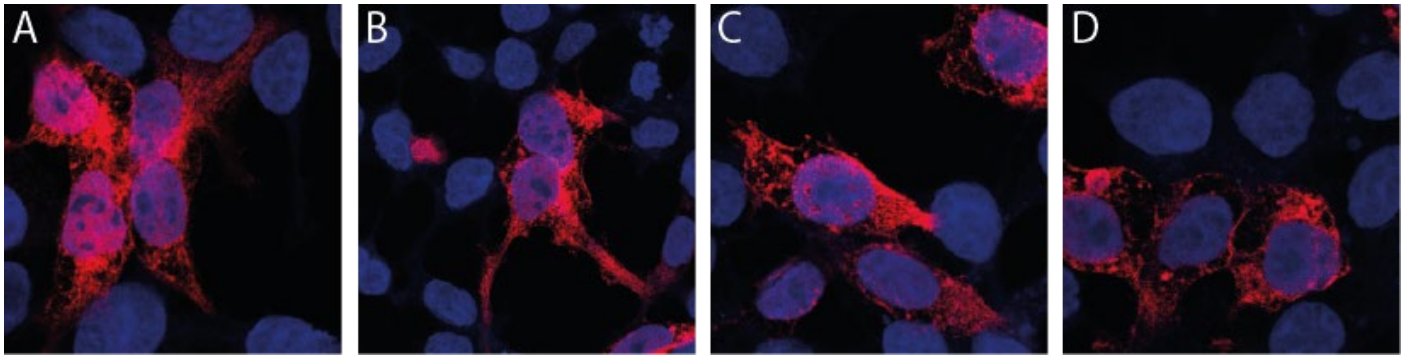


Figure 3. Confocal images of KIF1BP localisation after transfection of HEK293 cells with wild-type (WT) and mutant constructs, show no difference between WT and missense variants. A) WT. B) Mut1 - A68G. C) Mut2 - A1279G. D) Mut3 - C565T.

Discussion

In this manuscript, we report nine patients with variants in *KIF1BP*. A common feature to all these patients, is the presence of intellectual disability and developmental delays. However, the phenotypic spectrum is wide, with distinct facial morphology, microcephaly and other central nervous system malformations. Interestingly, this wide phenotypic range is even found in siblings carrying the same variant. As can be seen in table 1, the incidence of HSCR in GOSHS is approximately 70% (24/34), and of the ten patients without HSCR, seven have a family member with the same variant, that does have HSCR. This suggests the presence of modifying factors, or absence of protective factors, in these patients than can tilt the balance in favour of HSCR. Here, we hypothesized that common HSCR modifier variants in *RET*[27], *NRG1* and *SEMA3A*[28] may work as these modifying factors, as it has been shown for other syndromes[27,32,36]. However, our results did not show any correlation between the incidence of HSCR and presence of these common polymorphisms (P=0.526, Supplementary Table 2). As HSCR is a complex genetic disease, multiple factors are known to play a role in its development in addition to genetic risk factors, such as epigenetic changes[37], somatic variants[38], protective pathways[39], threshold numbers of cells[40] or stochastic chance[41]. It is therefore, possible that such factors

have a determinant role on the occurrence of HSCR in GOSHS, but further research is required to investigate this hypothesis.

Loss of function (LOF) variants in *KIF1BP* are known to cause GOSHS. However, there seems to be no correlation between the location of the variant, and the severity of syndromic characteristics (Figure 1), as they all result in a total loss of protein[13,17]. All patients reported here carry pathogenic variants in *KIF1BP*. In two of these patients, missense variants were identified instead of LOF and one of these cannot be considered to have GOSHS. This finding was quite interesting, as no missense variants have been previously reported in GOSHS and no other syndromes are known to be associated with *KIF1BP* variants. The characterisation and diagnosis of these two patients differs tremendously. While patient CYP3 has the hallmark features of GOSHS, including HSCR (Table 1), patient NL1 does not show any of the clinical defined features of this syndrome. In fact, NL1 cannot be classified as a GOSHS patient, despite the identification of compound heterozygous variants in *KIF1BP*. Functional studies performed to investigate the effect of these missense variants, showed that they all lead to reduced expression of both KIF1BP RNA and protein levels. However, this reduction is milder in patient NL1 in comparison with patient CYP3 (Figure 2A, 2B and 2C), which could explain the classical GOSHS

phenotype in CYP3. Based on this result, we speculate that a threshold expression of KIF1BP may be required for regulation of developmental functions, and that a small decrease in its expression due to missense variants is, to some extent, tolerable. Similarly, the variant in patient CYP3 may have not reached the threshold expression levels required, leading to the typical GOSHS phenotype. Therefore, missense variants can be as damaging as a truncating variant, depending on its effect on protein expression levels.

It has been previously noted that the diagnosis of GOSHS should rely on molecular and genetic findings in place of phenotypic recognition only, due to its similarity with other syndromes[20]. Based on genetic findings, patient NL1 would be considered to have GOSHS, however this patient has very few of the hallmark features of GOSHS. Therefore, we suggest the inclusion of *KIF1BP* in genetic screens if any neural crest derivative tissues are affected, such as enteric, craniofacial or cortical malformations. The classification of syndromes based on phenotype may be more useful for the family to appropriately meet the needs of the patient. Therefore, we believe that the accurate classification of a patient based on genotype, as well as phenotype, by a clinical geneticist, is vital to provide accurate information to the family of the patient, as well as for advising clinical treatment.

References

1. Goldberg RB, Shprintzen RJ. Hirschsprung megacolon and cleft palate in two sibs. *J Craniofac Genet Dev Biol.* 1981;1(2):185-189.
2. Brunoni D, Joffe R, Farah LM, Cunha AJ. Syndrome identification case report 92: Hirschsprung megacolon, cleft lip and palate, mental retardation, and minor congenital malformations. *The Journal of clinical dysmorphology.* 1983;1(1):20-22.
3. Hurst JA, Markiewicz M, Kumar D, Brett EM. Unknown syndrome: Hirschsprung's disease, microcephaly, and iris coloboma: a new syndrome of defective neuronal migration. *Journal of Medical Genetics.* 1988;25(7):494-497.
4. Kumasaka K, Clarren Sterling K, Opitz John M, Reynolds James F. Familial patterns of central nervous system dysfunction, growth deficiency, facial clefts and congenital megacolon: A specific disorder? *Am J Med Genet.* 1988;31(2):465-466.
5. Breslau L, Laan L. Facial cleft and congenital megacolon: A specific disorder? *Am J Med Genet.* 1989;34(4):613-613.
6. Halal F, Morel J. The syndrome of hirschsprung disease, microcephaly, unusual face, and mental retardation. *Am J Med Genet.* 1990;37(1):106-108.
7. Yomo A, Taira T, Kondo I. Goldberg-Shprintzen syndrome: Hirschsprung disease, hypotonia, and ptosis in sibs. *Am J Med Genet.* 1991;41(2):188-191.
8. Tanaka H, Ito J, Cho K, Mikawa M. Hirschsprung disease, unusual face, mental retardation, epilepsy, and congenital heart disease: Goldberg-Shprintzen syndrome. *Pediatr Neurol.* 1993;9(6):479-481.
9. Ohnuma K, Imaizumi K, Masuno M, Nakamura M, Kuroki Y. Magnetic resonance imaging abnormalities of the brain in Goldberg-Shprintzen syndrome (Hirschsprung disease, microcephaly, and iris coloboma). *Am J Med Genet.* 1997;73(2):230-232.
10. Fryer AE. Goldberg-Shprintzen syndrome: report of a new family and review of the literature. *Clin Dysmorphol.* 1998;7(2):97-101.
11. Brooks AS, Breuning MH, Osinga J, et al. A consanguineous family with Hirschsprung disease, microcephaly, and mental retardation (Goldberg-Shprintzen syndrome). *Journal of Medical Genetics.* 1999;36(6):485-489.
12. Silengo M, Ferrero Giovanni B, Tornetta L, et al. Pachygyria and cerebellar hypoplasia in Goldberg-Shprintzen syndrome. *American Journal of Medical Genetics Part A.* 2003;118A(4):388-390.
13. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous Nonsense Mutations in *KIAA1279* Are Associated with Malformations of the Central and Enteric Nervous Systems. *The American Journal of Human Genetics.* 2005;77(1):120-126.
14. Brooks A. *Genetics of syndromic and non-syndromic Hirschsprung disease*. Enschede: Clinical Genetics, Erasmus University Rotterdam; 2005.
15. Murphy HR, Carver MJ, Brooks AS, Kenny SE, Ellis IH. Two brothers with Goldberg-Shprintzen syndrome. *Clin Dysmorphol.* 2006;15(3):165-169.

16. Bruno DL, Stark Z, Amor DJ, et al. Extending the scope of diagnostic chromosome analysis: Detection of single gene defects using high-resolution SNP microarrays. *Hum Mutat.* 2011;32(12):1500-1506.
17. Drévillon L, Megarbane A, Demeer B, et al. KBP–cytoskeleton interactions underlie developmental anomalies in Goldberg–Shprintzen syndrome. *Human molecular genetics.* 2013;22(12):2387-2399.
18. Valence S, Poirier K, Lebrun N, et al. Homozygous truncating mutation of the KBP gene, encoding a KIF1B-binding protein, in a familial case of fetal polymicrogyria. *Neurogenetics.* 2013;14(3):215-224.
19. Dafsari Hormos S, Byrne S, Lin J-P, et al. Goldberg–Shprintzen megacolon syndrome with associated sensory motor axonal neuropathy. *American Journal of Medical Genetics Part A.* 2015;167(6):1300-1304.
20. Salehpour S, Hashemi-Gorji F, Soltani Z, Ghafouri-Fard S, Miryounesi M. Association of a Novel Nonsense Mutation in KIAA1279 with Goldberg-Shprintzen Syndrome. *Iranian Journal of Child Neurology.* 2017;11(1):70-74.
21. Alves MM, Burzynski G, Delalande J-M, et al. KBP interacts with SCG10, linking Goldberg–Shprintzen syndrome to microtubule dynamics and neuronal differentiation. *Human molecular genetics.* 2010;19(18):3642-3651.
22. Wozniak MJ, Melzer M, Dorner C, Haring H-U, Lammers R. The novel protein KBP regulates mitochondria localization by interaction with a kinesin-like protein. *BMC Cell Biology.* 2005;6(1):1-15.
23. Lyons DA, Naylor SG, Mercurio S, Dominguez C, Talbot WS. KBP is essential for axonal structure, outgrowth and maintenance in zebrafish, providing insight into the cellular basis of Goldberg-Shprintzen syndrome. *Development.* 2008;135(3):599-608.
24. Kevenaar Josta T, Bianchi S, van Spronsen M, et al. Kinesin-Binding Protein Controls Microtubule Dynamics and Cargo Trafficking by Regulating Kinesin Motor Activity. *Current Biology.* 2016;26(7):849-861.
25. Hirst CS, Stamp LA, Bergner AJ, et al. Kif1bp loss in mice leads to defects in the peripheral and central nervous system and perinatal death. *Scientific Reports.* 2017;7(1):16676.
26. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.
27. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer variants synergistically drive dysfunction of a gene regulatory network in Hirschsprung disease. *Cell.* 2016;167(2):355-368.e310.
28. Kapoor A, Jiang Q, Chatterjee S, et al. Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms. *Human molecular genetics.* 2015;24(10):2997-3003.
29. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature.* 2005;434:857.
30. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental Biology.* 2016;417(2):198-208.
31. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature.* 2005;434(7035):857.

32. De Pontual L, Pelet A, Clement-Ziza M, et al. Epistatic interactions with a common hypomorphic RET allele in syndromic Hirschsprung disease. *Hum Mutat.* 2007;28(8):790-796.
33. de Pontual L, Pelet A, Trochet D, et al. Mutations of the RET gene in isolated and syndromic Hirschsprung's disease in human disclose major and modifier alleles at a single locus. *Journal of medical genetics.* 2006;43(5):419-423.
34. Sribudiani Y, Chauhan RK, Alves MM, et al. Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease. *Gastroenterology.* 2018.
35. McCann CJ, Alves MM, Brosens E, et al. Neuronal Development and Onset of Electrical Activity in the Human Enteric Nervous System. *Gastroenterology.* 2019.
36. Tang CS-m, Gui H, Kapoor A, et al. Trans-ethnic meta-analysis of genome-wide association studies for Hirschsprung disease. *Human molecular genetics.* 2016;25(23):5265-5275.
37. Tang W, Li B, Tang J, et al. Methylation analysis of EDNRB in human colon tissues of Hirschsprung's disease. *Pediatric Surgery International.* 2013;29(7):683-688.
38. D'Gama AM, Walsh CA. Somatic mosaicism and neurodevelopmental disease. *Nature Neuroscience.* 2018:1.
39. Griseri P, Lantieri F, Puppo F, et al. A common variant located in the 3′ UTR of the RET gene is associated with protection from Hirschsprung disease. *Hum Mutat.* 2007;28(2):168-176.
40. Barlow AJ, Wallace AS, Thapar N, Burns AJ. Critical numbers of neural crest cells are required in the pathways from the neural tube to the foregut to ensure complete enteric nervous system formation. *Development.* 2008;135(9):1681-1691.
41. Cheeseman BL, Zhang D, Binder BJ, Newgreen DF, Landman KA. Cell lineage tracing in the developing enteric nervous system: superstars revealed by experiment and simulation. *Journal of The Royal Society Interface.* 2014;11(93).

Supplementary Table 1

Name	Purpose	Primer sequence
rs2506030_1FW	Sanger sequencing	GGAGGCGCATCCCTAGCC
rs2506030_1RV	Sanger sequencing	CAGCTACTCTGGGGCCTTGC
rs7069590_3FW	Sanger sequencing	CAATGGTCCCCTGCACACC
rs7069590_3RV	Sanger sequencing	GCACCCCTGGCAGTGACC
rs2505998 _1FW	Sanger sequencing	TGGCTACCTAGGCTACACACTCAGG
rs2505998 _1RV	Sanger sequencing	CCCCCAGACCTTTTTCCAAGG
rs2435357_1FW	Sanger sequencing	CAGCTGCTGCAGAGTTAATCACC
rs2435357_1RV	Sanger sequencing	AGAGGCACCAGGGTCAAAGC
rs9282834_1FW	Sanger sequencing	GTCCATGCCTTCCCCACTCC
rs9282834_1RV	Sanger sequencing	GGGAAAGTCTGTGTGGAAAACTGC
rs11766001_1FW	Sanger sequencing	CAATCAAAATGCAAGACACCATTAGC
rs11766001_1RV	Sanger sequencing	TGAAAGATGATGGTGTGGATGAGC
rs80227144 _1FW	Sanger sequencing	GGGCAGATGGATATGTAGGC
rs80227144 _1RV	Sanger sequencing	TTGAATAAAATGTCTTATTGTTTTCC
rs7005606 _2FW	Sanger sequencing	TCTGCACCATAATTACAGCAATGG
rs7005606 _2RV	Sanger sequencing	TGGAGGGTACCACTTCTAGTTTTGC
KBP(A68G)F	Site Directed Mutagenesis	TCGCGGGTGGGACTGCATAAAAATCC
KBP(A68G)R	Site Directed Mutagenesis	CAGAGCGAGCGCCGCCTG
KBP(A1279G)F	Site Directed Mutagenesis	CCAAGACCACGGTGCTCTGTT
KBP(A1279G)R	Site Directed Mutagenesis	ACAACTTCAATATGGTCAGTGAC

Supplementary Table 2

Sample	HSCR	rs2506030	rs7069590	rs2505998	rs2435357	rs9282834	rs1176600	rs8022714	rs7005606	rs2506030	rs7069590	rs2505998
97D0101	N	G/G	T/C	A/G	A/G	G/G	A/A	C/C	G/T	2	1	1
04D1273	N	G/G	T/C	A/G	A/G	G/G	A/A	C/C	T/T	2	1	1
96D3576	N	G/G	T/C	A/G	A/G	G/G	A/A	C/C	T/T	2	1	1
96D3571	Y	G/G	T/C	A/G	A/G	G/G	A/A	C/C	G/T	2	1	1
96D3572	Y	G/G	T/C	A/G	A/G	G/G	A/A	C/C	G/T	2	1	1
96D3573	Y	G/G	T/C	A/G	A/G	G/G	A/A	C/C	T/T	2	1	1
96D3574	Y	G/G	T/C	A/G	A/G	G/G	A/A	C/C	G/T	2	1	1
08D2080	N	G/A	C/C	G/G	G/G	G/G	A/C	C/C	G/T	1	0	2
08D1839	N	G/G	T/T	A/G	A/G	G/G	A/A	C/C	G/T	2	2	1
03D3615	Y	G/A	T/C	A/G	A/G	G/G	A/A	C/C	G/G	1	1	1
03D2667	Y	G/A	T/C	G/G	G/G	G/G	A/C	C/C	G/T	1	1	2
09D12521	N	G/A	T/C	G/G	G/G	G/G	A/A	C/C	T/T	1	1	2
23303	Y	G/A	T/C	A/G	A/G	G/G	A/A	C/C	G/T	1	1	1
27012	N	G/A	T/T	A/A	A/A	G/G	A/A	C/C	T/T	1	2	0
27013	N	A/A	T/C	G/G	G/G	G/G	A/C	C/C	G/T	0	1	2
27014	Y	G/A	T/C	A/G	A/G	G/G	A/A	C/C	T/T	1	1	1
30655	N	A/A	T/T	G/G	G/G	G/G	A/A	C/C	G/T	0	2	2
30656	N	G/A	T/T	G/G	G/G	G/G	A/A	C/C	G/T	1	2	2
30658	N	G/A	T/T	G/G	G/G	G/G	A/A	C/C	T/T	1	2	2
30660	Y	A/A	T/T	G/G	G/G	G/G	A/A	C/C	T/T	0	2	2
46677	Y	G/A	T/C	G/G	G/G	G/G	A/A	C/C	T/T	1	1	2
53649	N	G/A	T/T	G/G	G/G	G/G	A/A	C/C	T/T	1	2	2
55529	Y	G/A	T/T	G/G	G/G	G/G	A/A	C/C	G/G	1	2	2
57619	N	A/A	T/T	A/A	A/A	G/G	A/A	C/C	T/T	0	2	0
13D7703	N	A/A	T/C	A/G	A/G	G/G	C/C	C/C	T/T	0	1	1
cyprus?	Y	A/G	T/C		A/G	G/G				1	1	0

Supplementary Table 2 cont.

rs2435357	rs9282834	rs1176600	rs8022714	rs7005606	rs2506030	rs7069590	rs2505998	rs2435357	rs9282834	rs1176600	rs8022714	rs7005606	rs2506030
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
0	0	1	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	1,1755733
1	0	0	0	2	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	1	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
1	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
2	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	1	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0
1	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0
0	0	0	0	1	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0
0	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
0	0	0	0	2	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0,5877867
2	0	0	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0
1	0	2	0	0	0,5877867	0,5306283	1,427916	1,3887912	0,5877867	0,4700036	1,6486586	0,4700036	0
1	0	1	0	1	0,587787	0,530628	1,427916	1,388791	0,587787	0,470004	1,648659	0,470004	0,587787

Supplementary Table 2 cont.

rs7069590	rs2505998	rs2435357	rs9282834	rs1176600	rs8022714	rs7005606	PGRS
0,5306283	1,427916	1,3887912	0	0	0	0,4700036	4,992912
0,5306283	1,427916	1,3887912	0	0	0	0	4,522909
0,5306283	1,427916	1,3887912	0	0	0	0	4,522909
0,5306283	1,427916	1,3887912	0	0	0	0,4700036	4,992912
0,5306283	1,427916	1,3887912	0	0	0	0,4700036	4,992912
0,5306283	1,427916	1,3887912	0	0	0	0	4,522909
0,5306283	1,427916	1,3887912	0	0	0	0,4700036	4,992912
0	2,8558321	0	0	0,4700036	0	0,4700036	4,383626
1,0612565	1,427916	1,3887912	0	0	0	0,4700036	5,523541
0,5306283	1,427916	1,3887912	0	0	0	0,9400073	4,875129
0,5306283	2,8558321	0	0	0,4700036	0	0,4700036	4,914254
0,5306283	2,8558321	0	0	0	0	0	3,974247
0,5306283	1,427916	1,3887912	0	0	0	0,4700036	4,405126
1,0612565	0	2,7775825	0	0	0	0	4,426626
0,5306283	2,8558321	0	0	0,4700036	0	0,4700036	4,326468
0,5306283	1,427916	1,3887912	0	0	0	0	3,935122
1,0612565	2,8558321	0	0	0	0	0,4700036	4,387092
1,0612565	2,8558321	0	0	0	0	0,4700036	4,974879
1,0612565	2,8558321	0	0	0	0	0	4,504875
1,0612565	2,8558321	0	0	0	0	0	3,917089
0,5306283	2,8558321	0	0	0	0	0	3,974247
1,0612565	2,8558321	0	0	0	0	0	4,504875
1,0612565	2,8558321	0	0	0	0	0,9400073	5,444882
1,0612565	0	2,7775825	0	0	0	0	3,838839
0,5306283	1,427916	1,3887912	0	0,9400073	0	0	4,287343
0,5306283	0	1,3887912	0	0,4700036	0	0,4700036	3,447213

Supplementary Table 2 statistics

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
no HSCR	14	63,17114	4,512224	0,179667
HSCR	11	50,9675	4,633409	0,268306

ANOVA

rce of Varia	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between G	0,090464	1	0,090464	0,41458	0,526023	4,279344
Within Gro	5,018733	23	0,218206			
Total	5,109196	24				

no HSCR	HSCR
4,992912	4,992912
4,522909	4,992912
4,522909	4,522909
4,383626	4,992912
5,523541	4,875129
3,974247	4,914254
4,426626	4,405126
4,326468	3,935122
4,387092	3,917089
4,974879	3,974247
4,504875	5,444882
4,504875	
3,838839	
4,287343	


4

Chapter 4

Size matters: enrichment for large deletions encompassing genes expressed in the enteric nervous system in Hirschsprung disease patients with additional associated anomalies

Katherine C. MacKenzie[1], Clara S Tang[3], Jonathan D. Windster[1,2], Thuy Linh Le[4,6], Anwarul Karim[3, 7], Bianca M. de Graaf [1], Robert van der Helm[1], Yolande van Bever[1], Cornelius E.J. Sloots[2], Conny Meeussen[2], Dick Tibboel[2], Annelies de Klein[1], René M. H. Wijnen[2], Jeanne Amiel[4,5,6], Stanislas Lyonnet[4,5,6], Maria-Mercè Garcia-Barcelo[3], Paul K.H. Tam[3], Maria Alves[1], Alice Brooks[1], Robert M.W. Hofstra[1,8], Erwin Brosens[1]

[1] Department of Clinical Genetics, Erasmus University Medical Centre, Rotterdam, Netherlands
[2] Department of Paediatric Surgery, Erasmus University Medical Centre, Rotterdam, Netherlands
[3] Department of Surgery, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China
[4] Paris Descartes – Sorbonne Paris Cité University, Imagine Institute, Paris, France.
[5] Service de Génétique, Necker Enfants Malades University Hospital, APHP, Paris, France
[6] Laboratory of embryology and genetics of congenital malformations, INSERM UMR1163, Paris, France
[7] Chittagong Research Institute for Children Surgery (CRICS), Chittagong, Bangladesh
[8] Stem Cells and Regenerative Medicine, UCL Great Ormond Street Institute of Child Health, London, UK

Manuscript in preparation

Abstract

Hirschsprung disease (HSCR) is characterized by absence of enteric ganglia, primarily in the distal colon. Approximately 18% of patients have additional anatomical malformations or associated neurodevelopmental disorders, including autism and intellectual disability. A subset of these patients have a known genetic syndrome in which HSCR has a variable expression or penetrance. In others, the genetic aetiology is unknown and we hypothesize that rare Copy Number Variation (CNV) impacts their disease development. In this study, we found that rare Copy Number (CN) losses are indeed, significantly enriched in patients with HSCR and additional anomalies without a known causal variant. This is not the case in isolated HSCR, or in HSCR patients with a known *RET*, or other causal variant. Of the HSCR patients with additional anomalies tested, at least five (three males and two females) had a large *de novo* CNV and one male inherited an X-linked CN loss. Patients with a known causal variant, had a significant lower burden of the known HSCR predisposing risk haplotypes, and isolated HSCR patients a higher burden compared to patients with HSCR and additional anomalies without a known causal variant. The rare CN losses identified are enriched for variant intolerant genes, expressed in the developing mouse enteric nervous system: *SLC8A1*, *GNL1*, *GABBR1*, *MAPK8*, *UFD1L*, *AKT3*, *TBX2*, *USP32* and *TUBB*. A loss of function variant in *TUBB* (pLI 0.98) was identified in a HSCR patient without a deleterious CNV, and *AKT3* as well as *UFD1L* are impacted by CN losses in other patients, confirming our hypothesis that rare CNV contributes to HSCR with associated anomalies and an unknown genetic aetiology.

Introduction

The enteric ganglia form the enteric nervous system (ENS), which is derived from the neural crest. Enteric neural crest cells (ENCCs) invade the developing gut tube in the proximal foregut around week four of human development and migrate caudally to colonize the distal hindgut in week 7[1]. In quails and zebrafish, it has been shown that after colonization of the gut there is a subset of cells that then migrate in a caudal-cranial direction[2,3]. Failure in either migration, proliferation, differentiation or survival of these ENCCs[4] is known to cause Hirschsprung disease (HSCR), a congenital enteric neuropathy, characterized by an absence of enteric ganglia in the distal colon. The length of the affected region is variable and can extend to the whole intestine. However, it is more commonly restricted to the recto-sigmoid region. HSCR can segregate through families[5-8] and, depending on the length of the affected region and gender of the proband, siblings can have a recurrence risk of up to 33%[9]. Most patients have HSCR as an isolated anatomical malformation, and in this case, this disease is a textbook example of a complex genetic disorder in which both rare and common variants have an impact[10]. However, approximately 18% of patients have additional anatomical malformations or neurological symptoms[9]. A portion of these patients have a known monogenetic syndrome in which HSCR is a common feature[9]. The underlying cause in other patients with associated anomalies is unknown. Shared phenotypical characteristics in this subset of patients could be the result of a shared genetic defect – either *de novo* or inherited, such as a recessive aberration and/or variation from unaffected parents. The

unknown combination of disease entities could be the result of (a) point mutation(s), but is more likely the result of (large) structural rearrangements, duplications or deletions. In the past, chromosomal deletions have proven instrumental in the identification of disease-causing genes. However, finding such genes is not straightforward. The difficulty in searching for Copy Number Variation (CNV) is comparable to normal sequencing, namely how to distinguish between disease contributing variations and those that have no functional consequences. Another challenge is distinguishing the gene(s) within a CNV that contribute to the disease. Comparable to Single Nucleotide Polymorphisms (SNPs) and Single Nucleotide Variants (SNVs), segmental change in the amount of DNA can be common in the population, and are termed Copy Number Polymorphisms (CNPs) or CNVs. Paradoxically, having a CNV with a very low population frequency is not uncommon, even in seemingly unaffected individuals. However, large CNVs, especially those bigger than 250kb, are more frequent in individuals with congenital and developmental anomalies than in healthy controls[11]. A CNV can be either *de novo* or inherited and result in diseases, syndromes or traits, often due to changes in gene dosage[12]. Although larger CNVs are more prevalent in patients, unaffected individuals can also have relatively large CNVs affecting many genes without any clinical consequence[11]. Whilst having a rare CNV is common, having a specific CNV is not. We know from large case-control studies that there are hotspots in the human genome that, due to their genomic architecture, are more prone to *de novo* rearrangements. These loci harbour genes responsible for the phenotypes seen in recurrent microdeletion and duplication

syndromes[11-13]. These CNVs can occur *de novo* or be inherited, have a variable phenotypic penetrance and range from very rare to very common in the population depending on the region in question[11,13,14]. When patients suffer from multiple disease entities, a CNV affecting the expression of several dosage-sensitive genes which are important for the development of the different affected organ systems, could be suspected. This CNV could result in a contiguous gene syndrome in which multiple genes have a dosage effect that, in combination or as a single entity, result in the specific malformations seen in these patients.

In HSCR, deleterious deletions of chromosomal band 17q21[15] and 22q11[16], and duplications of 17q23[17], dup22p[15,18] and 22q11[19] have been previously described. Often, mutations in genes responsible for HSCR as well as CNVs overlapping them are described. Deletions of 10q11 [20,21] have been instrumental in identifying what we now know is the major responsible gene for familial and sporadic isolated HSCR, the REarranged during Transfection gene (*RET)* [22,23] Deletions of 13q[24-28] resulted in the identification of one of the genes responsible for Waardenburg-Shah syndrome type 4 (*EDNRB*)[29]. A deletion of another gene involved in this disorder,*SOX10*[30], has also been described[31]. Deletions of 2q[32-35] and 4p[36] contributed to the discovery of genes responsible for Mowatt-Wilson syndrome (*ZEB2*, formerly *ZFHX1B*)[37] and the gene responsible for Congenital Central Hypoventilation Syndrome *(PHOX2B)*[38]. In addition, more common CNPs are thought to be modifiers of the HSCR phenotype[39,40]. Therefore, we believe that rare CNVs could significantly contribute to patients with HSCR as well as other associated anomalies, where no pathogenic variant was

identified. Identification of these CNV can be instrumental to find yet unidentified genes that contribute to HSCR development in these patients.

In this manuscript, we test this hypothesis by comparing the Copy Number profiles of 23 patients with HSCR as a feature along with other associated anomalies and, for which no identified causal variant was identified (Group 1); patients with HSCR and a known mutation in *RET* or another causal gene (Group 2); patients with HSCR but without a *RET* mutation and without additional anatomical malformations or neurological defects (Group 3); and unaffected controls (Group 4). Our hypothesis pointed to the presence of more and larger CNVs in Group 1, and therefore, we aimed to identify the causal gene(s) for aganglionosis in these patients.

Materials & methods

Patient inclusion

In total 197 patients born between 1973 and 2018 were evaluated by a clinical geneticist in the department of Clinical Genetics, Erasmus Medical Centre, Rotterdam. Of these, 114 did not have associated anomalies nor a known syndrome. 29 patients had a known HSCR related genetic syndrome including Down syndrome (n=18). Of the 153 patients genetically evaluated, 21 had a (likely) deleterious RET mutation or variant. We selected 57 patients, of whom DNA and informed consent were available and in whom the RET gene was screened. Three subgroups of HSCR patients were included in the CNV detection study (1) patients with only HSCR, without a deleterious *RET* coding variant or other causal genetic defect (n=23), (2) patients with HSCR and a known mutation in *RET* or another causal gene (n=15), and patients with HSCR and additional anatomical malformations or neurological defects, but without a *RET* pathogenic variant, or other causal genetic defect (n=20). This project was approved by the Medical ethics committee of the Erasmus Medical Centre (Hirschsprung disease: no 2012-582, addendum No. 1 and no.193.948/2000/159, addendum Nos. 1 and 2). Parents of included patients gave informed consent prior to inclusion in this study.

Exclusion of the involvement of known disease genes

We determined the presence of *RET* coding mutations and intron-exon boundaries in all patients in this study. Furthermore, if a specific monogenetic syndrome was suspected based on the phenotypic spectrum observed, the coding sequence and intron-exon boundaries

of the suspected gene(s) were evaluated. In four HSCR patients with associated anomalies and nine HSCR patients without associated anomalies, the involvement of other known disease genes was excluded[10,41-43] using whole exome sequencing (WES), previously described pipelines[44,45] and variant prioritization methods[46].

Determination of Copy Number Variation

CNV profiles were determined with either the HumanCytoSNP-12 v2.1 or the Infinium Global Screening Array-24 v1.0 (Illumina Inc., San Diego, CA, USA), using methods and analysis settings previously described[47]. CNV profiles were inspected visually in Biodiscovery Nexus CN8.0 (Biodiscovery Inc., Hawthorne, CA, USA). CNVs with an overlap of at least 75%, with similar state CN changes, were either classified as rare, when absent from large control cohorts (n=19,584), or as a known modifier[11,13]. Size, type and gene content of rare CNVs were determined in HSCR patients (n=58) and unaffected controls (n=325). CNV count, size and gene content were compared between the control groups and previously described HSCR subgroups. All rare CNVs were uploaded to the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/). IDs and regions are depicted in supplementary table 1.

Evaluation of candidate gene expression

For CNVs to be involved in HSCR, we assumed that they should contain genes that are dosage sensitive, genes expressed in the developing intestine (copy number loss and gain), and genes with deleterious variation in other HSCR patients. All genes affected by rare deleterious CNVs were evaluated for expression in the mouse ENS between E11.5 and E15.5[48,49]. Since data from human intestines was only available for

embryonic week (EW) 12, EW14 and EW16, we evaluated gene expression in these time points[50]. Data was downloaded from the gene expression omnibus (GSE34208 and GSE100130).

Variant prioritization in loss of function genes

To determine whether or not a gene affected by a rare putative deleterious CNV was predicted to be intolerant to genetic variation[51,52] allowing some tolerance to account for reduced penetrance, we used capture-specific controls to eliminate technical noise in (1) a WES cohort of sporadic HSCR (n=76, 149 controls) and (2) a WES cohort of 443 short segment HSCR patients and 493 unaffected controls. Variants from WES data previously generated were prioritized as follows: an allele frequency below 1% in *in-house* unaffected controls (n=906); affect a gene that is intolerant to variation (pLI 0.85, Prec 0.90, synonymous or missense z-score at least 3 for missense variants and >1 for deletions and duplications); and have an allele frequency of maximum 0.01 for homozygous recessive variants and of 0.001 for heterozygous variants in GnomAD. We used a CADD score of 20 as a measure for deleteriousness for missense variants. All variants within two bases of an intron-exon boundary were considered to affect splicing and were included in the "loss of function" category when considering gene variant intolerance. Using RVTESTS[53], a variant burden test was done comparing the variant burden in 443 short segment HSCR patients and 493 controls (see supplementary table 2). Loss of function per candidate gene is described in supplementary table 3. All rare putative deleterious loss of function variants unique to the

HSCR cohort were uploaded to the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/).

Genotyping of HSCR associated SNPs

Sanger sequencing was used to genotype all patients for SNPs known to be associated to HSCR[39,54-56]. Primer sequences can be found in the supplementary table 4. The haplotypes per group of patients and relative weighted risk scores, are depicted in figure 1. The relative weighted risk score of published common and relatively rare risk alleles near *RET*, *NRG1* and *SEMA3C/D* was calculated using the formula below[54-57] and results are presented in supplementary table 5:

$$(Ln\ (OR\ risk\ allele\ 1) * allele\ count) + (Ln\ (OR\ risk\ allele\ 2) * allele\ count) + \ etc.$$

Statistical analysis

The number and size of rare CNV, the number of rare losses and gains, the number of genes intolerant to variation (SNVs and CNVs), the number of genes expressed in mouse ENS per rare CNV, and the relative weighted risk score were determined and compared for the different groups with a single ANOVA test. If group differences existed ($P<0.05$), we determined which subgroups were significantly different using a two-tailed T-test. The results are presented in Figure 1.

Results

Patient phenotypes

In total, 83 out of 197 patients with HSCR, from whom the medical charts were evaluated, had additional anomalies. Of these patients 29 had a known syndrome and had a combination of HSCR with associated anomalies that could not be diagnosed as a known syndrome (Group1). The phenotypes of the HSCR patients with additional anatomical malformations or neurological defects are described in table 1. Group 2 and 3 are described in supplementary table 6. Group 4 are unaffected parents of patients with unrelated conditions admitted to our hospital. There were no overlapping phenotypes, and no enrichment of aganglionic segment length between groups.

Copy Number variation profiling

We could detect 56 rare CNVs in HSCR patients (see supplementary table 1). In group 1, 10 CN losses, 8 CN gains and 1 maternally inherited hemizygous loss on chromosome X in a male patient, were detected. In group 2, 5 CN losses and 8 CN gains were identified, and in group 3, 7 CN losses, 12 CN Gains, 4 homozygous losses and 1 hemizygous loss on chromosome X in a male patient, were found. Two of the homozygous losses in patient P_000490 (group 3) disrupt the *CFTR* locus. The total list of rare CNVs in HSCR patients is depicted in supplementary table 1. The 17q23.1 - q23.2 loss, the 22q11.21 - q11.22 gain, the 6p22.1 - p21.33 loss and the 7q36.1 gain were *de novo*. Two CNVs were inherited maternally: the 10q11.22 - q11.23 loss and the Xq28 loss. The inheritance pattern of other rare CNV could not be determined due to unavailability of parental DNA. Two regions had

Table 1. Hirschsprung patients without a RET mutation and additional phenotypical features

Patient	HSCR type	Other phenotypical characteristics
P_000482	Short	Hydrocephalus, macrocephaly, autism
P_000540	Short	Facial dysmorphisms
P_000494	Short	Cardiac defects (VSD, ASD PDA, tricuspid atresia), dysplastic ears, renal malrotation
P_000512	Short	epilepsy, intellectual disability
P_000553	Short	Cardiac defects (VSD, dextrocardia, PDA, double outlet right ventricle), intestinal malrotation
P_000559	Total colonic	dysmorphic features, tracheomalacia, cardiac defects (dilated left ventricle, absence of AV conduction)
P_000561	Short	Facial dysmorphisms, small fontanelle, gastro-esophageal reflux, laryngeal web
P_000555	Short	hypoplastic thumb, hearing loss, developmental delay, facial dysmorphisms
P_002459	Short	Hypospadias, mild autism
P_000567	Short	Facial dysmorphisms, hearing loss, microcephaly, immunological hypersensitivity, nevus flammeus
P_000536	Abnormal	Telecanthus upslant; short segment HSCR although a longer segment is abnormal ganglionated.
P_000562	Short	Cafe au lait spots, cardia defect (VSD)
P_000572	TIA	Retrognathia, skin abnormality, facial dysmorphisms, cardiac defect (pulmonary valve stenosis)
P_000568	Short	dysmorphic features, hydrocele testis, hemangioma
P_000478	Short	Hypertelorism, facial dysmorphisms
P_000520	Short	Mild facial dysmorphisms, sandal-gap of toe
P_000534	Short	hypospadias, anteriorly placed anus
P_002455	Short	Hypermobility of fingers; mild developmental delay, downslant
P_001763	Short	White hair lock, mild developmental day
P_000537	Short	motor delay, spastic hemiplegia, bronchopulmonary dysplasia, cardiac defect (PDA)
P_000528	Total colonic	intellectual disability
P_000573	Short	Epicanthal folds, small ears, broad eyebrows with mild synophrys
P_002450	Long	Developmental delay
P_002343	Short	Hypertelorism, long deeply grooved philtrum

overlap in our small cohort: a gain in band 1p36.11 in patients P_000544, P_000566 (group 2) and P_001636 (group 3), covering the blood group type gene RHCE, and a 4q32.1 loss in patients P_000557 (group 2) and P_000515 (group 3) of which only the *ETFDH* gene is affected in both patients. The number of rare CNVs (P=0.385), number of rare losses (P= 0.420) and number of rare gains (P= 0.731), did not differ between unaffected individuals and any of the HSCR groups. Absence of a rare CNV per patient also did not differ between any of the groups (P= 0.363). However, when comparing CNV size, there was a significant difference between unaffected individuals (Group 4), and HSCR patients with associated anomalies without a known mutation (Group 1). The average CNV size in Group 1 is 689351bp compared to 156518bp in unaffected individuals (P= $7.297E^{-6}$). This effect was not present in Groups 2 or 3 (see figure 1). This difference could be specifically attributed to the size of Copy Number loss in HSCR patients with associated anomalies and without a known mutation.

CNVs with genes expressed in the ENS and Genes intolerant to variation and/or CN changes

In total, 1216 genes or transcripts were present in group 1-3 and group 4. Of these, 514 did not have a known mouse orthologue or probes in the microarray datasets (n=279 in group 1-3, n=235 in group 4). Differential expression of 472 genes (n=179 in group 1-3, n=293 in group 4) could also not be measured. A total of 230 genes affected by a rare CNV, had upregulated expression in the developing mouse ENS (n=91 in group 1-3, n=139 in group 4). These, were considered to be ENS genes (supplementary table 7). Rare CNVs present in HSCR

Figure 1 Enrichment of ENS genes in rare Copy Number losses

	P(T<=t) two-tailed							Mean values				
	Control vs Isolated	Control vs Syndromic	Control vs Mutation +	Control vs HSCR all	Syndromic vs Isolated	Syndromic vs Mutation +	Isolated vs Mutation +	Control (n=326)	Isolated (n=20)	Syndromic (n=23)	Mutation + (n=15)	HSCR all (n=58)
rare CNV Size	0.6470	0.0000	0.4620	0.0177	0.0391	0.4680	0.0558	156518	109500	708952	69472	336862
rare Loss Size	0.7000	0.0000	0.7050	0.0040	0.0590	0.8380	0.1000	55062	24817	574906	20779	241911
ENS genes in a rare CNV	0.9470	0.0000	0.5510	0.0020	0.0650	0.5260	0.0820	0.5736	0.5500	3.2174	0.3333	1.5517
ENS genes in a rare Loss	0.7480	0.0000	0.5330	0.0000	0.0660	0.5540	0.1010	0.2178	0.1500	2.7826	0.0667	1.1724
Variant intolerant gene in a rare CNV	0.3975	0.0000	0.9782	0.0284	0.0858	0.3776	0.2018	0.3405	0.1500	1.5217	0.3333	0.7414
Loss intolerant gene in a rare Loss	0.9706	0.0119	0.9442	0.1133	0.3239	0.8394	0.4129	0.0552	0.0500	0.4783	0.0667	0.2241
Intolerant gene in a rare CNV	0.4100	0.0002	0.8860	0.0543	0.1008	0.4434	0.2166	0.4509	0.2000	1.8696	0.4000	0.9138
ENS and Intolerant gene in a rare Loss	0.6550	0.0010	0.9920	0.0250	0.2640	0.7370	0.2650	0.0675	0.1000	0.3478	0.0667	0.1897
Relative weighted risk score	*NA*	*NA*	*NA*	*NA*	0.0265	6.94E-05	0.0219	*NA*	8.560	7.167	5.409	7.193

Abbreviations: CNV; Copy Number Variation, ENS; Enteric Nervous System, *NA*; not available. Higher values in red, lower values in green. The relative weighted risk score of the syndromic samples did not differ from the average score of all samples (P=0.8452). However, the scores of the isolated samples were significantly higher (P=0.0277) and the mutation carriers significantly lower (0.0232) compared to the total set of HSCR samples.

patients contain these ENS genes: 74 in group 1 CNVs, versus 5 in group 2, 11 in group 3 and 187 in group 4 (P= $4.565E^{-6}$). This result is mostly dependent on the overrepresentation of ENS genes in rare losses in HSCR patients with associated anomalies and without a known mutation: 68 in group 1 CNVs, versus 1 in group 2, 3 in group 3 and 71 in group 4 (P= $4.564E^{-6}$) (see figure 1). Next, we evaluated the number of ENS genes intolerant to variation and found no statistically significant differences: 10 in group 1 CNVs, versus 3 in group 2, 2 in group 3 and 51 in group 4 (P= 0.093). The CN losses did differ significantly between groups, as the majority of ENS genes were located in losses in group 1 (n=8), versus 1 in group 2, 2 in group 3 and 22 in group 4 (P= 0.014). All rare CNVs with genes expressed in the mouse ENS are depicted in table 2.

Variant analysis in an independent isolated HSCR patient cohort

Several putative deleterious loss of function and missense variants were detected in genes covered by a rare CNV, which are intolerant to either missense and/or loss of function variation using WES and whole genome sequencing. With the burden analysis using RVTESTS, neither all combined variants (missense and loss of function variants), nor missense variants alone resulted in significantly enriched genes (see supplementary table 2). However, we did find loss of function variants in the genes *INTS2*, *MED13*, *PRRC2A*, *TUBB*, *LINGO2*, *SGMS*, *KMT2C* and *SLC6A6* (see supplementary table 8).

Table 2. Rare CNV with genes expressed in mouse ENS between E11.5 and E15.5

P-number	Chromosome Region	Event	Length	Cytoband	Probes	Class	Sex	ENS gene(s)
P_000302	chr3:14,406,477-14,509,088	CN Gain	102612	p25.1	53	LD	F	**SLC6A6**
P_000479	chr12:9,245,492-9,308,543	CN Gain	63052	p13.31	24	VUS	M	A2M
P_000479	chr2:40,624,267-40,646,501	CN Loss	22235	p22.1	11	LD	M	**SLC8A1**
P_000494	chr12:128,208,742-128,917,555	CN Gain	708814	q24.32	96	VUS	M	TMEM132C
P_000498	chr1:152,286,216-152,323,703	CN Gain	37488	q21.3	11	VUS	F	FLG
P_000502	chr15:80,527,215-80,603,142	CN Gain	75928	q25.1	22	VUS	F	CTXND1
P_000512	chr6:28,005,012-31,683,185	CN Loss	3678174	p22.1 - p21.33	403	LD	F	6M1-18, ABHD16A, ATAT1, ATP6V1G2, **DDR1**, PCR1, FLOT1, **GABBR1**, **GNL1**, HLA-H, IER3, MUCL3, OR11A1, OR2J2, OR2J3, PGBD1, PPP1R11, P1R18, **TUBB**, ZKSCAN4, ZNRD1-AS1, ZNRD1ASP, ZSCAN31
P_000520	chr18:45,755,986-45,787,673	CN Gain	31688	q21.1	13	VUS	M	ZBTB7C
P_000537	chr10:49,033,586-52,417,694	CN Loss	3384109	q11.22 - q11.23	183	LD	M	ARHGAP22, C10orf128, CHAT, FAM21A, **MAPK8**, NCOA4, SLC18A3, TIMM23, TIMM23B, MEM273, VSTM4, WASHC2A
P_000540	chrX:154,277,428-154,299,482	Hemi-zygous	22055	q28	5	VUS	M	CMC4, FUNDC2, MTCP2
P_000557	chr11:62,251,301-62,298,871	CN Gain	47571	q12.3	25	LD	M	**AHNAK**
P_000561	chr22:18,861,209-21,630,630	CN Loss	2769422	q11.21	446	LD	M	ARVCF, BCRP2, C22orf29, C22orf39, CDC45, COMT, DGCR14, DGCR2, ESS2, GP1BB, LZTR1, RIMBP3, RTL10, SLC7A4, UFD1, **UFD1L**, ZDHHC8, ZNF74
P_000567	chr17:58,076,721-60,362,868	CN Loss	2286148	q23.1 - q23.2	74	LD	M	BCAS3, HEATR6, **TBX2**, **USP32**

P_000573	chr2:216,214,577-216,299,733	CN Loss	85157	q35	49	VUS	F	FN1
P_000579	chr3:60,468,409-60,490,104	CN Loss	21696	p14.2	16	LD	M	**FHIT**
P_000582	chr7:4,929,022-5,218,030	CN Gain	289009	p22.1	59	VUS	M	MMD2
P_000582	chr2:102,658,576-102,847,088	CN Gain	188513	q11.2 - q12.1	57	VUS	M	IL1R1, IL1RL2
P_000582	chr7:5,239,584-5,401,976	CN Gain	162393	p22.1	54	VUS	M	SLC29A4, WIPI2
P_001632	chr22:18,687,210-19,060,954	CN Gain	373745	q11.21	75	VUS	M	DGCR2; known modifier CNV
P_001637	chr2:10,664,398-10,914,786	CN Gain	250389	p25.1	70	VUS	M	NOL10
P_001763	chr18:65,699,090-66,534,856	CN Gain	835767	q22.1	228	VUS	F	TMX3
P_002431	chr1:243,963,527-244,016,804	CN Loss	53278	q44	9	LD	F	**AKT3**
P_002431	chr7:95,845,896-96,004,178	CN Loss	158283	q21.3	16	VUS	F	SLC25A13
P_002450	chr9:28,393,380-28,462,962	CN Loss	69583	p21.1	21	VUS	M	LINGO2
P_002455	chr22:21,802,791-22,555,544	CN Gain	752756	q11.21 - q11.22	169	VUS	M	CCDC116, **MAPK1**, PPM1F, SDF2L1, TMEM191C, **YDJC**, YPEL1
P_002459	chr12:94,767,704-94,880,489	CN Loss	112786	q22	27	VUS	M	CCDC41, CEP83

Abbreviations: CN; Copy Number, LD; Likely deleterious, VUS; variant of unknown significance, M; Male, F, Female. In Bold genes expressed in ENS and intolerant to variation. Chromosomal regions according to build hg19. Complete list of rare CNV in supplementary table 2.

Table 3. Genes expressed in mouse ENS in rare deletions.

			TPM value			*ExAC/ GnomAD intolerance scores*				
Patient	Gene		EW12	EW14	EW16	Missense Z	pLI	Deletion single	Deletion CNV	CNV region
P_000479	*SLC8A1*	13.48	17.04	28.30	2.23	1.00	-0.02	0.31	chr2:40,623,036-40,647,068	
P_000512	*TUBB*#	1023.02	1273.48	1043.17	5.71	0.98	-2.85	-1.72		
P_000512	*DDR1*	61.78	43.99	78.50	2.35	0.00	1.02	0.10	chr6:27,987,776-31,686,497	
P_000512	*GNL1*	32.36	35.27	32.39	2.52	1.00	1.03	0.70		
P_000512	*GABBR1*	14.30	19.44	17.81	4.98	1.00	1.36	1.23		
P_000537	*MAPK8*	11.11	12.86	12.83	2.92	1.00	0.84	-2.25	chr10:48,660,703-52,431,193	
P_000561	*UFD1L*	13.49	16.63	12.91	2.77	1.00	1.06	-2.53	chr22:18,844,632-21,798,907	
P_000567	*TBX2*	18.56	25.50	12.51	1.75	0.99	-0.01	0.53		
P_000567	*BCAS3*	5.72	6.84	7.80	2.76	0.00	-0.15	0.55	chr17:57,997,015-60,405,483	
P_000567	*USP32*	9.64	10.11	10.26	3.55	1.00	-2.85	-0.93		
P_000579	*FHIT*	8.31	6.42	3.81	-1.14	0.03	1.53	0.45	chr3:60,464,669-60,492,636	
P_002431	*AKT3*	9.98	13.73	16.45	4.03	1.00	-2.61	-1.20	chr1:243,956,683-244,017,616	

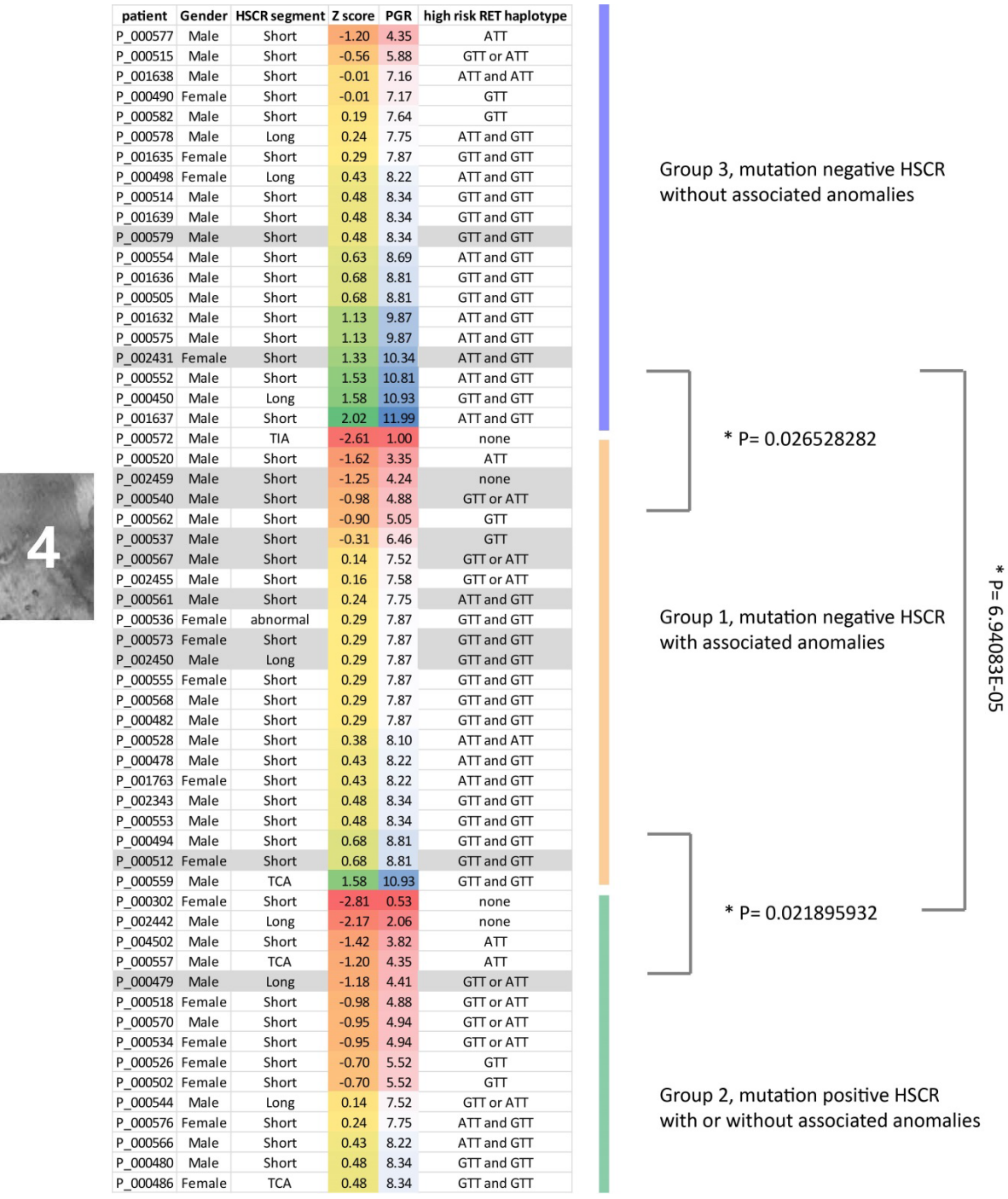
Genes marked with an # also have a loss of function variant in an independent HSCR whole exome sequencing cohort (see supplementary table 8) Depicted are the mean Transcripts Per Kilobase Million (TPM) values of human foetus in embryonic week (EW) 12-16 [50] and the CNV and variant intolerance scores derived from http://gnomad.broadinstitute.org/ and http://exac.broadinstitute.org/about.

Relative weighted risk score

The relative weighted risk score was determined for patients with HSCR as a feature along with other associated anomalies and no identified causal variant (Group 1); patients with HSCR and a known mutation in *RET* or another causal gene (Group 2); and HSCR patients without a *RET* mutation and without additional anatomical malformations or neurological defects (Group 3). The relative weighted risk score of group 1 (8.56) was lower than that of group 3 (7.1, P=0.0265) and higher than in group 2 (5.41). Moreover, the score of group 3 was significantly higher than in group 2 (P=6.941E-05). The patients individual polygenic risk scores (PGRS) are depicted in figure 3.

Figure 3 Polygenic risk scores

patient	Gender	HSCR segment	Z score	PGR	high risk RET haplotype
P_000577	Male	Short	-1.20	4.35	ATT
P_000515	Male	Short	-0.56	5.88	GTT or ATT
P_001638	Male	Short	-0.01	7.16	ATT and ATT
P_000490	Female	Short	-0.01	7.17	GTT
P_000582	Male	Short	0.19	7.64	GTT
P_000578	Male	Long	0.24	7.75	ATT and GTT
P_001635	Female	Short	0.29	7.87	GTT and GTT
P_000498	Female	Long	0.43	8.22	ATT and GTT
P_000514	Male	Short	0.48	8.34	GTT and GTT
P_001639	Male	Short	0.48	8.34	GTT and GTT
P_000579	Male	Short	0.48	8.34	GTT and GTT
P_000554	Male	Short	0.63	8.69	ATT and GTT
P_001636	Male	Short	0.68	8.81	GTT and GTT
P_000505	Male	Short	0.68	8.81	GTT and GTT
P_001632	Male	Short	1.13	9.87	ATT and GTT
P_000575	Male	Short	1.13	9.87	ATT and GTT
P_002431	Female	Short	1.33	10.34	ATT and GTT
P_000552	Male	Short	1.53	10.81	ATT and GTT
P_000450	Male	Long	1.58	10.93	GTT and GTT
P_001637	Male	Short	2.02	11.99	ATT and GTT
P_000572	Male	TIA	-2.61	1.00	none
P_000520	Male	Short	-1.62	3.35	ATT
P_002459	Male	Short	-1.25	4.24	none
P_000540	Male	Short	-0.98	4.88	GTT or ATT
P_000562	Male	Short	-0.90	5.05	GTT
P_000537	Male	Short	-0.31	6.46	GTT
P_000567	Male	Short	0.14	7.52	GTT or ATT
P_002455	Male	Short	0.16	7.58	GTT or ATT
P_000561	Male	Short	0.24	7.75	ATT and GTT
P_000536	Female	abnormal	0.29	7.87	GTT and GTT
P_000573	Female	Short	0.29	7.87	GTT and GTT
P_002450	Male	Long	0.29	7.87	GTT and GTT
P_000555	Female	Short	0.29	7.87	GTT and GTT
P_000568	Male	Short	0.29	7.87	GTT and GTT
P_000482	Male	Short	0.29	7.87	GTT and GTT
P_000528	Male	Short	0.38	8.10	ATT and ATT
P_000478	Male	Short	0.43	8.22	ATT and GTT
P_001763	Female	Short	0.43	8.22	ATT and GTT
P_002343	Male	Short	0.48	8.34	GTT and GTT
P_000553	Male	Short	0.48	8.34	GTT and GTT
P_000494	Male	Short	0.68	8.81	GTT and GTT
P_000512	Female	Short	0.68	8.81	GTT and GTT
P_000559	Male	TCA	1.58	10.93	GTT and GTT
P_000302	Female	Short	-2.81	0.53	none
P_002442	Male	Long	-2.17	2.06	none
P_004502	Male	Short	-1.42	3.82	ATT
P_000557	Male	TCA	-1.20	4.35	ATT
P_000479	Male	Long	-1.18	4.41	GTT or ATT
P_000518	Female	Short	-0.98	4.88	GTT or ATT
P_000570	Male	Short	-0.95	4.94	GTT or ATT
P_000534	Female	Short	-0.95	4.94	GTT or ATT
P_000526	Female	Short	-0.70	5.52	GTT
P_000502	Female	Short	-0.70	5.52	GTT
P_000544	Male	Long	0.14	7.52	GTT or ATT
P_000576	Female	Short	0.24	7.75	ATT and GTT
P_000566	Male	Short	0.43	8.22	ATT and GTT
P_000480	Male	Short	0.48	8.34	GTT and GTT
P_000486	Female	TCA	0.48	8.34	GTT and GTT



Polygenic risk scores (PGRS) of patients in group 1, 2 and 3 and their z-scores. Depicted in grey are the patients with a rare CN loss encompassing ENS expressed genes.

Discussion

Copy number variations (CNVs) contribute to the aetiology of many traits, common diseases and congenital anomalies[12]. However, for these CNVs to be involved in a disease they should contain genes that are both dosage sensitive and expressed in the tissue that is affected. Most known HSCR genes (see supplementary table 9) are intolerant to genetic variation[51] and are rarely impacted by CNV in unaffected individuals[11,13], but have been described to be impacted by CNV in HSCR patients[9,15]. This does not seem to be a frequent phenomenon, as we did not detect any CNV impacting a known disease gene. However, deleterious CNV seen in one patient can impact gene(s) that are more frequently affected by deleterious SNVs or small insertions or deletions, in other HSCR patients. This assumption is valid for many HSCR disease genes[9,15], mostly those associated to an inherited syndrome. Therefore, we hypothesized that rare CNV could significantly contribute to the aetiology of HSCR in patients with additional anomalies in which no pathogenic variant in a known disease gene could be identified. We also hypothesized that if present, this CNV could guide us to find new yet unidentified HSCR susceptibility genes.

Heterogeneous phenotypical spectrum

The additional anomalies seen in these patients are heterogeneous, and two patients have exactly the same spectrum of anomalies. No known syndromes were expected based on the patient' s phenotypes. The additional anomalies range from mild to very severe. Developmental delay, cardiac defects, hypospadias, intellectual disability and autism are seen in more than one patient. Most patients have short segment

HSCR and segment length seems not to differ between groups 1, 2 and 3. However, the number of patients and the high frequency of short segment HSCR does not allow for reliable statistical testing.

Rare CNVs are not increased in HSCR patients with associated anomalies

We determined the CN profiles of unaffected controls, HSCR patients without additional anomalies, patients with a known genetic aetiology and patients with additional anomalies without a known genetic aetiology. We could detect rare CNV in about half of HSCR patients (group 1-3) and unaffected controls (group 4). The number of rare CNVs did not differ between groups (P=0.385), nor did the number of rare CN losses (P=0.420) or gains (P= 0.731). This indicates that having a rare CNV is not related to disease development, which is already known. Previous studies, indicated that CNV size rather than number is associated with disease development[11,13], but this seems to contradict some of the results previously found in HSCR.

Tang and colleagues described that HSCR patients with associated anomalies have more and larger CNVs compared to isolated HSCR patients and unaffected controls[40]. An explanation might be that CNVs with a frequency below or comparable (0-0.026%) to HSCR prevalence (0.02-0.03%) are too rare to find a significant association in our study cohort size. CNV size, number, and type of affected genes, are likely more relevant.

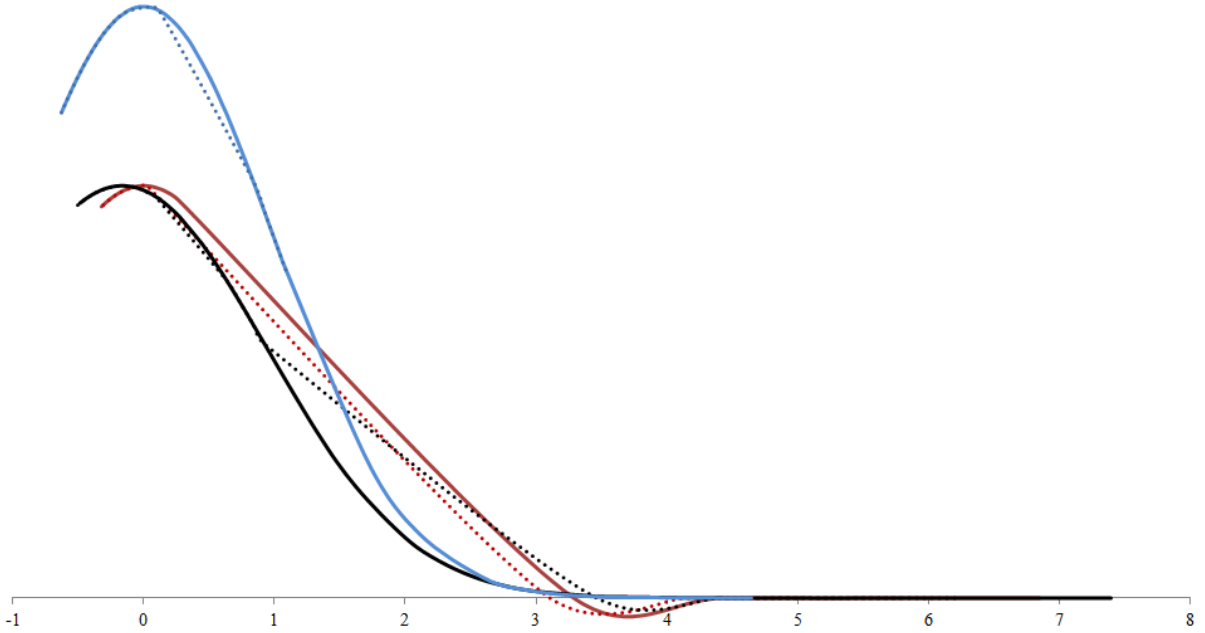
HSCR patients with associated anomalies without a known mutation, have larger rare CNVs.

If we compare CNV size, we see that this only correlates to group 1 and not to the other two HSCR subgroups. Not all HSCR patients with associated anomalies had a rare CNV, and the number of rare CNVs did not differ between unaffected individuals (Group 4) and HSCR subgroups 1 to 3, or even between HSCR subgroups. This indicates that if a patient has a rare CNV, this CNV is large. The Z-score distributions of total CNV size and CNV loss and gain of controls and HSCR patients, do not differ (see figure 2). The biggest difference is the average CNV size seen in group 1, compared to unaffected controls (group 4, $P<0.00001$), to isolated HSCR (group 2, $P=0.0177$), and likely also to HSCR patients with a mutation (group 3, $P=0.0558$). These results strongly suggest that rare CNV plays a role in some of the HSCR patients with associated anomalies.

The rare CNVs detected in HSCR patients with associated anomalies are often *de novo* and enriched for CN loss.

We determined segregation of the rare CNVs in nine patients, five of these were *de novo*. Although we did not determine the segregation of all rare CNVs, this already indicates the high frequency of *de novo* CNV in this cohort. Next, we compared the size of rare CN gains and of rare losses between groups. Only rare losses were enriched in HSCR patients with associated anomalies. The results are in-line with previous work by Tang and colleagues[40], but there was little CNV overlap between cohorts. Only five genes (*AIF1*, *APOM*, *ARHGAP22*, *CSNK2B*, *GALNT11*) were mutually affected. If CN losses contain dosage sensitive genes

Figure 2 Z- score distribution of CNV size



Depicted are the Z-scores of the all rare CNV (black lines), rare Copy Number Losses (red lines) and Gains (blue lines) Controls in continuous lines, all HSCR patients in dotted lines. Limited negative z-scores due to the size limit cut-off of 20kb.

expressed in the developing ENS, this could strongly suggest that these CNVs are related to HSCR development in these patients.

Genes expressed in the developing ENS are more frequently affected by rare losses

Using previously generated mouse transcriptome data, *in-house* [48] and from others [49], we determined which genes in rare CNVs were expressed in the developing mouse ENS between E11.5 and E15.5[48,49]. Using these transcriptome profiles, we could compare the number of ENS genes that were present in a rare CN loss. The rare CN losses in HSCR patients with associated anomalies (Group 1) contain more of these genes compared to the other two HSCR subgroups and unaffected controls (Group 4). Therefore, we can conclude that HSCR patients with associated anomalies are impacted by rare CN losses enriched in ENS expressed genes. Of the five genes mutually affected in our study and the previous study of Tang and colleagues[40], only the *Arhgap22* is upregulated in the developing mouse ENS. The expression of *Aif1* was highly upregulated in the developing mouse intestine compared to the ENS, and we could not measure differential expression of three genes (*Apom, Galnt11,* and *Csnk2b*). Unfortunately, as we used micro-array experiments as determinant of ENS expression, we cannot distinguish reliably between not expressed and not differentially expressed. However, *CSNK2B* is highly expressed in human embryonic intestine (EW12-EW16) (see supplementary table 7).

Rare Copy Number losses in HSCR patients with associated anomalies are enriched for ENS genes intolerant to genetic variation

Determining which gene(s) may be causal within a CNV can also be predicted by determining which genes are intolerant to rare genetic variation. Interestingly, many of the genes with a high intolerance to loss of function variants as well as missense variants are also sensitive for CN gains and losses[51,52]. Comparing the gene content of the rare CNVs between unaffected controls (Group 4) and HSCR subgroups 1 to 3, revealed that the rare CNVs from Group 1 contained more of these intolerant genes (P=0.0002). Moreover, the CN losses in this group contained more intolerant genes expressed in the ENS (P=0.001) compared to Group 4. We believe that these genes are excellent HSCR candidate genes, as many of the known HSCR genes are also intolerant to variation and expressed in the developing ENS (see supplementary table 9).

Overlap with previously described CNV and candidate genes in rare losses

Most of the genes intolerant to variation, present in a loss and overexpressed in the developing mouse ENS, can be considered to be excellent HSCR candidate genes. The genes we identified here are also expressed at moderate (*AKT3*, *GNL1*, *GABBR1*, *SLC8A1*, *MAPK8*, *UFD1L*, *TBX2*, *USP32*) or high levels (*TUBB*), in human embryonic intestine (see supplementary table 10). There was no overlap of CNV with any of the known HSCR genes (see supplementary table 9). In the DECIPHER database, patient phenotypes and genetic data are stored. This database is publicly available (https://decipher.sanger.ac.uk/) and

contains 18 patients with HSCR patients with associated anomalies with one or more CNV. Of these, 10 are *de novo*.

The 22q11 deletion described in one of these patients (ID 249397) had overlap with the 22q11 deletion found in patient P_000561. Patient P_000561 has the typical 22q11 deletion seen in 22q11 deletion syndrome. This deletion has a high penetrance with a variable phenotype[11,13,14]. HSCR has been previously described in patients with a 22q11 deletion [19] as well as disturbances in migration of neural crest cells[58]. One of the main candidate genes for HSCR is *UFD1L* as this encodes for a downstream target of HAND2[59]. *Hand2*$^{+/-}$ mice have decreased numbers of enteric neurons, neuronal differentiation defects and a disorganized ENS[60,61]. Mice with a targeted deletion encompassing *Ufd1l* as well as other genes of the 22q11 deletion syndrome seen in humans, are shown to model the syndrome[62] . A targeted deletion in mice encompassing less genes, but including *Ufd1l*, also modelled these features[63]. HSCR is a rare phenomenon in 22q11 deletion syndrome[64] and it is therefore, likely that other factors contribute to HSCR development.

One of the losses described in DECIPHER patient ID 249405 overlaps with the 1q44 loss in patient P_002431.The *AKT3* gene is the only gene affected by both CNVs. The AKT serine/threonine-protein kinase 3 is a core component of the phosphatidylinositol-3-kinase (PI3K)-AKT pathway and is expressed in the ENS. While activating pathogenic *de novo* missense variants are the cause of Megalencephaly-capillary malformation (MCAP)[65], 1q44 deletions are implicated in microcephaly and developmental anomalies of the brain[66]. These deletions can have

incomplete penetrance[67] and as far as we know, patient P_002431 has no associated anomalies in addition to HSCR. AKT3 and the PI3K-AKT pathway are directly activated by the RET. *RET* activating mutations are found in thyroid cancers and activate the PI3K-AKT pathway[68]. It is tempting to consider the opposite, that AKT deletions reduce the activation of this pathway and as such, influence neuronal development[69] or cell numbers[70]. *Akt3* knockout mice, as well as heterozygous mice, can develop normally[71]. However, the effect of the AKT deletion could act in combination with the common *RET* variants on *RET* expression, and downregulate the PI3K-AKT pathway substantially enough to affect enteric neuronal numbers.

Patient P_000512 has short segment HSCR, epilepsy and intellectual disability and a large *de novo* 6p22.1 - p21.33 deletion which affects the *GNL1*, *GABBR1* and *TUBB* genes amongst others. Little is known about the Guanine nucleotide binding protein like 1 (GNL1) function. There are no mouse models described in the mouse genome informatics database[72] (http://www.informatics.jax.org/). However, *in vitro* experiments indicate that GNL1 induces hyperphosphorylation of RB20 and promotes cell cycle progression and cell proliferation. Protein expression in HIPED[73] (https://www.genecards.org/) and mRNA expression in GTEX [74] (https://gtexportal.org/home/) of GNL1, highly correlates to KIF1BP. Moreover, both genes are expressed in human embryonic gut[50] and are upregulated in mouse developing ENS. Recessive *KIF1BP* (formerly *KIAA1279*) mutations are involved in Goldberg-Shprintzen syndrome which often includes HSCR as well as intellectual disability[75]. The gamma-aminobutyric acid B receptor 1

(*GABBR1*) is the receptor for the inhibitory neurotransmitter GABA, and is expressed in the submucosal and myenteric plexus of the ileum and colon[76]. Receptor agonist studies indicate there is a GABA subtype and intestinal region-specific impact on gastrointestinal motility[77-81]. The *TUBB* gene, coding for tubulin beta subtype 1, is deleted. Tubulins are the building blocks of the cytoskeleton and missense mutations in *TUBB* disrupt heterodimer formation by influencing chaperone interactions. Patients have circumferential skin creases, Kunze type, which includes circumferential skin creases, microcephaly, cleft palate, intellectual disability and other associated anomalies but also structural brain anomalies are described[82]. Breuss and colleagues described heterozygous *Tubb5* mice as having decreased brain volumes, and conclude that loss of function mutations in *TUBB* could also result in an abnormal phenotype[83]. The absence of enteric neurons in the distal colon and other neurological symptoms seen in Patient P_000512, can be likely related to the deletion of *TUBB* and subsequently to an increased neuronal apoptosis in combination with the effects of *GABBR1* and *GNL1*. Symptoms are less severe as seen in patients with dominant negative missense mutations, most likely due to the fact that TUBB heterodimers can still form, albeit at reduced levels.

Patient P_000567 has facial dysmorphisms, hearing loss, microcephaly, immunological hypersensitivity, nevus flammeus and this phenotype overlaps with other patients with 17q22-q23 deletions[84,85]. One of these patients is constipated, but HSCR disease has not been previously described[86,87]. Four genes expressed in the ENS are affected (*BCAS3*, *HEATR6*, *TBX2*, *USP32*) of which the latter two are also intolerant to

variation. There is little information about the Ubiquitin Specific Peptidase 32 (*USP32*), except that it is involved in endosomal transport and recycling[88]. T-Box 2 (*TBX2*) is a better candidate gene as this transcription factor is already known to be involved in the regulation of neural crest derived melanocytes[89], a process that is also hampered in specific patients with HSCR, and *EDN3*, *SOX10* and *EDNRB* mutations as these genes are part of the shared pathways[90,91]. *TBX2* heterozygous mice develop normally but homozygous mice have severe cardiovascular defects[92]. Similarly, *Bcass3, Microtubule Associated Cell Migration Factor* heterozygous mice do not have problems, but homozygous mice die *in utero* due to cardiovascular problems[93]. Taken together, it is likely that the combined effect of the *TBX2* and *BCAS3* on neural crest regulation and/or migration have a role in HSCR development in this patient[94].

Variant prioritization in loss of function genes

The population frequency of a specific deleterious variant related to a rare condition is low, especially if the causal gene is intolerant to variation[51]. In contrast, the occurrence of deleterious variants in such a gene is more common. The same holds true for a specific deleterious rare CNV related to the patient' s phenotype and having a rare CNV[11]. We believe that the chance occurrence of having both a deleterious loss of function variant in a gene intolerant to variation in one patient and a rare CNV overlapping that gene in another is very low, especially in such a relatively small cohorts of patients. Therefore, we screened two HSCR cohorts for the presence of variants in genes impacted by a rare CNV, and expressed in the developing mouse ENS (see table 3 and

supplementary table 7) and/or intolerant to variation (see supplementary table 10). We found several putative deleterious variants. However, burden analysis of missense variants with a high CADD score[95] did not result in significant enriched genes. However, we did find loss of function variants in genes intolerant to variation that were also impacted by the *de novo* 17 q23.1 - q23.2 loss (*INTS2*: NM_001330417.1: c.3172dupA and *MED13*: NM_005121: c.1968-1G>A), the *de novo* 6p22.1 - p21.33 loss (*PRRC2A*: (NM_004638.3: c.5836-1G>T *and TUBB* (NM_001293212.1: c.1330_1331delCAinsA), the 9p21 loss (*LINGO2*: NM_001258282: c.-395-2A>C) and maternal inherited 10q q11.22 - q11.23 loss (*SGMS1*: NM_147156: c.T529delT-:p.F177del, NM_147156: c.-313-2CAG>--G and NM_147156: c.-683+2T>C). These genes intolerant to variation, were either overexpressed in the developing mouse ENS and highly expressed in human embryonic intestine (*TUBB*), not overexpressed in developing mouse ENS but highly expressed in human embryonic ENS (*PRRC2A*), or not overexpressed in the mouse ENS and low to moderately expressed in human embryonic ENS (*INTS2*, *SGMS1, MED13*, see supplementary table 8 and 10). We also found loss of function variants in loss of function intolerant genes covered by the *de novo* 7q q36.1 Gain (*KMT2C*: NM_170606.2: c.1013-2A>G), and the 3q24 Gain (*SLC6A6*: NM_001134367: c.297-6A>G). The involvement of these genes is not as straightforward to determine as the loss of a gene in a CN loss, and a genetic variant resulting in the loss of function of a gene.

Common risk haplotypes' contributions differ between groups and fit the seesaw model

These findings strengthen the evidence for the involvement of these genes in HSCR. The question remains if these rare CNVs are the only cause for HSCR in these patients, or if they need other factors such as high impact rare variants or the known common risk haplotypes. Several haplotypes are associated with HSCR risk[54,55,96,97]. These associated common risk haplotypes have epistatic interactions, not only with each other, but also with known HSCR disease genes with deleterious rare variation such as *RET* and *NRG1*[55,57,98]. The SNPs near *RET* are located at a CTCF binding site (rs2505998), a highly conserved transcription factor binding site[57]. These reduce enhancer activity by disturbing the binding of RARB (rs25060300), GATA2 (rs7069690)[54,56] or SOX10 and P300 (rs2435357)[99]. Having more of these risk haplotypes increases disease risk substantially, especially if homozygous (rs2435357) in specific combinations[54,56], or together with other risk loci near the semaphorin gene cluster (rs7069590)[54,55] or *NRG1* (rs7005606, rs80227144)[96-98]. The rs8022744 haplotype has the highest effect on risk, and impacts a GATA6 binding site and a *SOX7* regulatory motif[57]. The rare Asian-specific SNP near *RET* (rs9282834) increases HSCR risk only modestly, but if present in trans with rs2435357, this risk increases drastically from an odds ratio of 1.1 to 26.7[57]. We hypothesized that *RET* acts as a fulcrum in a seesaw, balancing protective and harmful factors and high impact genetic factors such as other deleterious rare variants, trisomy 21 and pathogenic CNVs[100]. Variation in *RET* or common risk haplotypes, disturb this balance. Therefore, the more risk haplotypes present, the more the fulcrum shifts, resulting in a more easily disturbed

balance of the seesaw. In the same line of reasoning, the stronger the effect of high impact genetic factors, the easier the balance is disturbed. This model fits with earlier observations in specific genetic syndromes in which HSCR has a variable penetrance. In Down syndrome, Bardet Biedl syndrome and congenital central hypoventilation syndrome, epistatic interactions of the *RET* intron one risk haplotype was observed, whilst in Waardenburg syndrome (IV), Goldberg-Shprintzen syndrome and Mowat-Wilson syndrome it did not[101-105]. De Pontual and colleagues hypothesized that the effect of the risk haplotype is small in high penetrant monogenetic disorders and relatively larger in disorders in which the penetrance of HSCR is lower[101,102]. Our study also confirms this seesaw model: the HSCR subgroup without a high impact genetic variant, Group 3, has, on average, the highest contribution of the common haplotypes and the subgroup with a known pathogenic variant, Group 2, has the lowest. Patients with HSCR as a feature along with other associated anomalies and no identified causal variant (Group 1) have an intermediate common haplotype contribution. This makes sense, as this group likely consists of a mixture of high impact CNV, unidentified high impact mutation carriers, and patients with milder additional phenotypical features. It would be interesting to see if pathogenic genetic variation can be detected in undiagnosed patients with a low z-score (low contribution of the common variants).

PGRS and prediction of CNV involvement

We can hypothesize that the more risk variants (and higher PGRS) in a patient, the less likely it is that there will be a strong genetic factor identified. Moreover, if a strong genetic factor is identified, it is not

likely that another factor (a CNV) has a high impact. For instance, patient P_000479 has no associated anomalies, and carries a known pathogenic *RET* mutation: RET: NM_020630: c.656-21C>T. The PGRS in this patient is 4.41 (z-score of -1.18) which implies a low involvement of the common risk haplotypes, but a strong influence of a high impact genetic factor. The 2p22.1 loss in this patient has one ENS expressed gene, S*LC8A1*, which codes for a Na^{+}/Ca^{2+} exchanger. Homozygous knockout mice for this gene, have underdeveloped hearts due to increased cell apoptosis and a dilated pericardium, these mice die *in utero*[106]. However, heterozygous *Slc8a1* deficient mice develop and mature normally, without evident phenotypical abnormalities[106]. At a cellular level, they have increased B cell function, proliferation and mass[107]. It is therefore, not likely that the 2p22 loss contributes to the HSCR phenotype. Patient P_000537 has motor delay, spastic hemiplegia, bronchopulmonary dysplasia, cardiac defect (PDA) and short segment HSCR. The maternally inherited 10q11.22 - q11.23 loss impacts the Mitogen-Activated Protein Kinase 8 gene (*MAPK8*). The Mitogen activated protein (MAP) kinase pathway needs to be activated for ENCCs to migrate properly[108]. The fact that this deletion is inherited from an unaffected mother and this patient has a relatively low PGRs (6.46, z score of -0.31), would indicate that there are likely more (genetic) factors to be identified in this patient' s HSCR aetiology (and would make this patient an excellent candidate for further genetic analysis using whole exome or genome sequencing). Patients P_000567 (PGRS: 7.52, z-score 0.14), P_000512 (PGRS: 8.81, z-score 0.68) and P_000561 (PGRS:7.75, z-score 0.24), have intermediate risk scores and *de novo* deletions. This would imply that their CNVs have an impact,

but that other factors are needed to develop HSCR. Not all patients with a 22q11 deletion, or any of the published patients with a 17q22-q23 deletion have HSCR. It could very well be that the relatively high impact of the common risk haplotypes, in combination with the *de novo* CNV impacting ENS expressed genes, results in a shift from normal to defective ENS development.

Overlap of rare CNVs with existing ENS animal models

Within the rare CNVs identified in our patient cohort, only one of the genes are associated with existing HSCR or ENS mouse or zebrafish knockouts. *RADIL* in which we find a CN gain in one patient, has been knocked out in a zebrafish model with neural crest defects, including in the ENS[109]. It is not clear if a gain in this gene may also cause disruption to developmental pathways. The use of CRISPR-Cas systems to create zebrafish with deletions of the most likely contributing genes from CNVs (e.g. *TUBB*, *AKT3*, *UFD1L*, *TBX2*, *PRRC2A*) would help to elucidate their roles in ENS development. Whilst finding abnormal numbers of neurons would strengthen the evidence for these genes, lack of a phenotype in fish would not exclude their involvement, as it is not unlikely that the contribution of the common risk haplotypes may also be required for HSCR development in these cases.

Conclusions

In addition to the increased burden for rare CNV in sporadic isolated HSCR[40] and the role of CNPs as modifiers[39,40,110], rare, and often *de novo* CNVs impact HSCR patients with associated anomalies substantially. Rare Copy Number losses, particularly, are enriched for genes expressed in the ENS and intolerant to genetic variation. The high frequency of these rare CNVs highlights the importance of CN profiling in HSCR patients with associated anomalies in this whole genome and whole exome sequencing era. Given the low cost of SNP-array chips and/or the ability to already evaluate large CN changes with exome and genome wide sequencing technologies, screening HSCR patients seems warranted, especially when additional anomalies are present.

References

1. Nagy N, Goldstein AM. Enteric nervous system development: A crest cell's journey from neural tube to colon. *Seminars in cell & developmental biology.* 2017;66:94-106.
2. Burns AJ, Douarin NM. The sacral neural crest contributes neurons and glia to the post-umbilical gut: spatiotemporal analysis of the development of the enteric nervous system. *Development (Cambridge, England).* 1998;125(21):4335-4347.
3. Shepherd I, Eisen J. Development of the zebrafish enteric nervous system. *Methods in cell biology.* 2011;101:143-160.
4. Lake JI, Heuckeroth RO. Enteric nervous system development: migration, differentiation, and disease. *American journal of physiology Gastrointestinal and liver physiology.* 2013;305(1):G1-G24.
5. Moore SW, Zaahl MG. A review of genetic mutation in familial Hirschsprung's disease in South Africa: towards genetic counseling. *Journal of pediatric surgery.* 2008;43(2):325-329.
6. Moore SW, Zaahl M. Clinical and genetic correlations of familial Hirschsprung's disease. *Journal of pediatric surgery.* 2015;50(2):285-288.
7. Engum SA, Petrites M, Rescorla FJ, Grosfeld JL, Morrison AM, Engles D. Familial Hirschsprung's disease: 20 cases in 12 kindreds. *Journal of pediatric surgery.* 1993;28(10):1286-1290.
8. Russell MB, Russell CA, Fenger K, Niebuhr E. Familial occurrence of Hirschsprung's disease. *Clinical genetics.* 1994;45(5):231-235.
9. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *J Med Genet.* 2008;45(1):1-14.
10. Alves MM, Sribudiani Y, Brouwer RW, et al. Contribution of rare and common variants determine complex diseases-Hirschsprung disease as a model. *Developmental biology.* 2013;382(1):320-329.
11. Cooper GM, Coe BP, Girirajan S, et al. A copy number variation morbidity map of developmental delay. *Nat Genet.* 2011;43(9):838-846.
12. Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med.* 2010;61:437-455.
13. Coe BP, Witherspoon K, Rosenfeld JA, et al. Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet.* 2014;46(10):1063-1071.
14. Rosenfeld JA, Coe BP, Eichler EE, Cuckle H, Shaffer LG. Estimates of penetrance for recurrent pathogenic copy-number variations. *Genetics in medicine : official journal of the American College of Medical Genetics.* 2013;15(6):478-481.
15. Amiel J, Lyonnet S. Hirschsprung disease, associated syndromes, and genetics: a review. *J Med Genet.* 2001;38(11):729-739.
16. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal deletion map of human malformations. *Am J Hum Genet.* 1998;63(4):1153-1159.

17. Brewer C, Holloway S, Zawalnyski P, Schinzel A, FitzPatrick D. A chromosomal duplication map of malformations: regions of suspected haplo- and triplolethality--and tolerance of segmental aneuploidy--in humans. *Am J Hum Genet.* 1999;64(6):1702-1708.
18. Mahboubi S, Templeton JM, Jr. Association of Hirschsprung's disease and imperforate anus in a patient with "cat-eye" syndrome. A report of one case and review of the literature. *Pediatr Radiol.* 1984;14(6):441-442.
19. Kerstjens-Frederikse WS, Hofstra RM, van Essen AJ, Meijers JH, Buys CH. A Hirschsprung disease locus at 22q11? *J Med Genet.* 1999;36(3):221-224.
20. Fewtrell MS, Tam PK, Thomson AH, et al. Hirschsprung's disease associated with a deletion of chromosome 10 (q11.2q21.2): a further link with the neurocristopathies? *J Med Genet.* 1994;31(4):325-327.
21. Martucciello G, Bicocchi MP, Dodero P, et al. Total colonic aganglionosis associated with interstitial deletion of the long arm of chromosome 10. *Pediatric Surgery International.* 1992;7(4):308-310.
22. Edery P, Lyonnet S, Mulligan LM, et al. Mutations of the RET proto-oncogene in Hirschsprung's disease. *Nature.* 1994;367(6461):378-380.
23. Romeo G, Ronchetto P, Luo Y, et al. Point mutations affecting the tyrosine kinase domain of the RET proto-oncogene in Hirschsprung's disease. *Nature.* 1994;367(6461):377-378.
24. Bottani A, Xie YG, Binkert F, Schinzel A. A case of Hirschsprung disease with a chromosome 13 microdeletion, del(13)(q32.3q33.2): potential mapping of one disease locus. *Hum Genet.* 1991;87(6):748-750.
25. Kiss P, Osztovics M. Association of 13q deletion and Hirschsprung's disease. *J Med Genet.* 1989;26(12):793-794.
26. Lamont MA, Fitchett M, Dennis NR. Interstitial deletion of distal 13q associated with Hirschsprung's disease. *J Med Genet.* 1989;26(2):100-104.
27. Shanske A, Ferreira JC, Leonard JC, Fuller P, Marion RW. Hirschsprung disease in an infant with a contiguous gene syndrome of chromosome 13. *Am J Med Genet.* 2001;102(3):231-236.
28. Sparkes RS, Sparkes MC, Kalina RE, Pagon RA, Salk DJ, Disteche CM. Separation of retinoblastoma and esterase D loci in a patient with sporadic retinoblastoma and del(13)(q14.1q22.3). *Hum Genet.* 1984;68(3):258-259.
29. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. *Cell.* 1994;79(7):1257-1266.
30. Inoue K, Tanabe Y, Lupski JR. Myelin deficiencies in both the central and the peripheral nervous systems associated with a SOX10 mutation. *Annals of neurology.* 1999;46(3):313-318.
31. Bondurand N, Dastot-Le Moal F, Stanchina L, et al. Deletions at the SOX10 gene locus cause Waardenburg syndrome types 2 and 4. *Am J Hum Genet.* 2007;81(6):1169-1185.
32. Amiel J, Espinosa-Parrilla Y, Steffann J, et al. Large-scale deletions and SMADIP1 truncating mutations in syndromic Hirschsprung disease with

involvement of midline structures. *Am J Hum Genet.* 2001;69(6):1370-1377.
33. Lurie IW, Supovitz KR, Rosenblum-Vos LS, Wulfsberg EA. Phenotypic variability of del(2) (q22-q23): report of a case with a review of the literature. *Genet Couns.* 1994;5(1):11-14.
34. McMilin KD, Reiss JA, Brown MG, et al. Clinical outcomes of four patients with microdeletion in the long arm of chromosome 2. *Am J Med Genet.* 1998;78(1):36-43.
35. Mowat DR, Croaker GD, Cass DT, et al. Hirschsprung disease, microcephaly, mental retardation, and characteristic facial features: delineation of a new syndrome and identification of a locus at chromosome 2q22-q23. *J Med Genet.* 1998;35(8):617-623.
36. Benailly HK, Lapierre JM, Laudier B, et al. PMX2B, a new candidate gene for Hirschsprung's disease. *Clinical genetics.* 2003;64(3):204-209.
37. Wakamatsu N, Yamada Y, Yamada K, et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat Genet.* 2001;27(4):369-370.
38. Amiel J, Laudier B, Attie-Bitach T, et al. Polyalanine expansion and frameshift mutations of the paired-like homeobox gene PHOX2B in congenital central hypoventilation syndrome. *Nat Genet.* 2003;33(4):459-461.
39. Jiang Q, Ho YY, Hao L, Nichols Berrios C, Chakravarti A. Copy number variants in candidate genes are genetic modifiers of Hirschsprung disease. *PloS one.* 2011;6(6):e21219.
40. Tang CS, Cheng G, So MT, et al. Genome-wide copy number analysis uncovers a new HSCR gene: NRG3. *PLoS genetics.* 2012;8(5):e1002687.
41. Gui H, Schriemer D, Cheng WW, et al. Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. *Genome biology.* 2017;18(1):48.
42. Lai FP, Lau ST, Wong JK, et al. Correction of Hirschsprung-Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function. *Gastroenterology.* 2017;153(1):139-153 e138.
43. Tang CS, Li P, Lai FP, et al. Identification of Genes Associated With Hirschsprung Disease, Based on Whole-Genome Sequence Analysis, and Potential Effects on Enteric Nervous System Development. *Gastroenterology.* 2018;155(6):1908-1922.e1905.
44. Brouwer RWW, van den Hout M, Kockx CEM, et al. Nimbus: A design-driven analyses suite for amplicon based NGS data. *Bioinformatics (Oxford, England).* 2018.
45. Brouwer RW, van den Hout MC, Grosveld FG, van Ijcken WF. NARWHAL, a primary analysis pipeline for NGS data. *Bioinformatics (Oxford, England).* 2012;28(2):284-285.
46. Halim D, Brosens E, Muller F, et al. Loss-of-Function Variants in MYLK Cause Recessive Megacystis Microcolon Intestinal Hypoperistalsis Syndrome. *Am J Hum Genet.* 2017;101(1):123-129.

47. Brosens E, Marsch F, de Jong EM, et al. Copy number variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula. *Eur J Hum Genet.* 2016;24(12):1715-1723.
48. Schriemer D, Sribudiani Y, A IJ, et al. Regulators of gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes. *Developmental biology.* 2016;416(1):255-265.
49. Memic F, Knoflach V, Morarach K, et al. Transcription and Signaling Regulators in Developing Neuronal Subtypes of Mouse and Human Enteric Nervous System. *Gastroenterology.* 2018;154(3):624-636.
50. McCann CJ, Alves MM, Brosens E, et al. Neuronal Development and Onset of Electrical Activity in the Human Enteric Nervous System. *Gastroenterology.* 2019.
51. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536(7616):285-291.
52. Ruderfer DM, Hamamsy T, Lek M, et al. Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. *Nat Genet.* 2016;48(10):1107-1111.
53. Zhan X, Hu Y, Li B, Abecasis GR, Liu DJ. RVTESTS: an efficient and comprehensive tool for rare variant association analysis using sequence data. *Bioinformatics (Oxford, England).* 2016;32(9):1423-1426.
54. Kapoor A, Jiang Q, Chatterjee S, et al. Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms. *Human molecular genetics.* 2015;24(10):2997-3003.
55. Jiang Q, Arnold S, Heanue T, et al. Functional loss of semaphorin 3C and/or semaphorin 3D and their epistatic interaction with ret are critical to Hirschsprung disease liability. *Am J Hum Genet.* 2015;96(4):581-596.
56. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease. *Cell.* 2016;167(2):355-368.e310.
57. Tang CS, Gui H, Kapoor A, et al. Trans-ethnic meta-analysis of genome-wide association studies for Hirschsprung disease. *Human molecular genetics.* 2016;25(23):5265-5275.
58. Scambler PJ. The 22q11 deletion syndromes. *Human molecular genetics.* 2000;9(16):2421-2426.
59. Yamagishi H, Garg V, Matsuoka R, Thomas T, Srivastava D. A molecular pathway revealing a genetic basis for human cardiac and craniofacial defects. *Science.* 1999;283(5405):1158-1161.
60. Hendershot TJ, Liu H, Sarkar AA, et al. Expression of Hand2 is sufficient for neurogenesis and cell type-specific gene expression in the enteric nervous system. *Dev Dyn.* 2007;236(1):93-105.
61. D'Autreaux F, Margolis KG, Roberts J, et al. Expression level of Hand2 affects specification of enteric neurons and gastrointestinal function in mice. *Gastroenterology.* 2011;141(2):576-587, 587 e571-576.
62. Lindsay EA, Botta A, Jurecic V, et al. Congenital heart disease in mice deficient for the DiGeorge syndrome region. *Nature.* 1999;401(6751):379.

63. Vitelli F, Lindsay E, Baldini A. Genetic dissection of the DiGeorge syndrome phenotype. Paper presented at: Cold Spring Harbor symposia on quantitative biology2002.
64. McDonald-McGinn DM, Sullivan KE, Marino B, et al. 22q11.2 deletion syndrome. *Nature reviews Disease primers.* 2015;1:15071.
65. Riviere JB, Mirzaa GM, O'Roak BJ, et al. De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat Genet.* 2012;44(8):934-940.
66. Boland E, Clayton-Smith J, Woo VG, et al. Mapping of deletion and translocation breakpoints in 1q44 implicates the serine/threonine kinase AKT3 in postnatal microcephaly and agenesis of the corpus callosum. *Am J Hum Genet.* 2007;81(2):292-303.
67. Gai D, Haan E, Scholar M, Nicholl J, Yu S. Phenotypes of AKT3 deletion: a case report and literature review. *Am J Med Genet A.* 2015;167a(1):174-179.
68. Mulligan LM. RET revisited: expanding the oncogenic portfolio. *Nature reviews Cancer.* 2014;14(3):173-186.
69. Markus A, Zhong J, Snider WD. Raf and akt mediate distinct aspects of sensory axon growth. *Neuron.* 2002;35(1):65-76.
70. Easton RM, Cho H, Roovers K, et al. Role for Akt3/protein kinase Bgamma in attainment of normal brain size. *Molecular and cellular biology.* 2005;25(5):1869-1878.
71. Dummler B, Tschopp O, Hynx D, Yang ZZ, Dirnhofer S, Hemmings BA. Life with a single isoform of Akt: mice lacking Akt2 and Akt3 are viable but display impaired glucose homeostasis and growth deficiencies. *Molecular and cellular biology.* 2006;26(21):8042-8051.
72. Blake JA, Eppig JT, Kadin JA, et al. Mouse Genome Database (MGD)-2017: community knowledge resource for the laboratory mouse. *Nucleic Acids Res.* 2017;45(D1):D723-D729.
73. Fishilevich S, Zimmerman S, Kohn A, et al. Genic insights from integrated human proteomics in GeneCards. *Database : the journal of biological databases and curation.* 2016;2016:baw030.
74. Consortium GT, Laboratory DA, Coordinating Center -Analysis Working G, et al. Genetic effects on gene expression across human tissues. *Nature.* 2017;550(7675):204-213.
75. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous nonsense mutations in KIAA1279 are associated with malformations of the central and enteric nervous systems. *Am J Hum Genet.* 2005;77(1):120-126.
76. Casanova E, Guetg N, Vigot R, et al. A mouse model for visualization of GABA(B) receptors. *Genesis.* 2009;47(9):595-602.
77. Ong J, Kerr DI. GABAA- and GABAB-receptor-mediated modification of intestinal motility. *Eur J Pharmacol.* 1982;86(1):9-17.
78. Krantis A, Harding RK. GABA-related actions in isolated in vitro preparations of the rat small intestine. *Eur J Pharmacol.* 1987;141(2):291-298.
79. Giotti A, Luzzi S, Spagnesi S, Zilletti L. GABAA and GABAB receptor-mediated effects in guinea-pig ileum. *Br J Pharmacol.* 1983;78(3):469-478.

80. Pencheva N, Itzev D, Milanov P. Comparison of gamma-aminobutyric acid effects in different parts of the cat ileum. *Eur J Pharmacol.* 1999;368(1):49-56.
81. Auteri M, Zizzo MG, Mastropaolo M, Serio R. Opposite role played by GABAA and GABAB receptors in the modulation of peristaltic activity in mouse distal colon. *Eur J Pharmacol.* 2014;731:93-99.
82. Isrie M, Breuss M, Tian G, et al. Mutations in Either TUBB or MAPRE2 Cause Circumferential Skin Creases Kunze Type. *Am J Hum Genet.* 2015;97(6):790-800.
83. Breuss M, Fritz T, Gstrein T, et al. Mutations in the murine homologue of TUBB5 cause microcephaly by perturbing cell cycle progression and inducing p53-associated apoptosis. *Development (Cambridge, England).* 2016;143(7):1126-1133.
84. Ballif BC, Theisen A, Rosenfeld JA, et al. Identification of a recurrent microdeletion at 17q23.1q23.2 flanked by segmental duplications associated with heart defects and limb abnormalities. *Am J Hum Genet.* 2010;86(3):454-461.
85. Laurell T, Lundin J, Anderlid BM, et al. Molecular and clinical delineation of the 17q22 microdeletion phenotype. *Eur J Hum Genet.* 2013;21(10):1085-1092.
86. Schonewolf-Greulich B, Ronan A, Ravn K, et al. Two new cases with microdeletion of 17q23.2 suggest presence of a candidate gene for sensorineural hearing loss within this region. *Am J Med Genet A.* 2011;155A(12):2964-2969.
87. Khattab M, Xu F, Li P, Bhandari V. A de novo 3.54 Mb deletion of 17q22-q23.1 associated with hydrocephalus: a case report and review of literature. *Am J Med Genet A.* 2011;155A(12):3082-3086.
88. Sapmaz A, Berlin I, Bos E, et al. USP32 regulates late endosomal transport and recycling through deubiquitylation of Rab7. *Nature communications.* 2019;10(1):1454.
89. Pan L, Ma X, Wen B, et al. Microphthalmia-associated transcription factor/T-box factor-2 axis acts through Cyclin D1 to regulate melanocyte proliferation. *Cell proliferation.* 2015;48(6):631-642.
90. Stanchina L, Baral V, Robert F, et al. Interactions between Sox10, Edn3 and Ednrb during enteric nervous system and melanocyte development. *Developmental biology.* 2006;295(1):232-249.
91. Watanabe Y, Stanchina L, Lecerf L, et al. Differentiation of Mouse Enteric Nervous System Progenitor Cells Is Controlled by Endothelin 3 and Requires Regulation of Ednrb by SOX10 and ZEB2. *Gastroenterology.* 2017;152(5):1139-1150.e1134.
92. Harrelson Z, Kelly RG, Goldin SN, et al. Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development. *Development (Cambridge, England).* 2004;131(20):5041-5052.
93. Shetty R, Joshi D, Jain M, et al. Rudhira/BCAS3 is essential for mouse development and cardiovascular patterning. *Scientific reports.* 2018;8(1):5632.

94. Jain M, Bhat GP, Vijayraghavan K, Inamdar MS. Rudhira/BCAS3 is a cytoskeletal protein that controls Cdc42 activation and directional cell migration during angiogenesis. *Experimental cell research.* 2012;318(6):753-767.
95. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46(3):310-315.
96. Tang CS, Tang WK, So MT, et al. Fine mapping of the NRG1 Hirschsprung's disease locus. *PloS one.* 2011;6(1):e16181.
97. Garcia-Barcelo MM, Tang CS, Ngan ES, et al. Genome-wide association study identifies NRG1 as a susceptibility locus for Hirschsprung's disease. *Proceedings of the National Academy of Sciences of the United States of America.* 2009;106(8):2694-2699.
98. Gui H, Tang WK, So MT, et al. RET and NRG1 interplay in Hirschsprung disease. *Hum Genet.* 2013;132(5):591-600.
99. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature.* 2005;434(7035):857-863.
100. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental biology.* 2016;417(2):198-208.
101. de Pontual L, Pelet A, Trochet D, et al. Mutations of the RET gene in isolated and syndromic Hirschsprung's disease in human disclose major and modifier alleles at a single locus. *J Med Genet.* 2006;43(5):419-423.
102. de Pontual L, Pelet A, Clement-Ziza M, et al. Epistatic interactions with a common hypomorphic RET allele in syndromic Hirschsprung disease. *Human mutation.* 2007;28(8):790-796.
103. de Pontual L, Zaghloul NA, Thomas S, et al. Epistasis between RET and BBS mutations modulates enteric innervation and causes syndromic Hirschsprung disease. *Proceedings of the National Academy of Sciences of the United States of America.* 2009;106(33):13921-13926.
104. Arnold S, Pelet A, Amiel J, et al. Interaction between a chromosome 10 RET enhancer and chromosome 21 in the Down syndrome-Hirschsprung disease association. *Human mutation.* 2009;30(5):771-775.
105. Jannot AS, Pelet A, Henrion-Caude A, et al. Chromosome 21 scan in Down syndrome reveals DSCAM as a predisposing locus in Hirschsprung disease. *PloS one.* 2013;8(5):e62519.
106. Cho CH, Kim SS, Jeong MJ, Lee CO, Shin HS. The Na+ -Ca2+ exchanger is essential for embryonic heart development in mice. *Mol Cells.* 2000;10(6):712-722.
107. Nguidjoe E, Sokolow S, Bigabwa S, et al. Heterozygous inactivation of the Na/Ca exchanger increases glucose-induced insulin release, beta-cell proliferation, and mass. *Diabetes.* 2011;60(8):2076-2085.
108. Natarajan D, Marcos-Gutierrez C, Pachnis V, de Graaff E. Requirement of signalling by receptor tyrosine kinase RET for the directed migration of enteric nervous system progenitor cells during mammalian embryogenesis. *Development (Cambridge, England).* 2002;129(22):5151-5160.

109. Smolen GA, Schott BJ, Stewart RA, et al. A Rap GTPase interactor, RADIL, mediates migration of neural crest precursors. *Genes Dev.* 2007;21(17):000-000.
110. Bae JS, Koh I, Cheong HS, et al. A genome-wide association analysis of chromosomal aberrations and Hirschsprung disease. *Transl Res.* 2016;177:31-40 e36.
111. Firth HV, Richards SM, Bevan AP, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet.* 2009;84(4):524-533.


4

Supplementary Table 1.

Chr	Gene	mis_z	syn_z	pLI	pRec	pNull	del	dup	del.sing	dup.sing	del.sing.scc	dup.sing.sc	del.score	dup.score	cnv.score	flag	EW12_TPM	EW12_CPM	EW14_TPM	EW14_CPM	EW16_TPM	EW16_CPM	DDD_contr	DDD_contr	Inheritance	hg19 location
1	*ECE1*	2,410	0,017	0,956	0,044	0,000	5	8	5	7	2,371	2,699	0,900	0,646	0,720	0	65,208	142,968	70,109	145,284	60,633	132,936	0	0	D	chr1:21,543,740 21,616,766
2	*GLI2*	0,769	0,319	0,990	0,010	0,000	1	6	1	6	0,075	2,544	0,486	0,450	0,148	0	9,135	20,239	12,841	26,857	10,039	22,270	1	0	D; polygenic?	chr2:121,554,867 121,750,229
2	*IHH*	1,601	0,652	0,332	0,665	0,003	0	0	0	0	0,168	0,110	0,207	0,600	0,647	0	109,104	66,470	92,228	53,103	277,568	169,218	1	0	D; modifier	chr2:219,919,142 219,925,238
2	*ZEB2*	4,222	0,301	1,000	0,000	0,000	4	5	3	4	1,791	2,009	0,949	0,415	0,632	0	24,538	103,565	43,600	173,905	31,492	132,952	0	0	D	chr2:145,141,942 145,277,958
4	*PHOX2B*	2,210	1,135	0,944	0,056	0,000	0	0	0	0	0,185	0,091	0,190	0,622	0,663	0	28,286	26,380	37,190	32,805	43,565	40,618	0	0	D and repeat expansion	chr4:41,746,099 41,750,987
5	*GDNF*	1,116	0,991	0,197	0,761	0,042	0	2	0	0	0,263	0,135	0,147	0,123	0,030	0	8,788	11,123	18,489	22,111	20,473	25,890	0	0	D; polygenic?	chr5:37,812,779 37,835,929
7	*GLI3*	0,715	2,013	1,000	0,000	0,000	0	2	0	0	0,939	0,961	0,985	0,507	0,824	0	8,627	24,053	14,322	37,765	11,158	31,124	0	0	D; polygenic?	chr7:42,000,548 42,276,618
7	*SEMA3A*	1,904	0,646	0,983	0,017	0,000	1	4	1	2	0,526	0,077	0,686	0,084	0,358	0	16,864	34,642	28,043	54,450	28,352	58,241	1	1	D; uncertain	chr7:83,587,659 83,824,217
7	*SEMA3C*	1,078	0,790	0,001	0,999	0,000	4	3	3	2	1,421	0,392	0,811	0,196	0,191	0	40,877	77,330	45,260	80,897	74,331	140,503	2	0	D; uncertain; common SNP	chr7:80,371,854 80,548,667
7	*SEMA3D*	0,826	0,212	0,000	1,000	0,000	7	3	7	2	2,846	0,083	1,173	0,638	0,102	1	10,238	24,909	10,517	24,166	21,230	51,626	1	0	D; uncertain; common SNP	chr7:84,624,872 84,751,247
8	*DENND3*	2,557	0,151	0,000	1,000	0,000	3	6	3	4	1,347	1,504	0,382	0,384	0,385	0	2,737	12,116	3,959	16,575	4,197	18,585	0	1	D; modifier	chr8:142,138,720 142,205,900
8	*NRG1*	0,624	0,831	0,996	0,004	0,000	0	1	0	1	1,986	1,049	1,468	1,014	1,429	0	1,547	3,576	2,152	4,703	2,460	5,684	1	0	D and common SNP	chr8:32,405,728 32,622,558
10	*CYP26A1*	0,397	1,450	0,000	0,201	0,799	0	2	0	0	0,216	0,245	0,517	0,166	0,390	0	0,026	0,020	0,256	0,181	0,081	0,061	0	0	D; polygenic?	chr10:94,833,647 94,837,641
10	*GFRA1*	1,449	0,808	0,132	0,868	0,000	0	0	0	0	0,326	0,517	0,465	0,826	0,931	0	16,546	49,610	25,192	71,333	46,401	139,125	0	0	D; polygenic?	chr10:117,816,442 118,033,126
10	*KIF1BP*	1,543	1,746	0,092	0,908	0,000	2	0	2	0	1,299	0,382	0,294	0,961	0,547	0	27,609	22,492	33,930	26,128	39,952	32,536	0	1	D	chr10:70,748,477 70,776,739
10	*NRG3*	0,215	0,875	0,094	0,906	0,000	0	5	0	2	1,328	0,194	0,941	0,491	0,003	0	1,054	1,396	1,199	1,505	1,276	1,683	many small	3	D; modifier	chr10:83,635,070 84,746,935
10	*RET*	1,083	0,081	1,000	0,000	0,000	0	3	0	2	1,120	0,315	1,334	0,587	0,956	0	12,975	25,872	16,143	30,458	21,091	42,075	0	many	D and common SNPs	chr10:43,572,517 43,625,797
10	*SUFU*	1,961	0,727	1,000	0,000	0,000	4	2	1	0	0,085	0,978	0,621	0,643	0,155	0	14,465	29,974	15,086	29,555	15,695	32,516	1	0	D; polygenic?	chr10:104,263,719 104,393,214
10	*TBATA*	0,329	0,264	0,000	0,090	0,910	0	6	0	0	0,491	0,506	0,734	0,646	0,227	0	0,000	0,000	0,065	0,040	0,062	0,041	1	1	D; modifier	chr10:72,530,995 72,545,157
10	*VCL*	2,955	1,567	0,080	0,920	0,000	1	1	0	1	1,160	0,439	0,566	0,969	0,968	0	132,706	232,990	169,809	281,724	248,314	436,123	0	3	D; uncertain	chr10:75,757,872 75,879,914
11	*DHCR7*	0,322	1,613	0,000	0,125	0,875	1	2	1	0	0,750	0,213	0,083	0,119	0,086	0	33,003	41,791	40,837	48,893	39,410	49,924	0	0	D	chr11:71,145,457 71,159,477
11	*NUP98*	1,575	0,085	1,000	0,000	0,000	5	16	2	6	0,167	1,586	0,531	1,406	1,056	0	32,216	96,360	29,885	84,494	30,565	91,442	0	1	D; modifier	chr11:3,696,240 3,819,022
12	*CCT2*	1,516	0,099	0,999	0,001	0,000	0	3	0	2	0,883	0,554	0,975	0,225	0,591	0	131,107	174,450	128,800	162,075	133,138	177,127	0	3	D; uncertain	chr12:69,979,208 69,995,357
12	*GLI1*	1,043	0,957	0,000	0,780	0,220	1	1	0	0	0,552	0,506	0,238	0,689	0,635	0	13,269	16,829	18,191	21,814	12,289	15,602	1	4	D; polygenic?	chr12:57,853,918 57,866,047
13	*EDNRB*	1,239	0,726	0,012	0,988	0,000	3	0	2	0	1,373	0,372	0,822	0,897	0,208	0	43,726	66,442	48,587	69,795	75,311	114,395	0	1	D and AR	chr13:78,469,616 78,493,903
14	*VASH1*	1,701	0,960	0,309	0,688	0,003	0	0	0	0	0,194	0,240	0,485	0,847	0,953	0	28,794	57,921	42,794	81,384	23,474	47,242	0	0	D; modifier	chr14:77,228,235 77,249,363
19	*NCLN*	1,834	0,625	0,986	0,014	0,000	2	2	2	2	1,141	0,840	0,316	0,329	0,105	0	94,376	118,897	91,223	108,701	59,658	75,233	0	0	D; modifier	chr19:3,185,875 3,209,573
19	*NRTN*	0,820	1,088	0,028	0,811	0,160	.	.	.	.	.	.	.	.	.	.	6,677	2,236	9,416	2,986	7,018	2,350	0	1	D	chr19:5,823,818 5,828,335
19	*PSPN*	0,283	0,720	0,000	0,352	0,648	.	.	.	.	.	.	.	.	.	.	2,144	0,305	1,535	0,204	0,514	0,074	0	0	D	chr19:6,375,305 6,375,860
20	*EDN3*	1,046	0,053	0,028	0,926	0,047	0	0	0	0	0,357	0,354	0,665	0,974	1,119	0	8,169	6,593	11,553	8,819	25,675	20,727	0	0	D and AR	chr20:57,875,499 57,901,047
21	*BACE2*	1,670	0,987	0,006	0,990	0,004	0	4	0	1	0,434	0,275	0,650	0,285	0,056	0	16,077	17,940	29,369	30,980	18,627	20,809	0	1	D; uncertain	chr21:42,539,728 42,648,524
22	*SOX10*	2,883	1,216	0,991	0,009	0,000	0	1	0	0	0,224	0,153	0,152	0,177	0,265	0	6,408	7,172	8,786	9,289	9,188	10,287	0	0	D	chr22:38,368,319 38,380,539
X	*L1CAM*	2,757	1,423	1,000	0,000	0,000	.	.	.	.	.	.	.	.	.	.	18,772	42,158	28,338	60,211	31,734	71,338	.	.	XLR	chrX:153,126,969 153,151,628

Supplementary Table 2.

P number	Chromosor	Event	Length	Cytoband	Probes	Min Size	Min Regior	Max Size	Max Regio	Classificati	Inheritance	Gender	Group	ENS gene(s)
P_000302	chr3:14,40	CN Gain	102612	p25.1	53	101689	chr3:14,40	103535	chr3:14,40	likely delet	undetermined	F	2	SLC6A6
P_000450	chr1:185,1	CN Gain	22846	q25.3	41	22050	chr1:185,1	23641	chr1:185,1	VUS	undetermined	M	3	
P_000479	chr1:3,776,	CN Gain	272933	p36.32	57	261945	chr1:3,783,	283922	chr1:3,769,	VUS	undetermined	M	2	
P_000479	chr12:9,24	CN Gain	63052	p13.31	24	59301	chr12:9,24	66803	chr12:9,24	VUS	undetermined	M	2	A2M
P_000479	chr2:40,62	CN Loss	22235	p22.1	11	20437	chr2:40,62	24033	chr2:40,62	likely delet	undetermined	M	2	SLC8A1
P_000480	chr3:137,7	CN Loss	50512	q22.3	14	47556	chr3:137,7	53469	chr3:137,7	VUS	undetermined	M	2	
P_000490	chr2:220,2	Homozygou	1050	q35	9	480	chr2:220,2	1620	chr2:220,2	VUS	undetermined	F	3	
P_000490	chr6:49,66	Homozygou	2471	p12.3	6	93	chr6:49,66	4850	chr6:49,66	VUS	undetermined	F	3	
P_000490	chr7:117,2	Homozygou	3495	q31.2	13	110	chr7:117,2	6881	chr7:117,2	VUS	undetermined	F	3	
P_000490	chr7:117,2	Homozygou	5945	q31.2	6	59	chr7:117,2	11830	chr7:117,2	VUS	undetermined	F	3	
P_000494	chr12:128,	CN Gain	708814	q24.32	96	689879	chr12:128,	727749	chr12:128,	VUS	paternal	M	1	TMEM132C
P_000498	chr1:152,2	CN Gain	37488	q21.3	11	36431	chr1:152,2	38545	chr1:152,2	VUS	undetermined	F	3	FLG
P_000502	chr15:80,5	CN Gain	75928	q25.1	22	72813	chr15:80,5	79044	chr15:80,5	VUS	undetermined	F	2	CTXND1
P_000512	chr3:108,4	CN Loss	445452	q13.13	30	409895	chr3:108,4	481010	chr3:108,4	VUS	maternal	F	1	
P_000512	chr6:28,00	CN Loss	3678174	p22.1 p21	403	3657627	chr6:28,02	3698722	chr6:27,98	likely delet	de novo	F	1	6M1 18, ABHD16A, ATAT1, ATP6V1G2, DDR1, DPCR1, FLOT1, GABBR1, GNL1, HLA H, IER3, MUCL3, OR11A1, OR2J2, OR2J3, PGBD1, PPP1R11, PPP1R18, TUBB, ZKSCAN4, ZNRD1 AS1, ZNRD1ASP, ZSCAN31
P_000514	chr21:47,8	CN Loss	25143	q22.3	14	20293	chr21:47,8	29993	chr21:47,8	VUS	undetermined	M	3	
P_000515	chr3:57,41	CN Loss	77410	p14.3	15	74402	chr3:57,41	80417	chr3:57,41	VUS	undetermined	M	3	
P_000515	chr4:159,5	CN Loss	25325	q32.1	20	18612	chr4:159,6	32037	chr4:159,5	VUS	undetermined	M	3	
P_000520	chr18:45,7	CN Gain	31688	q21.1	13	28932	chr18:45,7	34445	chr18:45,7	VUS	undetermined	M	1	ZBTB7C
P_000537	chr10:49,0	CN Loss	3384109	q11.22 q1	183	2997727	chr10:49,4	3770491	chr10:48,6	likely delet	maternal	M	1	ARHGAP22, C10orf128, CHAT, FAM21A, MAPK8, NCOA4, SLC18A3, TIMM23, TIMM23B, TMEM273, VSTM4, WASHC2A
P_000540	chrX:154,2	Hemizygou	22055	q28	5	13897	chrX:154,2	30213	chrX:154,2	VUS	maternal	M	1	CMC4, FUNDC2, MTCP2
P_000544	chr1:25,71	CN Gain	29090	p36.11	26	28584	chr1:25,71	29596	chr1:25,71	VUS	undetermined	M	2	
P_000544	chr15:62,4	CN Gain	111807	q22.2	33	105178	chr15:62,4	118437	chr15:62,4	VUS	undetermined	M	2	
P_000552	chr18:6,33	CN Gain	95403	p11.31	20	92431	chr18:6,33	98375	chr18:6,33	VUS	undetermined	M	3	
P_000552	chr18:6,45	CN Gain	94532	p11.31	37	91889	chr18:6,45	97174	chr18:6,45	VUS	undetermined	M	3	
P_000555	chr7:15179	CN Gain	460773	q36.1	52	450810	chr7:151,8	470735	chr7:151,7	VUS	de novo	F	1	
P_000557	chr11:62,2	CN Gain	47571	q12.3	25	42335	chr11:62,2	52806	chr11:62,2	likely delet	undetermined	M	2	AHNAK
P_000557	chr2:189,8	CN Loss	23992	q32.2	53	21606	chr2:189,8	26379	chr2:189,8	VUS	undetermined	M	2	
P_000557	chr2:206,8	CN Loss	124758	q33.3	18	121507	chr2:206,8	128008	chr2:206,8	VUS	undetermined	M	2	
P_000557	chr4:159,5	CN Loss	90194	q32.1	48	86742	chr4:159,5	93646	chr4:159,5	VUS	undetermined	M	2	
P_000561	chr12:80,5	CN Loss	237894	q21.31	12	225135	chr12:80,6	250653	chr12:80,5	VUS	undetermined	M	1	
P_000561	chr22:18,8	CN Loss	2769422	q11.21	446	2584567	chr22:18,8	2954276	chr22:18,8	likely delet	de novo	M	1	ARVCF, BCRP2, C22orf29, C22orf39, CDC45, COMT, DGCR14, DGCR2, ESS2, GP1BB, LZTR1, RIMBP3, RTL10, SLC7A4, UFD1, UFD1L, ZDHHC8, ZNF74
P_000566	chr1:25,71	CN Gain	27399	p36.11	24	27102	chr1:25,71	27697	chr1:25,71	VUS	undetermined	M	2	
P_000567	chr17:58,0	CN Loss	2286148	q23.1 q23	74	2163827	chr17:58,1	2408469	chr17:57,9	likely delet	de novo	M	1	BCAS3,HEATR6, TBX2,USP32
P_000567	chr2:177,1	CN Gain	131505	q31.1	15	123759	chr2:177,1	139250	chr2:177,1	VUS	undetermined	M	1	
P_000568	chr8:95,18	CN Gain	115668	q22.1	41	111368	chr8:95,18	119967	chr8:95,18	VUS	undetermined	M	1	
P_000573	chr2:216,2	CN Loss	85157	q35	49	84491	chr2:216,2	85822	chr2:216,2	VUS	undetermined	F	1	FN1
P_000573	chrX:3,184,	CN Gain	56053	p22.33	16	55061	chrX:3,185,	57044	chrX:3,184,	VUS	undetermined	F	1	
P_000579	chr3:60,46	CN Loss	21696	p14.2	16	15424	chr3:60,47	27968	chr3:60,46	likely delet	undetermined	M	3	FHIT
P_000579	chrX:94,40	Hemizygou	37358	q21.33	5	14273	chrX:94,41	60443	chrX:94,40	VUS	undetermined	M	3	
P_000582	chr2:102,6	CN Gain	188513	q11.2 q12	57	183275	chr2:102,6	193750	chr2:102,6	VUS	undetermined	M	3	IL1R1, IL1RL2
P_000582	chr7:4,736,	CN Gain	123671	p22.1	47	120625	chr7:4,736,	126717	chr7:4,735,	VUS	undetermined	M	3	
P_000582	chr7:4,929,	CN Gain	289009	p22.1	59	283810	chr7:4,931,	294208	chr7:4,926,	VUS	undetermined	M	3	MMD2
P_000582	chr7:5,239,	CN Gain	162393	p22.1	54	160806	chr7:5,240,	163980	chr7:5,238,	VUS	undetermined	M	3	SLC29A4, WIPI2
P_001632	chr22:18,6	CN Gain	373745	q11.21	75	341266	chr22:18,7	406223	chr22:18,6	VUS; modif	undetermined	M	3	DGCR2
P_001636	chr1:25,71	CN Gain	29090	p36.11	26	28584	chr1:25,71	29596	chr1:25,71	VUS	undetermined	M	3	
P_001637	chr2:10,66	CN Gain	250389	p25.1	70	246862	chr2:10,66	253917	chr2:10,66	VUS	undetermined	M	3	NOL10

P_001639	chr6:22,00	CN Loss	84880	p22.3		32	73287	chr6:22,01	96473	chr6:22,00	VUS	undetermined	M	3	
P_001763	chr18:65,6	CN Gain	835767	q22.1		228	826179	chr18:65,7	845355	chr18:65,6	VUS	undetermined	F	1	TMX3
P_002431	chr1:243,9	CN Loss	53278	q44		9	45623	chr1:243,9	60934	chr1:243,9	likely delet	undetermined	F	3	AKT3
P_002431	chr3:145,7	CN Gain	26593	q24		9	19512	chr3:145,7	33674	chr3:145,7	VUS	undetermined	F	3	
P_002431	chr7:95,84	CN Loss	158283	q21.3		16	151029	chr7:95,84	165537	chr7:95,84	VUS	undetermined	F	3	SLC25A13
P_002450	chr9:28,39	CN Loss	69583	p21.1		21	64590	chr9:28,39	74575	chr9:28,39	VUS	undetermined	M	1	LINGO2
P_002455	chr22:21,8	CN Gain	752756	q11.21	q1	169	732827	chr22:21,8	772685	chr22:21,8	VUS	de novo	M	1	CCDC116, MAPK1, PPM1F, SDF2L1, TMEM191C, YDJC, YPEL1
P_002455	chr7:3,627,	CN Loss	132054	p22.2		32	125747	chr7:3,627,	138361	chr7:3,626,	VUS	undetermined	M	1	
P_002459	chr12:94,7	CN Loss	112786	q22		27	110549	chr12:94,7	115022	chr12:94,7	VUS	undetermined	M	1	CCDC41, CEP83

Supplementary Table

Gene	NVAR	NCASEHET	NCTRLHET	PVALUE
MED13	11	1	11	0,013323
CDC45	7	1	8	0,016175
TNRC18	33	28	15	0,022411
INTS2	8	1	7	0,022945
HEATR6	12	3	11	0,039536
NOL10	5	4	1	0,082978
TBX1	2	3	0	0,105246
TBX4	3	1	3	0,113098
TMEM132C	8	6	2	0,141436
ZSCAN31	3	4	1	0,141762
SLC6A6	2	2	0	0,147457
SGMS1	5	5	1	0,147487
COL3A1	4	4	1	0,164587
INO80D	5	1	4	0,170651
ARHGAP22	10	4	7	0,195018
GALNT11	9	4	9	0,198599
RIMBP3	7	7	3	0,223134
BCAS3	7	5	2	0,226416
MMD2	2	1	3	0,230996
MAPK8	2	1	4	0,258441
LZTR1	11	7	5	0,266329
APPBP2	2	0	2	0,268345
YPEL1	1	0	1	0,27441
DGCR8	4	3	1	0,276763
CEP83	2	0	2	0,27776
SLC8A1	4	1	3	0,278436
FLG	39	22	19	0,278939
SDF2L1	1	1	0	0,281767
CLDN5	1	0	1	0,287018
C22orf39	1	1	0	0,312221
VSTM4	4	1	3	0,333988
MED15	1	1	0	0,341908
PGBD1	4	1	3	0,342391
FAM21A	3	1	3	0,358184
USP32	10	8	5	0,364467
SCARF2	4	3	1	0,387793
DGCR14	6	4	2	0,392625
ARVCF	22	9	16	0,393638
PPM1F	6	4	2	0,394534
A2M	3	2	1	0,403244
GP1BB	3	2	1	0,429398
RXFP1	3	2	1	0,43508
UFD1L	1	0	1	0,4355
IL1RL2	4	1	3	0,451528
WDFY4	30	18	14	0,455625
SEPT5	1	0	1	0,463793
ZNF74	4	3	2	0,470632
CHAT	4	3	2	0,477867
FN1	16	10	8	0,498278

Supplementary Table cont.

TBX2	4	3	2	0,498533
WIPI2	3	1	3	0,53723
SLC25A13	14	11	7	0,56787
SLC7A4	10	6	6	0,599833
CLTCL1	24	14	14	0,604724
ZBTB7C	4	3	2	0,610935
C22orf29	1	2	3	0,623266
HIC2	5	2	3	0,625334
CA4	6	3	3	0,669649
SLC29A4	9	5	5	0,682379
ZDHHC8	11	7	6	0,699988
AHNAK	64	58	47	0,704864
PI4KA	14	6	9	0,719658
CCDC116	4	2	3	0,744967
ZKSCAN4	2	1	1	0,757303
RTN4R	2	1	1	0,76249
HIRA	8	5	7	0,782911
DGCR2	2	1	1	0,801046
C10orf128	2	1	1	0,804269
TMX3	4	2	2	0,821874
YDJC	2	1	1	0,841144
AKT3	2	1	1	0,868532
SWT1	2	1	1	0,90647
COMT	5	3	3	0,939442
LINGO2	9	5	6	0,941814
NXF1	2	1	1	0,943867
EPHA5	4	2	2	0,948386

Supplementary Table

Genesymbol	Patient number	Genesymbol	CNV Type	Mouse orthologue	Mouse ENS E11 E15.5 gene	EW12_TPM	EW12_CPM	EW14_TPM	EW14_CPM	EW16_TPM	EW16_CPM	ENS_gene	mis_z	syn_z	pLI	pRec	pNull	del	dup	del.sing	dup.sing	del.sing.scc	dup.sing.sc	del.score	dup.score	cnv.score	flag	Variant_int	Loss intoler	Gain intole	Intolerant_	ENS_intolerant_genes
AHNAK	P_000557	AHNAK	CN Gain	Ahnak	YES	121,949122	717,552172	143,8831	800,3591	218,8149	1287,816	YES	2,6889	1,2365	0,90114	0,098859	2,61E 16	0		1	0	0	0,04906	0,129105	0,234312	0,295019	0,398553	0	YES	.	.	YES	YES
AKT3	P_002431	AKT3	CN Loss	Akt3	YES	9,98360131	22,5992908	13,72781	29,34257	16,45234	37,22031	YES	4,0256	0,10227	0,99956	0,000442	2,66E 10	5		3	5	3	2,60549	0,93311	1,2035	0,105106	0,43328	0	YES	.	.	YES	YES
APPBP2	P_000567	APPBP2	CN Loss	Appbp2	NO	18,7210697	55,7433199	19,96505	56,17531	23,74236	70,66897	.	3,2772	0,14954	0,99657	0,003425	3,24E 10	0		5	0	1	1,081554	0,300336	1,384406	0,050871	0,549403	1	YES	.	.	YES	.
BAG6	P_000512	BAG6	CN Loss	Bag6	NO	129,110888	207,730503	127,9784	194,6738	123,2605	198,4574	.	2,6849	0,72379	0,99999	9,08E 06	1,71E 16	0		1	0	1	1,337149	0,533509	1,281018	1,131581	1,424271	0	YES	YES	.	YES	.
C6orf136	P_000512	C6orf136	CN Loss	2310061I04Rik	NO	31,9310967	20,4068481	29,03775	17,56561	25,90071	16,57012	.	1,2978	1,1867	0,9524	0,047593	5,14E 06	0		2	0	2	0,910147	0,74802	1,321416	0,75705	1,126391	0	YES	YES	.	YES	.
CLDN5	P_000561	CLDN5	CN Loss	Cldn5	NO	23,6061131	24,4276291	39,3693	38,56095	16,53239	17,1322	.	1,5995	0,86061	0,90443	0,094849	0,000725	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
COL3A1	P_000557	COL3A1	CN Loss	Col3a1	NO	2063,80386	4245,04413	4323,169	8406,266	2908,534	5984,314	.	4,4011	0,59541	1	6,99E 13	8,04E 34	11		23	9	19	2,84641	2,87279	1,90399	1,90913	1,72864	0	YES	.	.	YES	.
CSNK2B	P_000512	CSNK2B	CN Loss	Csnk2b	NO	109,390923	97,4070387	109,5618	92,30638	87,14019	77,66748	.	3,0895	0,77356	0,97899	0,02099	1,64E 05	0		0	0	0	0,05809	0,054358	0,286863	0,70718	0,774563	0	YES	.	.	YES	.
DDR1	P_000512	DDR1	CN Loss	Ddr1	YES	61,7782928	170,046612	43,98524	114,4218	78,50281	216,2703	YES	2,3529	1,9042	5,18E 06	0,99998	1,21E 05	2		6	0	0	1,022847	0,978386	0,097256	0,34321	0,18854	0	.	YES	.	YES	YES
DDX39B	P_000512	DDX39B	CN Loss	Ddx39b	NO	117,269991	253,434949	109,9179	224,7495	68,85175	148,8791	.	4,6354	1,0264	0,99748	0,002518	1,85E 08	0		1	0	0	0,632489	0,637405	0,825798	0,735237	0,971055	0	YES	.	.	YES	.
DGCR8	P_000561	DGCR8	CN Loss	Dgcr8	NO	31,1004312	68,6520889	30,80723	64,29783	20,7316	45,76406	.	3,058	0,31842	0,99967	0,000327	5,23E 12	13		33	0	0	1,122676	1,048404	2,46581	2,53125	2,47163	1	YES	.	.	YES	.
DHX16	P_000512	DHX16	CN Loss	Dhx16	NO	39,0514425	47,2182602	39,53798	45,19012	33,21936	40,216	.	3,1491	1,369	7,97E 13	0,99999	1,41E 05	1		0	1	0	0,07281	1,033105	0,436915	1,272042	1,133005	0	YES	.	.	YES	.
FHIT	P_000579	FHIT	CN Loss	Fhit	YES	8,31448977	13,6433165	6,418951	9,97824	3,805209	6,24057	YES	1,139	1,2841	0,029167	0,81428	0,15656	1		5	0	0	1,525674	1,157754	0,445892	0,55642	0,2134	0	.	YES	.	YES	YES
GABBR1	P_000512	GABBR1	CN Loss	Gabbr1	YES	14,3044388	36,5563308	19,44168	47,00717	17,80889	45,55914	YES	4,9802	1,0093	0,99989	0,000113	3,22E 15	1		0	0	0	1,360855	1,421753	1,229685	1,998958	1,904234	0	YES	YES	.	YES	YES
GNL1	P_000512	GNL1	CN Loss	Gnl1	YES	32,3563173	84,4317419	35,2716	87,02605	32,38833	84,54276	YES	2,5212	0,70573	0,99922	0,000784	4,41E 11	1		0	0	0	1,02519	0,914214	0,696838	1,418668	1,34504	0	YES	YES	.	YES	YES
HIC2	P_002455	HIC2	CN Gain	Hic2	NO	39,7507474	86,5831207	23,68168	48,76105	36,81726	80,24758	.	2,1152	2,2164	0,99679	0,003211	1,65E 07	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
HIRA	P_000561	HIRA	CN Loss	Hira	NO	56,4488058	75,4825433	40,04365	50,62276	43,23488	57,8455	.	3,8363	0,33966	1	1,24E 07	2,41E 20	13		34	0	0	1,408548	1,51161	2,51548	2,53125	2,47163	1	YES	.	.	YES	.
INO80D	P_000557	INO80D	CN Loss	Ino80d	NO	9,35712602	40,5861369	9,07603	37,23781	9,811772	42,53822	.	2,9131	0,11631	0,99958	0,000421	1,96E 12	0		0	0	0	0,085323	0,190245	0,348871	0,728334	0,816173	0	YES	.	.	YES	.
INTS2	P_000567	INTS2	CN Loss	Ints2	NO	8,40368443	18,0816483	7,214446	14,67664	6,075531	13,06234	.	2,4922	1,7744	0,99701	0,002991	1,6E 14	1		4	1	0	0,347388	1,353848	0,841984	0,350435	0,580853	1	YES	.	.	YES	.
MAPK1	P_002455	MAPK1	CN Gain	Mapk1	YES	34,22101	117,065824	34,9678	113,043	43,44961	148,6445	YES	3,7127	0,37659	0,9973	0,002698	1,08E 07	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	YES
MAPK8	P_000537	MAPK8	CN Loss	Mapk8	YES	11,1110256	32,229146	12,85747	35,28936	12,83318	37,20982	YES	2,9224	0,45067	0,99827	0,001726	7,36E 09	9		8	0	0	0,835657	0,907072	2,24621	0,91348	1,3964	0	YES	.	.	YES	YES
MED13	P_000567	MED13	CN Loss	Med13	NO	23,7316747	78,9650857	25,13972	79,09438	27,21403	90,56939	.	2,8161	1,2723	1	1,98E 16	1,09E 39	0		5	0	2	1,747769	0,318278	1,735003	0,40757	0,935395	1	YES	.	.	YES	.
MED15	P_000561	MED15	CN Loss	Med15	NO	26,6238312	83,3388482	25,97913	76,86108	17,81748	55,79005	.	2,6169	0,31626	0,99962	0,000379	1,26E 14	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
NXF1	P_000568	NXF1	CN Gain	Nxf1	NO	52,6249822	98,3041638	46,87168	82,78065	39,8716	74,48067	.	2,6711	0,53229	0,98814	0,011863	2,77E 10	0		4	0	1	1,342514	0,548133	1,328249	0,280578	0,728529	0	YES	.	.	YES	.
PI4KA	P_000561	PI4KA	CN Loss	Pi4ka	NO	37,6765045	129,693522	29,86899	97,19763	31,58556	108,8127	.	3,8158	0,074324	1,95E 11	1	3,33E 15	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
POU5F1	P_000512	POU5F1	CN Loss	Pou5f1	NO	0,02746929	0,03926764	0,126856	0,170907	0,115137	0,165357	.	1,9285	1,8986	0,98254	0,017456	2,11E 06	0		0	0	0	0,31784	0,17184	0,085038	0,557813	0,571258	0	YES	.	.	YES	.
PPP1R10	P_000512	PPP1R10	CN Loss	Ppp1r10	NO	51,3022114	80,7313282	48,28542	71,83158	43,23828	68,08656	.	3,0103	0,54528	1	1,74E 06	1,51E 17	2		1	2	0	0,43384	1,191345	0,334479	1,137744	0,94151	0	YES	.	.	YES	.
PRRC2A	P_000512	PRRC2A	CN Loss	Prrc2a	NO	95,5524989	262,783873	87,77293	228,2009	78,36644	215,6328	.	0,23086	0,23638	1	5E 11	1,82E 33	11		23	11	22	2,84641	2,87279	2,17051	2,15057	1,97157	0	YES	.	.	YES	.
RTN4R	P_000561	RTN4R	CN Loss	Rtn4r	NO	8,75218903	8,82137973	9,532229	9,075124	12,53104	12,62657	.	1,2449	0,28375	0,95915	0,04077	8,52E 05	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
SCARF2	P_000561	SCARF2	CN Loss	Calm5	NO	24,4809878	26,6094065	31,52783	32,38189	19,3091	21,02184	.	3,1636	2,8721	0,99973	0,000272	1,64E 11	15		46	0	0	1,001017	0,958605	1,42334	1,93669	1,61174	1	YES	.	.	YES	.
SEPT5	P_000561	SEPT5	CN Loss	.	.	16,9297424	30,2192401	20,92352	35,28224	22,31379	39,85699	.	2,7529	0,52611	0,92863	0,071367	2,89E 06	12		33	0	0	0,52758	0,625092	2,62354	2,53125	2,47163	1	YES	.	.	YES	.
SGMS1	P_000537	SGMS1	CN Loss	Sgms1	NO	14,9712264	19,3519225	16,81085	20,54514	20,70387	26,76167	.	2,7237	0,16909	0,93736	0,062635	2,08E 06	2		65	0	0	0,424344	0,588971	0,20088	2,53125	2,47163	1	YES	.	.	YES	.
SLC6A6	P_000302	SLC6A6	CN Gain	Slc6a6	YES	16,6002021	41,3832211	14,64367	34,49	5,641203	14,08104	YES	3,4669	0,18998	0,98494	0,015062	1,22E 08	1		2	0	1	1,228165	0,388763	1,042711	1,015823	1,1606	0	YES	.	YES	YES	YES
SLC8A1	P_000479	SLC8A1	CN Loss	Slc8a1	YES	13,4810762	28,7513924	17,03655	34,35355	28,29569	60,36847	YES	2,2298	3,1149	0,99605	0,003952	4,6E 10	1		6	1	3	0,0152	1,05906	0,31223	0,6599	0,34997	0	YES	.	.	YES	YES
SWT1	P_000450	SWT1	CN Gain	Swt1	NO	5,58815574	6,83736006	6,375164	7,369928	8,475753	10,37314	.	1,5742	0,4418	0,99949	0,000507	6,21E 13	58		42	58	38	2,84641	2,87279	2,62354	2,53125	2,47163	0	YES	.	.	YES	.
TBX2	P_000567	TBX2	CN Loss	Tbx2	YES	18,55854	33,7672082	25,49512	43,84805	12,50785	22,76906	YES	1,7515	1,3708	0,99134	0,008655	3,77E 07	0		0	0	0	0,01215	0,085138	0,531942	0,92725	0,988479	0	YES	.	.	YES	YES
TNRC18	P_000582	TNRC18	CN Gain	Tnrc18	NO	67,1905664	254,898274	59,53933	213,5033	57,84216	219,6266	.	0,10479	6,4425	0,99661	0,003394	7,38E 18	2		2	1	0	0,113032	1,179455	0,388382	1,087289	0,838203	0	YES	.	YES	YES	.
TRIM26	P_000512	TRIM26	CN Loss	Trim26	NO	56,7911705	75,0134458	54,70279	68,28719	57,34557	75,79019	.	3,1729	1,0883	0,19022	0,80976	1,75E 05	0		0	0	0	0,291646	0,352274	0,787697	1,108101	1,242702	0	YES	.	.	YES	.
TRIM27	P_000512	TRIM27	CN Loss	Trim27	NO	41,2208551	99,3383234	36,45532	83,02299	31,93536	76,99771	.	3,8091	0,15942	0,99758	0,002415	3,38E 09	1		3	1	2	0,57752	1,09945	0,077547	0,06005	0,022875	0	YES	.	.	YES	.
TRIM39	P_000512	TRIM39	CN Loss	Trim39	NO	12,3581488	15,8473668	12,394	15,01226	11,99055	15,3932	.	3,9127	1,6778	0,99395	0,006048	6,43E 09	2		0	0	0	0,330481	0,418135	0,21304	1,128721	0,663071	0	YES	.	.	YES	.
TUBB	P_000512	TUBB	CN Loss	Tubb2a	YES	1023,01729	1036,93414	1273,477	1220,576	1043,167	1058,009	YES	5,7068	0,00984	0,98045	0,019549	2,79E 06	5		12	5	12	2,84641	2,87279	1,71507	1,92232	1,9144	0	YES	.	.	YES	YES
UFD1L	P_000561	UFD1L	CN Loss	Ufd1	YES	13,4874668	35,4337574	16,63382	41,31418	12,90988	33,94235	YES	2,767	1,1387	0,9965	0,003495	4,11E 08	13		34	0	0	1,057765	1,018976	2,52508	2,53125	2,47163	1	YES	.	.	YES	YES
USP32	P_000567	USP32	CN Loss	Usp32	YES	9,64355777	22,5704017	10,1056	22,33813	10,26487	24,03409	YES	3,5546	2,2783	1	1,19E 08	5,47E 27	9		8	9	4	2,84641	0,34655	0,92535	0,292861	0,08835	1	YES	.	.	YES	YES
WDFY4	P_000537	WDFY4	CN Loss	Wdfy4	NO	1,22895039	5,17656992	2,355761	9,380276	0,758482	3,194854	.	3,3374	1,8891	0,6551	0,3449	1,67E 24	.	.	.	.	.	.	.	.	.	.	YES	.	.	YES	.
YDJC	P_002455	YDJC	CN Gain	Ydjc	YES	72,0801259	38,1504871	56,8613	28,47175	43,24999	22,90299	YES	1,5775	2,2537	0,060829	0,87436	0,064811	0		3	0	0	0,390439	0,423384	1,812725	1,253831	1,416572	0	.	.	YES	YES	YES

Supplementary Table

Name	Purpose	Primer sequence
rs2506030_1FW	Sanger sequencing	GGAGGCGCATCCCTAGCC
rs2506030_1RV	Sanger sequencing	CAGCTACTCTGGGGCCTTGC
rs7069590_3FW	Sanger sequencing	CAATGGTCCCCTGCACACC
rs7069590_3RV	Sanger sequencing	GCACCCCTGGCAGTGACC
rs2505998 _1FW	Sanger sequencing	TGGCTACCTAGGCTACACACTCAGG
rs2505998 _1RV	Sanger sequencing	CCCCCAGACCTTTTTCCAAGG
rs2435357_1FW	Sanger sequencing	CAGCTGCTGCAGAGTTAATCACC
rs2435357_1RV	Sanger sequencing	AGAGGCACCAGGGTCAAAGC
rs9282834_1FW	Sanger sequencing	GTCCATGCCTTCCCCACTCC
rs9282834_1RV	Sanger sequencing	GGGAAAGTCTGTGTGGAAAACTGC
rs11766001_1FW	Sanger sequencing	CAATCAAAATGCAAGACACCATTAGC
rs11766001_1RV	Sanger sequencing	TGAAAGATGATGGTGTGGATGAGC
rs80227144 _1FW	Sanger sequencing	GGGCAGATGGATATGTAGGC
rs80227144 _1RV	Sanger sequencing	TTGAATAAAATGTCTTATTGTTTTCC
rs7005606 _2FW	Sanger sequencing	TCTGCACCATAATTACAGCAATGG
rs7005606 _2RV	Sanger sequencing	TGGAGGGTACCACTTCTAGTTTTGC

Supplementary table 6; odds ratio's used to calculate polygenic risk scores

Gene	risk/non-risk allele	Risk allele frequency	Odds ratio (95% CI)	P	OR used in PGRS	Ref OR used
RET	rs2506030: G/A	0.56/0.41	1.8 (1.5-2.2)	4.46×10-11	1.8	[1]
RET	rs7069590: T/C	0.84/0.76	1.7 (1.4-2.2)	4.36×10-6	1.7	[1]
RET	rs2505998: A/C	0.64/0.22	4.17 (3.23–5.26)	1.1×10-28	4.17	[2]
RET	rs2435357: T/C	0.58/0.25	4.01 (3.33–4.84)	2.98×10-48	4.01	[1]
RET	rs9282834: A/G	0.05/0.03	1.80 (1.06–3.04)	0.029	1.8	[2]
SEMA3	rs11766001:C/A	0.22/0.15	1.6 (1.3–2.0)	1. 1	1.6	[3]
SEMA3	rs80227144: C/A	0.14/0.03	5.2 (3.09–8.73)	4.0×10-4	5.2	
NRG1	rs7005606: T/G	0.54/0.42	1.64 (1.25–2.15)	4.7×10-10	1.64	[2]

Gene	Risk haplotype	Risk haplotype frequency	Odds ratio (95% CI)	P	OR used in PGRS	Ref OR used
RET	ATT#	0.14/0.08	3.13 (2.17-4.50)	8.31 1 1	not used	[1]
RET	GTT#	0.42/0.16	4.40 (3.26-5.94)	3.62 1 22	not used	[1]
RET	TA@	0.05/0.03	20.3 (9.31–44.4)	2.98×10-48	not used	[2]

Risk alleles and Odds ratio's used in the polygenic risk score calculation. #We did not account for the increased risk of having the two main risk haplotype combinations[1] # this risk haplotype was not present in this cohort

References

1. Chatterjee, S., et al., *Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease.* Cell, 2016. **167**(2): p. 355-368.e10.
2. Tang, C.S., et al., *Trans-ethnic meta-analysis of genome-wide association studies for Hirschsprung disease.* Hum Mol Genet, 2016. **25**(23): p. 5265-5275.
3. Kapoor, A., et al., *Population variation in total genetic risk of Hirschsprung disease from common RET, SEMA3 and NRG1 susceptibility polymorphisms.* Hum Mol Genet, 2015. **24**(10): p. 2997-3003.

Supplementary table 7 Hirschsprung patients without a RET mutation or other disease causing mutations and without other associated phenotypes

Control group 1: Isolated Hirschsprung Disease

Patient	subgroup	HSCR type
P_000577	isolated	Short
P_001632	Isolated	Short
P_001635	Isolated	Short
P_001636	isolated	Short
P_001637	isolated	Short
P_001638	Isolated	Short
P_001639	isolated	Short
P_000498	Isolated	Long
P_000575	Isolated	Short
P_000514	isolated	Short
P_000450	isolated	Long
P_000582	isolated	Short
P_000578	isolated	Long
P_000505	isolated	Short
P_000490	isolated	Short
P_000552	isolated	Short
P_000554	isolated	Short
P_000579	isolated	Short
P_000515	isolated	Short
P_002431	isolated	Short

Supplementary Table cont.

Control group 2: HSCR patients with a deleterious variant

Patient	HSCR type	Other phenotypical characteristics	Genetic defect
P_000302	Short	-	IHH:NM_002181.3:c151C>A, GLI3:NM_000168.5:c2119C>T, GDNF:NM_001190468.1:c676_681delGGATGT
P_000526	Short	-	RET:NM_020630:exon?:c.1196C>T
P_000479	Long	-	RET:NM_020630:exon?:c.656-21C>T in intron 3 Ret gen
P_000502	Short	-	RET:NM_020630:exon11:c.1880_1892del:p.D627Afs*6
P_002442	Long	-	c.2599C>T (p.Glu867X) in ex 14 RET gen
P_000566	Short	-	RET:NM_020630:exon19:c.A3173G:p.E1058G pat
P_000544	Long	-	RET:NM_020630:exon15:c.G2690A:p.R897Q
P_000534	Short	Anteriorly placed anus	*RET: c.2906G>A*:p.R969Q
P_000480	Short	Short stature	*NRG1*::c.811C>T, *p.R271W*
P_004502	Short	Hypertelorism, triangular face, pointy chin, straight eye brows, deepset eyes, small dysmorphic ears, agenesis of corpus callosum, hypospadia, dysmorphic nose	ZEB2; c.1570del/p.Ser524fs
P_000557	Total colonic	postaxial polydactyly	RET:NM_020630:exon2:c. c.C229T :p.Arg77Cys de novo
P_000518	Short	Facial dysmorphisms, microcephaly, bilateral generalized polymicogyria, developmental elay, short stature, hypotonia, eye anomaly	R90X(nucleotideverandering 268C>T) in het KIAA1279 gen. c.268C>T, p.Arg90* (KIAA1279, exon 01)
P_000486	Total colonic	Waardenburg syndrome? No abnormal phenotypedescribed. Normal psychomotor development	EDNRB:NM_001122659:exon2:c.534_535insGGTGCCT:p.F179
P_000570	Short	congenital centraal hypoventilatie syndroom. Tevens heeft zij de ziekte van Hirschsprung.	Phox2B in exon 3 (c.738_761dup PHOX2b) (
P_000576	Short	Microcephaly, epicantus folds, upslant eyes, broad nose, synophrys naevus sacralis hyperpigmentosis back and shoulders, abnormal palmar creases	RET:NM_020630:exon?:c.A1321C:p.Lys441Gln and RET:NM_020630:exon?:c.C1941T (inheritance unspecified)

Supplementary Table

Genesymbol	Patient number	CNV Type	Mouse orthologue	Mouse ENS E11 E15.5 gene	EW12_TPM	EW12_CPM	EW14_TPM	EW14_CPM	EW16_TPM	EW16_CPM	ENS_gene	mis_z	syn_z	pLI	pRec	pNull	del	dup	del.sing	dup.sing	del.sing.sc	dup.sing.sc	del.score	dup.score	cnv.score	flag	Variant_int	Loss intole	Gain intole	Intolerant_	ENS_intolerant_genes
SLC6A6	P_000302	CN Gain	Slc6a6	YES	16,6002021	41,3832211	14,64367	34,49	5,641203	14,08104	YES	3,4669	0,18998	0,98494	0,015062	1,22E 08	1		2	0	1	1,228165	0,388763	1,042711	1,015823	1,1606	0	YES		YES	YES	YES
A2M	P_000479	CN Gain	A2m	YES	89,4899495	172,413533	171,7171	312,7306	98,10483	189,0585	YES	2,0872	1,8665	5,61E 11	1	6,36E 09	0		3	0	2	1,563109	0,229215	1,303954	0,571732	0,944284	0					
SLC8A1	P_000479	CN Loss	Slc8a1	YES	13,4810762	28,7513924	17,03655	34,35355	28,29569	60,36847	YES	2,2298	3,1149	0,99605	0,003952	4,6E 10	1		6	1	3	0,0152	1,05906	0,31223	0,6599	0,34997	0	YES			YES	YES
TMEM132C	P_000494	CN Gain	Tmem132c	YES	2,53958554	3,78982097	4,0012	5,654928	3,769971	5,628447	YES	1,9672	0,91297	0,63788	0,36212	4,35E 07	.	.	.	.	.	.	.	.	.		.	.	.	.	.
FLG	P_000498	CN Gain	Fgfr1	YES	0,03176246	0,12182404	0,050848	0,184859	0,036953	0,141293	YES	17,639	19,143	0,000312	0,21024	0,78945	.	.	.	.	.	.	.	.	.		.	.	.	.	.
CTXND1	P_000502	CN Gain	Gm2115	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
6M1 18	P_000512	CN Loss	Olfr96	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
ABHD16A	P_000512	CN Loss	Abhd16a	YES	45,057204	56,4502743	40,6746	48,20835	57,31949	71,86699	YES	2,5085	0,95092	0,30441	0,69559	2,52E 07	7		1	0	0	0,947791	1,019841	1,72028	0,877555	0,31185	0					
ATAT1	P_000512	CN Loss	Atat1	YES	22,0214445	26,8073863	22,71187	26,12722	11,5945	14,13506	YES	1,5422	1,2362	1,03E 07	0,98409	0,015913	1		4	1	0	0,30594	0,698546	0,186716	0,17517	0,04016	0					
ATP6V1G2	P_000512	CN Loss	Atp6v1g2	YES	5,0708905	3,28553489	9,984063	6,114678	13,71048	8,89412	YES	0,87581	0,72551	0,003097	0,82919	0,16771	0		0	0	0	0,020954	0,078233	0,350708	0,746561	0,820699	0					
DDR1	P_000512	CN Loss	Ddr1	YES	61,7782928	170,046612	43,98524	114,4218	78,50281	216,2703	YES	2,3529	1,9042	5,18E 06	0,99998	1,21E 05	2		6	0	0	1,022847	0,978386	0,097256	0,34321	0,18854	0		YES		YES	YES
DPCR1	P_000512	CN Loss	Mucl3	YES	0,02148897	0,03432969	0,009014	0,013464	0,040384	0,064975	YES	2,5404	3,9058	1,51E 05	0,99843	0,001551	0		6	0	0	0,94017	0,78456	0,41361	1,73355	1,51652	0					
FLOT1	P_000512	CN Loss	Flot1	YES	115,93708	111,470929	134,4491	122,2216	123,9001	119,1749	YES	2,3971	2,0478	1,58E 08	0,95572	0,044278	2		17	2	3	0,61739	1,08341	0,28246	1,79079	1,20926	0					
GABBR1	P_000512	CN Loss	Gabbr1	YES	14,3044388	36,5563308	19,44168	47,00717	17,80889	45,55914	YES	4,9802	1,0093	0,99989	0,000113	3,22E 15	1		0	0	0	1,360855	1,421753	1,229685	1,998958	1,904234	0	YES	YES		YES	YES
GNL1	P_000512	CN Loss	Gnl1	YES	32,3563173	84,4317419	35,2716	87,02605	32,38833	84,54276	YES	2,5212	0,70573	0,99922	0,000784	4,41E 11	1		0	0	0	1,02519	0,914214	0,696838	1,418668	1,34504	0	YES	YES		YES	YES
HLA H	P_000512	CN Loss	Hfe	YES	161,511233	89,5270484	27,72585	14,55688	54,23698	30,07962	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
IER3	P_000512	CN Loss	Ier3	YES	22,1225878	9,05355829	58,75149	22,74148	46,96551	19,221	YES	1,3077	1,1836	0,63556	0,33787	0,026568	0		4	0	0	0,22991	0,17069	0,178249	0,69833	0,46253	0					
MUCL3	P_000512	CN Loss	Mucl3	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
OR11A1	P_000512	CN Loss	Olfr96	YES	0	0	0	0	0	0	YES	0,32758	0,26585	0,002805	0,58641	0,41078	1		0	0	0	0,41267	0,29623	0,46605	0,395931	0,132173	0					
OR2J2	P_000512	CN Loss	Olfr137	YES	0	0	0	0	0	0	YES	0,19614	0,4006	0,021788	0,77031	0,2079	.	.	.	.	.	.	.	.	.		.	.	.	.	.
OR2J3	P_000512	CN Loss	Olfr137	YES	0	0	0,052374	0,013951	0	0	YES	0,25204	0,22958	0,011604	0,85456	0,13384	6		8	0	0	1,019768	0,73352	0,48917	0,2411	0,30993	1					
PGBD1	P_000512	CN Loss	Pgbd1	YES	8,71891841	8,24746772	11,04933	9,881491	9,481974	8,966505	YES	0,99961	1,4825	5,69E 10	0,95653	0,043472	1		1	1	1	0,75204	0,62423	0,02217	0,440979	0,360283	0					
PPP1R11	P_000512	CN Loss	Ppp1r11	YES	51,5367334	39,4671635	50,15439	36,3053	67,23282	51,50022	YES	1,8259	0,95184	0,37541	0,58161	0,042985	1		1	0	0	0,326175	0,33765	0,343871	0,804354	0,738122	0					
PPP1R18	P_000512	CN Loss	Ppp1r18	YES	21,5910717	29,9638091	33,73464	44,2379	20,15498	27,97517	YES	1,0921	1,2706	0,000401	0,99792	0,001679	0		0	0	0	0,12606	0,08284	0,257902	0,630086	0,698196	0					
TUBB	P_000512	CN Loss	Tubb2a	YES	1023,01729	1036,93414	1273,477	1220,576	1043,167	1058,009	YES	5,7068	0,00984	0,98045	0,019549	2,79E 06	5		12	5	12	2,84641	2,87279	1,71507	1,92232	1,9144	0	YES			YES	YES
ZKSCAN4	P_000512	CN Loss	Zkscan4	YES	11,1784928	8,18799829	13,44443	9,295878	15,29835	11,22165	YES	1,1749	1,385	7,78E 05	0,98287	0,017057	1		0	0	0	0,228128	0,231449	0,03168	0,876728	0,698965	0					
ZNRD1 AS1	P_000512	CN Loss	4930564C03Rik	YES	3,2229443	18,9998799	2,154027	11,99638	1,454748	8,581442	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
ZNRD1ASP	P_000512	CN Loss	4930564C03Rik	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
ZSCAN31	P_000512	CN Loss	Zkscan2	YES	.	.	.	.	.	.	YES	0,18994	0,36549	4,82E 08	0,62186	0,37814	9		1	9	1	2,84641	0,014617	1,46842	1,321569	0,015154	0					
ZBTB7C	P_000520	CN Gain	Zbtb7c	YES	1,54394454	3,56425789	1,948362	4,260125	8,843397	20,41549	YES	1,3549	0,099726	0,43087	0,56777	0,001362	.	.	.	.	.	.	.	.	.		.	.	.	.	.
ARHGAP22	P_000537	CN Loss	Arhgap22	YES	1,86853723	3,82551906	1,721007	3,332757	1,614208	3,307688	YES	0,71178	0,56246	1,89E 08	0,96483	0,035172	9		8	0	0	0,564583	0,639477	2,47259	1,07043	1,6158	0					
C10orf128	P_000537	CN Loss	Tmem273	YES	2,30317318	2,86330648	4,695011	5,513037	1,92888	2,400195	YES	0,7349	0,2993	0,006449	0,75985	0,2337	9		11	0	2	0,253213	1,12085	2,62354	1,63713	1,99255	0					
CHAT	P_000537	CN Loss	Chat	YES	0,5608568	0,62537152	0,559517	0,5876	0,730879	0,812196	YES	0,39532	1,466	4,86E 11	0,94139	0,058612	9		8	0	0	1,164134	1,177045	1,74239	0,4349	0,88968	1					
FAM21A	P_000537	CN Loss	Washc2	YES	59,476606	92,6502131	32,32941	47,61106	52,45521	81,72932	YES	0,62924	0,94611	0,000318	0,98837	0,011307	.	.	.	.	.	.	.	.	.		.	.	.	.	.
MAPK8	P_000537	CN Loss	Mapk8	YES	11,1110256	32,229146	12,85747	35,28936	12,83318	37,20982	YES	2,9224	0,45067	0,99827	0,001726	7,36E 09	9		8	0	0	0,835657	0,907072	2,24621	0,91348	1,3964	0	YES			YES	YES
NCOA4	P_000537	CN Loss	Ncoa4	YES	131,22476	166,184635	136,0625	162,9178	168,153	212,926	YES	0,8603	0,12173	2,58E 06	0,9992	0,000801	11		26	1	0	0,333634	1,192302	0,76341	0,92684	0,75483	1					
SLC18A3	P_000537	CN Loss	Slc18a3	YES	5,20299969	3,80028676	8,499963	5,844219	10,52375	7,695024	YES	0,06671	2,4128	0,4593	0,53538	0,005321	8		7	0	0	0,30889	0,229	2,61082	1,30775	1,94703	1					
TIMM23	P_000537	CN Loss	Timm23	YES	57,4714935	43,0152603	59,79142	42,30748	56,81521	42,52308	YES	0,81887	0,70801	0,13936	0,78559	0,075052	.	.	.	.	.	.	.	.	.		.	.	.	.	.
TIMM23B	P_000537	CN Loss	Timm23	YES	7,6078945	3,16857718	7,752038	3,057628	5,80349	2,419286	YES	0,51539	0,80523	0,12538	0,62617	0,24845	.	.	.	.	.	.	.	.	.		.	.	.	.	.
TMEM273	P_000537	CN Loss	Tmem273	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
VSTM4	P_000537	CN Loss	Vstm4	YES	8,24792318	10,8995449	13,86466	17,2925	9,219797	12,19122	YES	0,3106	0,33551	0,14212	0,85422	0,003664	.	.	.	.	.	.	.	.	.		.	.	.	.	.
WASHC2A	P_000537	CN Loss	Washc2	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
CMC4	P_000540	Homozygous Copy	Cmc4	YES	.	.	.	.	.	.	YES	0,19276	0,68013	0,53303	0,41164	0,055329	.	.	.	.	.	.	.	.	.		.	.	.	.	.
FUNDC2	P_000540	Homozygous Copy	Fundc2	YES	16,5556968	51,2251122	23,04253	67,39093	20,80147	64,37172	YES	0,48069	0,82104	0,066351	0,87615	0,057495	.	.	.	.	.	.	.	.	.		.	.	.	.	.
MTCP1	P_000540	Homozygous Copy	Cmc4	YES	1,81556096	1,63818784	1,359117	1,161806	0,681888	0,616746	YES	1,6886	1,382	0,48813	0,49224	0,019638	.	.	.	.	.	.	.	.	.		.	.	.	.	.
AHNAK	P_000557	CN Gain	Ahnak	YES	121,949122	717,552172	143,8831	800,3591	218,8149	1287,816	YES	2,6889	1,2365	0,90114	0,098859	2,61E 16	0		1	0	0	0,04906	0,129105	0,234312	0,295019	0,398553	0	YES			YES	YES
ARVCF	P_000561	CN Loss	Arvcf	YES	17,216227	29,2062691	17,93627	28,76877	15,35844	26,04585	YES	0,05568	1,0011	2,69E 10	0,99427	0,005735	12		34	0	1	0,793055	0,093225	2,62354	2,53125	2,47163	1					
BCRP2	P_000561	CN Loss	Dcaf5	YES	1,07358655	0,87029683	0,018244	0,013951	0,073027	0,059625	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
C22orf29	P_000561	CN Loss	Rtl10	YES	25,8930784	53,1514991	16,61557	32,21284	16,99542	34,90238	YES	0,00392	0,45583	NA	NA	NA	12		33	0	0	0,09331	0,04563	2,62354	2,53125	2,47163	1					
C22orf39	P_000561	CN Loss	2510002D24Rik	YES	11,3669223	19,106681	12,41853	19,76872	11,24334	18,91287	YES	0,017109	0,41434	2E 05	0,28592	0,71406	13		34	0	0	1,343959	1,431679	2,53807	2,53125	2,47163	1					
CDC45	P_000561	CN Loss	Cdc45	YES	14,3508455	12,3240207	19,01479	15,43633	8,00299	6,885399	YES	0,86522	0,70043	3,75E 05	0,99988	8,49E 05	13		35	0	1	1,203157	0,482911	2,57293	2,53125	2,47163	1					
COMT	P_000561	CN Loss	Comt	YES	87,9988565	137,96541	71,42726	105,8717	73,48676	115,3141	YES	0,45875	0,50656	1,27E 05	0,39336	0,60663	12		33	0	0	0,00901	0,107433	2,62354	2,53125	2,47163	1					
DGCR14	P_000561	CN Loss	Ess2	YES	11,2971982	20,4688654	10,45443	17,90053	10,24815	18,57143	YES	0,3983	1,8894	2,75E 07	0,91936	0,080638	14		38	1	0	0,03673	0,809327	2,62354	2,53125	2,47163	1					
DGCR2	P_000561	CN Loss	Dgcr2	YES	100,100722	167,194992	82,8498	130,8425	94,50394	157,9775	YES	0,61839	0,12304	3,83E 06	0,95133	0,048668	.	.	.	.	.	.	.	.	.		.	.	.	.	.
ESS2	P_000561	CN Loss	Ess2	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
GP1BB	P_000561	CN Loss	Gp1bb	YES	0,136676	0,06399966	0	0	0	0	YES	0,44731	0,56201	0,51626	0,42214	0,061594	.	.	.	.	.	.	.	.	.		.	.	.	.	.
LZTR1	P_000561	CN Loss	Lztr1	YES	30,2005931	68,8657913	28,78831	62,07815	19,14104	43,67006	YES	0,63008	0,84189	2,21E 89	6,38E 24	1	20		48	0	0	1,879799	1,827243	1,2512	1,28021	0,94608	1					
RIMBP3	P_000561	CN Loss	Rimbp3	YES	0,30205083	0,5571905	0,135037	0,233276	0,233524	0,43125	YES	1,5298	2,7593	2,38E 08	0,15373	0,84627	.	.	.	.	.	.	.	.	.		.	.	.	.	.
RTL10	P_000561	CN Loss	Rtl10	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
SLC7A4	P_000561	CN Loss	Slc7a4	YES	42,4812944	31,3739056	14,44247	10,08804	37,13916	27,45172	YES	0,03804	0,09944	2,11E 07	0,46145	0,53855	13		35	0	0	0,572257	0,526655	0,8079	1,26589	1,04317	1					
UFD1	P_000561	CN Loss	Ufd1	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
UFD1L	P_000561	CN Loss	Ufd1	YES	13,4874668	35,4337574	16,63382	41,31418	12,90988	33,94235	YES	2,767	1,1387	0,9965	0,003495	4,11E 08	13		34	0	0	1,057765	1,018976	2,52508	2,53125	2,47163	1	YES			YES	YES
ZDHHC8	P_000561	CN Loss	Zdhhc8	YES	23,6521633	39,549907	23,07231	36,50398	16,89778	28,30982	YES	1,936	1,2401	0,81885	0,18115	1,37E 06	12		35	0	2	0,514641	0,89633	2,62354	2,53125	2,47163	1					
ZNF74	P_000561	CN Loss	Klf8	YES	15,3704688	28,9629072	14,3174	25,51083	10,22434	19,27864	YES	1,2229	1,043	1,86E 06	0,97071	0,029292	15		46	0	0	0,972241	0,863865	1,30438	1,89912	1,51976	1					
BCAS3	P_000567	CN Loss	Bcas3	YES	5,721111	13,8168018	6,842214	15,62457	7,79709	18,83755	YES	2,7576	0,29747	0,00071	0,99929	1,03E 08	3		6	3	3	0,15245	0,257037	0,550955	0,256312	0,421628	1					
HEATR6	P_000567	CN Loss	Heatr6	YES	12,0525631	17,6986672	14,87138	20,63176	12,62595	18,53575	YES	0,88854	1,0705	3,32E 20	0,80992	0,19008	.	.	.	.	.	.	.	.	.		.	.	.	.	.
TBX2	P_000567	CN Loss	Tbx2	YES	18,55854	33,7672082	25,49512	43,84805	12,50785	22,76906	YES	1,7515	1,3708	0,99134	0,008655	3,77E 07	0		0	0	0	0,01215	0,085138	0,531942	0,92725	0,988479	0	YES			YES	YES
USP32	P_000567	CN Loss	Usp32	YES	9,64355777	22,5704017	10,1056	22,33813	10,26487	24,03409	YES	3,5546	2,2783	1	1,19E 08	5,47E 27	9		8	9	4	2,84641	0,34655	0,92535	0,292861	0,08835	1	YES			YES	YES
FN1	P_000573	CN Loss	Fn1	YES	126,453748	660,694646	164,8331	814,1571	167,6804	876,4151	YES	1,98	0,22832	0,001216	0,99878	7,61E 21	2		3	2	1	0,56224	1,52285	0,846928	1,028008	1,068277	0					
FHIT	P_000579	CN Loss	Fhit	YES	8,31448977	13,6433165	6,418951	9,97824	3,805209	6,24057	YES	1,139	1,2841	0,029167	0,81428	0,15656	1		5	0	0	1,525674	1,157754	0,445892	0,55642	0,2134	0		YES		YES	YES
IL1R1	P_000582	CN Gain	Il1r1	YES	17,4868512	37,3904656	32,2564	65,18287	14,11016	30,15519	YES	1,968	0,24861	0,14331	0,85654	0,000149	0		1	0	0	0,757852	0,769199	0,912427	0,723	1,009247	0					
IL1RL2	P_000582	CN Gain	Il1r1	YES	0,12585511	0,10043225	0,54846	0,414495	0,764328	0,608523	YES	0,46054	1,8635	3,59E 11	0,74593	0,25407	0		3	0	1	0,534166	0,19855	0,724736	0,026388	0,348158	0					
MMD2	P_000582	CN Gain	Mmd2	YES	0,02151843	0,01599991	0,288188	0,203743	0,55525	0,414071	YES	0,01978	0,92125	4,19E 07	0,38035	0,61965	27		2	18	0	2,84641	0,222629	2,62354	0,094173	2,47163	0					
SLC29A4	P_000582	CN Gain	Slc29a4	YES	35,622874	32,3056953	39,43418	33,83117	59,15677	53,69849	YES	0,94576	7,0834	1,04E 05	0,94357	0,056423	2		2	1	1	0,14757	0,009682	0,384392	0,975714	0,781237	0					
WIPI2	P_000582	CN Gain	Wipi2	YES	27,7469606	75,1657101	28,87427	73,95565	33,44457	90,62838	YES	1,7171	0,94026	0,007051	0,99229	0,00066	7		2	5	1	2,5123	0,278307	1,00038	1,195347	0,238675	1					
DGCR2	P_001632	CN Gain	Dgcr2	YES	100,100722	167,194992	82,8498	130,8425	94,50394	157,9775	YES	0,61839	0,12304	3,83E 06	0,95133	0,048668	.	.	.	.	.	.	.	.	.		.	.	.	.	.
NOL10	P_001637	CN Gain	Nol10	YES	22,4464006	25,8304213	26,6142	28,93691	25,16044	28,93601	YES	1,3735	0,58821	0,045571	0,95443	3,41E 07	0		12	0	4	0,901214	1,52093	0,867732	1,50908	0,93108	0					
TMX3	P_001763	CN Gain	Tmx3	YES	23,6472627	54,5035803	22,35977	48,7118	17,88561	41,19992	YES	1,2165	0,16596	0,003792	0,99588	0,000328	3		16	0	7	1,010952	2,86293	0,35588	1,89479	1,43636	0					
AKT3	P_002431	CN Loss	Akt3	YES	9,98360131	22,5992908	13,72781	29,34257	16,45234	37,22031	YES	4,0256	0,10227	0,99956	0,000442	2,66E 10	5		3	5	3	2,60549	0,93311	1,2035	0,105106	0,43328	0	YES			YES	YES
SLC25A13	P_002431	CN Loss	Slc25a13	YES	37,9571398	42,5426172	29,01802	30,73074	38,9876	43,72577	YES	0,42308	0,17083	2,89E 21	0,014202	0,9858	4		5	2	2	0,51296	0,24877	0,5875	0,12264	0,27409	0					
LINGO2	P_002450	CN Loss	Lingo2	YES	0,34424148	0,37273993	0,266137	0,27354	1,188749	1,293123	YES	0,69887	2,1427	0,82307	0,17678	0,000155	.	.	.	.	.	.	.	.	.		.	.	.	.	.
CCDC116	P_002455	CN Gain	Ccdc116	YES	0,36455557	0,26430772	0,545746	0,377553	0,792574	0,579712	YES	0,08626	0,57813	4,1E 06	0,61978	0,38021	0		8	0	0	0,42421	0,409997	1,864258	0,174686	0,546447	1					
MAPK1	P_002455	CN Gain	Mapk1	YES	34,22101	117,065824	34,9678	113,043	43,44961	148,6445	YES	3,7127	0,37659	0,9973	0,002698	1,08E 07	.	.	.	.	.	.	.	.	.		YES	.	.	YES	YES
PPM1F	P_002455	CN Gain	Ppm1f	YES	27,9384014	61,705511	31,69088	66,19394	17,54845	38,79915	YES	1,4268	1,0022	0,003383	0,98737	0,009252	.	.	.	.	.	.	.	.	.		.	.	.	.	.
SDF2L1	P_002455	CN Gain	Ccdc116	YES	98,1332742	32,5478232	103,7136	32,49768	74,7211	24,80722	YES	1,4786	0,94787	0,16422	0,77789	0,057887	0		6	0	0	0,234752	0,30483	1,708757	0,484567	0,742023	1					
TMEM191C	P_002455	CN Gain	Tmem191c	YES	3,36343871	4,36550404	2,277703	2,80752	1,509646	1,960636	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
YDJC	P_002455	CN Gain	Ydjc	YES	72,0801259	38,1504871	56,8613	28,47175	43,24999	22,90299	YES	1,5775	2,2537	0,060829	0,87436	0,064811	0		3	0	0	0,390439	0,423384	1,812725	1,253831	1,416572	0			YES	YES	YES
YPEL1	P_002455	CN Gain	Ypel1	YES	8,00279018	11,381293	11,60481	15,60819	11,65712	16,58609	YES	2,0655	0,63447	0,5906	0,40037	0,009031	.	.	.	.	.	.	.	.	.		.	.	.	.	.
CCDC41	P_002459	CN Loss	Cep83	YES	8,89876425	17,0215548	10,72293	19,37122	7,141952	13,64833	YES	0,46726	0,80071	1,28E 13	0,90788	0,092118	1		5	1	3	0,387142	0,74279	0,964811	0,113625	0,46566	1					
CEP83	P_002459	CN Loss	Cep83	YES	.	.	.	.	.	.	YES	.	.	.	.	.	.	.	.	.	.	.	.	.	.		.	.	.	.	.
EPHA5	P_004500	CN Loss	Epha5	YES	0,20589913	0,61159249	0,194939	0,548054	1,667952	4,950714	YES	1,5318	1,2617	3,15E 05	0,99996	4,48E 06	5		3	4	1	1,93526	0,426006	1,133	0,198584	0,34295	0					

Supplementary table 9 Variants of Unknown significance in rare CNV genes

Sample	Chr	Start	Stop	Ref	Alt	Exon	Gene	dbSNP	Type	location	Effect	HGVS cDNA-level	CADD	gnomAD_E	gnomAD_G
SE14-0527	17	59946466	59946466	.	T	23	INTS2	.	insertion	exonic	frameshift	NM_001330417.1:c.3172dupA	.	0	0
SE14-0656	6	31604286	31604286	G	T	27	PRRC2A	rs569706200	snp	splicing	splicing	NM_004638.3:c.5836-1G>T	23.6	0.0001	3.24E-05
SE16-3109	22	21065731	21065731	G	A	51	PI4KA	rs767451281	snp	exonic	stopgain	NM_058004.3:c.5821C>T	51	0.0002	0.0002
SE16-3114	17	59469360	59469360	C	T	26	BCAS3	rs372166016	snp	exonic	stopgain	NM_001320470.1:c.2773C>T	18.4	0.0002	0.0007
SE16-3123	6	30692109	30692110	CA	A	4	TUBB	.	substitution	exonic	frameshift	NM_001293212.1:c.1330_1331delCAinsA	.	0	0
SE17-3220	7	151962296	151962296	T	C	8	KMT2C	rs751158858	snp	splicing	splicing	NM_170606.2:c.1013-2A>G	22	8.21E-06	0
HK19-0001	3	14485130	14485130	A	G		SLC6A6	.	snp	intronic	splicing	NM_001134367:c.297-6A>G	15.92	0	0
HK19-0002	9	28476025	28476025	T	G		LINGO2	.	snp	intronic	splicing	NM_001258282:c.-395-2A>C	23.2	0	0
HK19-0003	10	52103343	52103344	TA	T	7	SGMS1	.	deletion	exonic	frameshift	NM_147156:c.T529delT-:p.F177del	33	0	0
HK19-0004	10	52104106	52104108	CTG	C		SGMS1	.	deletion	intronic	splicing	NM_147156:c.-313-2CAG>--G	24.5	0	0
HK19-0005	10	52349911	52349911	A	G		SGMS1	.	snp	intronic	splicing	NM_147156:c.-683+2T>C	23	0	0
HK19-0006	17	60072727	60072727	C	T	10	MED13	.	snp	intronic	splicing	NM_005121:c.1968-1G>A	22.4	0	0

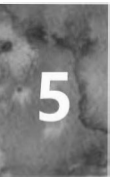
Supplementary table 10

ID	Phenotype	Chromosomal region (hg19)	Size	Type	Inheritance	Class
00826	Abnormality of the antihelix, Aganglionic megacolon, Atrial septal defect, Blue sclerae, Brachycephaly, Bulbous nose, Craniosynostosis, Cryptorchidism, Delayed speech and language development, Frontal bossing, Frontal upsweep of hair, Hiatus hernia, Hypoplasia of the corpus callosum, Hypospadias, Intellectual disability, Macrotia, Micrognathia, Micropenis, Nephrolithiasis, Patent ductus arteriosus, Recurrent urinary tract infections, Ridged cranial sutures, Seizures, Short nose, Short stature, Strabismus, Thick eyebrow, Upslanted palpebral fissure	chr7:138818902-149006346	10187444	Loss	de novo-het	LP
01046	Aganglionic megacolon, Behavioral abnormality, Constipation, Headache, Intellectual disability	chr16:15504454-16284248	779794	Gain	unknown-het	LB
01639	2-3 toe syndactyly, Aganglionic megacolon, Ataxia, Hypospadias, Inguinal hernia, Intellectual disability, Muscular hypotonia, Seizures, Sleep disturbance	chr15:23739358-28525460	4786102	Loss	de novo-het	LP
04119	Adrenogenital syndrome, Aganglionic megacolon, Aplasia/Hypoplasia of the breasts, Frontal balding, Generalized hirsutism, Hypotelorism, Intellectual disability, Intention tremor, Kyphosis, Mitral regurgitation, Patent ductus arteriosus, Prematurely aged appearance, Proportionate short stature, Scoliosis, Seizures, Short palm, Synophrys, Uterine neoplasm	chr5:204849-6753953	6549104	Gain	unknown-triplication	P
		chr5:7521238-43644925	36123687	Gain	unknown-het	P
249397	Aganglionic megacolon, Hyperactivity, Primary amenorrhea, Proportionate short stature, Psychosis, Short attention span	chr22:21032298-21939922	907624	Gain	inherited-het	VUS
249405	Aganglionic megacolon, Brachycephaly, Intellectual disability, Microcephaly, Micropenis, Midface retrusion, Seizures, Tetralogy of Fallot	chr1:242277317-243035431	758114	Loss	unknown-het	LP
		chr1:242987796-246992667	4004871	Loss	unknown-het	VUS
258348	2-3 toe syndactyly, Aganglionic megacolon, Anteverted nares, Hearing impairment, Intellectual disability, Long philtrum, Low-set ears, Posteriorly rotated ears, Renal hypoplasia, Short nose	chr10:1-9190854	9190853	Loss	de novo-het	LP
262159	Abnormal facial shape, Aganglionic megacolon, Global developmental delay, Hearing impairment	chr4:174150963-175434556	1283593	Gain	maternal-triplication	VUS
262199	Abnormality of the foot, Aganglionic megacolon, Growth hormone deficiency, Hydronephrosis, Intellectual disability	chr3:192437045-192641995	204950	Gain	inherited-het	LB
		chr9:220253-17167649	16947396	Gain	de novo-het	LP
269501	Abnormality of the inner ear, Aganglionic megacolon, Delayed speech and language development	chr16:8868640-9270011	401371	Gain	de novo-het	VUS
280593	Abnormality of the mastoid, Aganglionic megacolon, Atresia of the external auditory canal, Dentinogenesis imperfecta, Ectodermal dysplasia, Global developmental delay	chr11:60146368-60228162	81794	Loss	maternal-het	VUS

Supplementary table 10. Copy Numbers Variations from the Decipher Cohort

ID	Phenotype	Chromosomal region (hg19)	Size	Type	Inheritance	Class
281449	Aganglionic megacolon, Delayed speech and language development	chr16:8868640-9270011	401371	Gain	de novo-het	VUS
281751	Abnormal facial shape, Aganglionic megacolon, Agenesis of corpus callosum	chr10:128298104-128614339	316235	Gain	maternal-het	LB
283754	Aganglionic megacolon, Dyscalculia, Enuresis nocturna, Intellectual disability, Microcephaly, Mild global developmental delay, Mild short stature, Synophrys	chr15:22652084-23146103	494019	Gain	unknown-het	LB
307767	Abnormality of the eyelid, Aganglionic megacolon, Clinodactyly of the 5th finger, Hypertelorism, Micrognathia, Open mouth, Premature birth, Thick lower lip vermilion, Thick upper lip vermilion, Wide nasal bridge	chr4:117116790-134810139	17693349	Loss	de novo-het	P
		chr6:57270030-58086843	816813	Loss	paternal-het	VUS
		chr12:132827801-133297838	470037	Gain	maternal-het	LB
326621	Abnormality of the ear, Aganglionic megacolon	chr15:22784523-23179948	395425	Loss	de novo-het	LB
345528	Abnormality of dental morphology, Abnormality of the pupil, Aganglionic megacolon, Blepharitis, High-frequency hearing impairment, Micrognathia, Specific learning disability, Yellow-brown discoloration of the teeth	chr4:39909985-42050575	2140590	loss	de novo-het	LP
350273	Long-segment aganglionic megacolon, Non-midline cleft lip, Unilateral cleft lip	chr16:21475060-21837551	362491	Loss	maternal-het	LB
360735	Abnormal pulmonary artery morphology, Aganglionic megacolon, Arteriovenous fistulas of celiac and mesenteric vessels, Bronchial atresia, Laryngotracheal stenosis, Low-set ears, Retrognathia	chr2:138485012-145271130	6786118	loss	unknown-het	P
288160	Aganglionic megacolon	Chr7:22918653-24645984	1727332	Gain	de novo-het	VUS
		Chr19:54753548-54835792	82245	Gain	Paternal-het	LB
351781	Aganglionic megacolon	chrX:66126285-66952677	826393	Loss	unknown-het	NA
293431	Aganglionic megacolon	Chr2:96545379-98206183	1660805	Gain	unknown-het	VUS
290362	Aganglionic megacolon	Chr16:78307944-78431561	123618	Loss	Loss-het	LB

Supplementary table 6b. Copy Numbers Variations from the Decipher Cohort


5

Chapter 5

Creation and characterisation of induced pluripotent stem cells from Hirschsprung disease patient fibroblasts and their potential for cell replacement therapy

Katherine C. MacKenzie[1], Almira Zada[1], Bianca M. de Graaf[1], Tim Rugenbrink[1], Mehrnaz Ghazvani[2], Lieke Dons[2], Tracy Li[2], Joost Gribnau[2], Cornelius E.J. Sloots[3], Conny J.H.M. Meeuwsen[3], René M. Wijnen[3], Maria M. Alves[1], Erwin Brosens[1], Alan J. Burns[1,4,5], Robert M.W. Hofstra[1,4]

[1]Department of Clinical Genetics, Erasmus University Medical Centre, Rotterdam, Netherlands
[2]Department of Reproduction and Development, Erasmus University Medical Centre, Rotterdam, Netherlands
[3]Department of Paediatric Surgery, Sophia Children's Hospital, Erasmus University Medical Centre, Rotterdam, Netherlands
[4]Stem Cells and Regenerative Medicine, UCL Great Ormond Street Institute of Child Health, London, UK
[5]Takeda Pharmaceuticals, Cambridge, Massachusetts, USA

Corresponding author: Prof. Dr. Robert M.W. Hofstra
Department of Clinical Genetics, Erasmus University Medical Centre, PO BOX 2040, 3000CA Rotterdam, The Netherlands.
Tel: +31-10-7030683. E-mail: r.hofstra@erasmusmc.nl

Manuscript in preparation

Abstract

Hirschsprung disease (HSCR) is a rare congenital gastrointestinal disorder characterised by aganglionosis of the distal colon. The current treatment for HSCR is surgical removal of the aganglionic segment, however, complications following surgery are common. Research into alternative therapeutic strategies are largely focussed on cell transplantation therapy. Recent efforts have focussed on induced pluripotent stem cells (iPSCs) due to their potential to generate large numbers of patient-specific cells as well as other benefits.

We generated and characterised iPSC-derived ENS progenitors from HSCR patients to determine whether the genetic defect causing the HSCR phenotype negatively affects the function of these cells. Fibroblasts from four HSCR patients, with putative deleterious and/or possible pathogenic variants in *RET*, *GFRA1*, *ZEB2* and *EDNRB/EDN3*, and two control individuals with no enteric neuropathy, were lentivirally transduced to create iPSC lines from each individual.

We show that differentiation towards neural crest was indicated by the expression of appropriate markers (*HNK-1*, *RET*, *SNAI2*). The differentiation capacity, and proliferation potential, of HSCR patient cells was not significantly different from control cells, with the exception of cells harbouring *EDNRB/EDN3* variants which expressed fewer neural crest markers and were significantly more proliferative (P=<0.0001). These data demonstrate that although these cells harbour potentially disease-causing DNA alterations, in three out of four HSCR patient lines cell proliferation and cell differentiation were not affected, suggesting that these cells could be competent for therapy without genetic rescue.

Introduction

The enteric nervous system (ENS) is the intrinsic innervation of the gastrointestinal (GI) tract and is formed from cells of the neural crest in a mass migration between weeks four and seven of human embryonic development[1,2]. In Hirschsprung disease (HSCR) the enteric neural crest cells (NCCs) fail to colonise a variable length of the distal GI tract, resulting in an aganglionic region which remains constricted causing a functional obstruction. Approximately 1 in 5000 newborn infants are affected by this congenital malformation. HSCR is considered to be a genetic disease and more than 15 genes have been found so far[3]. Of these, *RET* has been proven to be the major gene with high penetrant coding DNA alterations, ~50% of familial cases and ~15% in sporadic cases[4]. Up to 80% of all cases are found to have low penetrant non-coding RET variants[5]. These data suggest a central role for *RET* in HSCR and ENS development[4]. Pathogenic variants in other genes are known to be causative for HSCR, however their effects on gut and ENS development are less well understood. Despite the fact that more than 15 genes have been found to date, the majority of HSCR patients do not receive a genetic diagnosis[3].

As the pathology of HSCR is primarily the absence of enteric ganglia in the distal gut there is a drive in research towards cell transplantation therapy to replace the missing cells in the aganglionic segment of the gut[6]. It has been shown that a subset of neuronal precursors in the ENS, termed enteric neural stem cells (ENSCs), can be isolated from HSCR patient gut and expanded in culture. These ENSCs can migrate and differentiate into neurons and glia following transplantation into explanted embryonic mouse hindgut[7,8]. Initial *in vivo* transplantations

of ENSCs into mouse models have also shown that these cells can form functional synaptic connections[9-12]. This, together with the apparent ability to isolate neural progenitors from aganglionic HSCR gut[13], has made ENSCs a candidate cell population for transplantation therapy. However, the majority of studies in mouse have shown only modest cell proliferation and spread, suggesting that higher numbers of cells may be necessary to achieve functional rescue in human gut. This may be problematic as it has been reported that ENSCs lose their differentiation capacity over time[14].

An attractive alternative source of patient-derived cells are induced pluripotent stem cells (iPSCs). As well as their ability to differentiate into multiple cell types and their vast proliferative potential[15], iPSCs can be generated from patient fibroblasts or peripheral blood mononuclear cell (PBMS) with minimally invasive techniques. Moreover, being patient derived, they can generate a source of cells for functional investigation into the effects of genetic variants. The use of iPSCs for disease modelling, developmental pathway elucidation and assessing possibilities for cell transplantation therapy has become a focus of ENS research in recent years[16-21]. To our knowledge most of this research has concerned the influence of *RET* mutations on the ENS and the creation and correction of iPSC lines from patients with *RET* variations or with an induced *RET* variant or knockout[18,19].

In this study we induced human dermal fibroblasts obtained from four HSCR patients with variants in different causative genes into iPSCs using an established protocol with alternative plating densities[17]. The pathology of HSCR is thought to be caused by defects in the differentiation, proliferation, migration and/or survival of NCCs as they

colonize the developing gut. We determined the pathogenic nature of the variants and correlated this with the ability of the cells to differentiate to a vagal neural crest fate, as well as assessing their capacity for proliferation, compared with iPSCs from healthy controls.

Materials & methods

Patients & Ethical Approval

Four HSCR patients undergoing routine colonic pull-through surgery in the Paediatric Surgery Department of the Erasmus Medical Centre, Sophia Children's Hospital, Rotterdam were included in this study. This project was approved by the medical ethics committee of the Erasmus Medical Centre (METC-2012-582). Written informed consent for inclusion was obtained from the parents of the patients. Fibroblasts from age-matched controls were included from anonymous in-house stocks.

Fibroblast culture and DNA isolation

Skin biopsies were dissected and plated in Ham's F-10 nutrient mix (Gibco, Thermo Fisher Scientific, USA) supplemented with 15% foetal calf serum (FCS) and 1% penicillin/streptomycin. Medium was refreshed every 2-3 days. Once confluent, cells were split at a ratio of 1:3, using TrypLE Express (Gibco, Thermo Fisher Scientific, USA) according to manufacturer's instructions. Genomic DNA was isolated from fibroblasts with QIAamp DNA Micro kit (Qiagen, Venlo, NL) according to manufacturer's instructions.

RET gene sequencing

All exons of the RET gene were screened for putative deleterious variation using Sanger sequencing. In brief, DNA of patients was amplified using touch-down PCR and an annealing temperature ranging from 70 to 60 °C. Next, sequencing was done using the Big Dye Terminator v3.1 Kit (Applied Biosystems, Foster City, CA, USA) on an ABI 3130XL automated sequencer.

Whole Exome Sequencing analysis

DNA libraries for WES were constructed of seventeen patients using 3 ug of dsDNA and captured with the Haloplex exome target enrichment kit or the Sureselect Clinical Research Exome v1 (Agilent Technologies, Santa Clara, USA). Captured fragments were paired-end sequenced on the Illumina HiSeq4000 sequencers (Illumina, San Diego, USA). Raw sequence data was processed and aligned to the hg19 reference sequence using the Nimbus Suite[22]. We focussed on the known HSCR genes[23] and only considered exonic and putative splicing effect variants with a minor allele frequency below 0.01 in GnomAD exome and/or GnomAD genome (http://gnomad.broadinstitute.org/).

Generation of iPSC lines

We selected four patients for reprogramming: one with a known pathogenic variant and three with a variant of unknown significance. Human primary skin fibroblasts were reprogrammed through infection with a single, multicistronic lentiviral vector encoding OCT4, SOX2, KLF4, and MYC, as described previously[24]. Cells were cultured on γ-irradiated mouse embryonic feeder (MEF) cells until iPSC colonies were picked.

In Vitro Differentiation of Embryonic Bodies

To form embryonic bodies (EBs), iPSC colonies from two wells per line were dissociated by treatment with collagenase IV and transferred to ultralow attachment 6-well plates (Corning). Floating EBs were cultured in iPSC medium without bFGF for a minimum of 6 days. For ectoderm differentiation the medium was supplemented with SB431542. For endoderm differentiation, cells after 6 days were transferred to gelatine coated 12-well plates and cultured in RPMI 1640 medium (Gibco-Invitrogen) supplemented with 20% FBS, alpha-thioglycerol (0.4mM)

and 1:100 dilution of penicillin/streptomycin/glutamine. Mesoderm differentiation from the EBs was induced in gelatine coated 12-well plates with DMEM low glucose medium supplemented with 15% FBS, 1: 100 dilution of penicillin/streptomycin/glutamine and 1:100 dilution of MEM-non-essential amino acids (NEAA). The formation of ectoderm was induced in matrigel coated plates with neurobasal medium (Gibco) and DMEM/F12 (v/v 50/50) supplemented with 1:100 dilution of penicillin/streptomycin/glutamine and 1:100 dilution of MEM-non-essential aminoacids, 0.02% BSA (Gibco), 1:200 N2 (Gibco) and 1:100 B27 (Gibco). After 2 weeks in culture, the cells were fixed with 4% paraformaldehyde for staining for germ-layer specific markers.

Chromosome spreads

For chromosome count analysis, cells in a well of a 6-well plate were treated with colcemid (100 ng/ml) for 1 hour. Cells were then harvested with trypsin, treated with hypotonic solution and fixed. Metaphases were spread onto glass slides and stained with DAPI (Dako). At least 10 metaphases were analysed per cell line.

iPSC culture

Cells were clonally expanded and cultured on vitronectin in Essential 8 feeder-free medium (StemCell Technologies). Medium was refreshed daily or every second day depending on cell confluence. Cells were split at a ratio of 1:10 using ReleSR (StemCell Technologies) when colonies were large and the plate neared 80% confluence.

Neural Crest Differentiation

For NCC differentiation, we used a modification to the published protocol from Hackland, et al.[17]. Cells were dissociated with Accutase (StemCell Technologies) for 10-15 minutes at RT and counted using a

BioRad TC20™ automated cell counter (BioRad). After comparison of varying plating densities, single cells were plated at a density of 100,000 cells/cm^2 onto laminin-521 coated plates (Biolamina) with Essential 6 medium (StemCell Technologies) and Y-27632 dihydrochloride (10μM; Tocris). The next day the medium was changed to NCN2 medium, consisting of DMEM-F12 (Gibco) containing MEM-NEAA (100x; Gibco), Glutamax (100x; Gibco) N2 supplement (100x; Gibco), CHIR99021 (1.0 mM; Tocris), and SB431542 (2.0 mM; Tocris). NCN2 medium was refreshed every second day and cells were sorted on day 10 of differentiation.

Flow cytometry

In order to isolate iPS-NCCs from the mixed cell population the culture was dissociated with Accutase (StemCell Technologies), washed in PBS containing 10% FCS and stained with an antibody against HNK1, a neural crest and early neural marker, conjugated with FITC (1:100, Aviva Systems Biology). Cells were sorted using a BD FACSAria™ III (BD Biosciences, USA), and either re-plated on vitronectin cover-slips or plates for staining, proliferation and migration assays or lysed directly for DNA and RNA isolation.

Immunocytochemistry

iPSCs or differentiated cells were washed with PBS once, fixed with 4% paraformaldehyde solution for 5 min and washed again with PBS. Cells were incubated with 50mM glycine for 5 min, washed with PBS and permeabilized with 0.5% Triton X-100 for 5 min. After blocking for 45 min at room temperature with 0.1% PBS-Tween containing 2% FBS (Invitrogen), primary antibody staining (See Supplementary Table 1 for list of primary antibodies used) was performed for 1 hour in room

temperature with antibodies diluted in blocking solution. Cells were then washed and incubated with the appropriate secondary Cy3 or Alexa Fluor A555 antibody (1:200, Jackson Immunoresearch Laboratories or Invitrogen) for 45 min. Cells were then washed with 0.1% PBS-Tween, with a nuclear staining step in between (Hoechst or DAPI). Cells were covered with Mowiol or ProlongGold (Invitrogen) and a glass slide. Staining for alkaline phosphatase was carried out using the Alkaline Phosphatase kit (Sigma-Aldrich) according to the manufacturer's instructions.

Proliferation assay

Cell proliferation was measured with a standard MTT assay[25]. Wavelengths were measured on a Varioskan™ plate reader (Thermo Fisher Scientific) and proliferation rate of patient lines was compared to that of control lines.

Results

RET sequencing

We included four patients diagnosed with HSCR and undergoing pull-through surgery within the Sophia Children's Hospital, Rotterdam. *RET* screening using Sanger sequencing was negative in 3 patients. Two patient had variants in RET. Patient (P) 1 had two rare variants (NM_020975.4: c.1321A>C, p.Lys441Gln, NM_020975.4: c.1941C>T).

Identification of variants for HSCR with WES

Whole exome sequencing also confirmed: a patient with variants in HSCR disease gene *GFRA1* (NM_005264.4: c.1A>T, p.Met1?, P2); a *ZEB2* mutation in a patient with Mowatt-Wilson syndrome (NM_014795.3: c.1570delA, p.S524Vfs*20, P3); and a patient with variants in *EDNRB* (NM_001201397.1: c.596G>A, p.C199Y) and *EDN3* (NM_001010848.3: c.1770A>G, P4). There characteristics are depicted in table 1, their variant characteristics in table 2.

Establishment of patient-derived iPSC lines

Fibroblasts were reprogrammed to iPSC lines according to established protocols[15,24]. First, four iPSC clones were generated that showed typical characteristics of pluripotent stem cells: expression of alkaline phosphatase, silencing of the multicistronic lentiviral transgene, expression of pluripotency markers (Figure 1) and morphology similar to that of embryonic stem cells in densely packed colonies (Figure 2A). These cells were able to be maintained in long term culture (up to passage 30), and showed a normal diploid chromosome count. All four cell lines generated embryonic bodies that, after differentiation *in vitro*, expressed markers of endoderm, mesoderm and ectoderm (Figure 1). These four lines were further characterised as described below. Multiple

Table 1. Patients included in this study, their predicted pathogenic variants and clinical features (HAEC = Hirschsprung Associated Enterocolitis)

Patient	Line	Variant	Aganglionic Segment	Complications post-surgery
C1	32i	Control	-	-
C2	33i	Control	-	-
P1	75i	RET: c.1321A>C and RET: c.1941C>T	Rectosigmoid; 7cm	1 x HAEC
P2	76i	GFRA1: c.1A>T	Rectosigmoid; 13cm	2 x HAEC
P3	80i	ZEB2: c.1570del Mowat-Wilson syndrome	Short segment; 4.5cm	-
P4	81i	EDNRB: c.596G>A and EDN3: c.565dupA	Rectosigmoid; 9cm	1 x HAEC

Table 2. Patient variant information and predicted pathogenicity

Patient	Gene	Chr	Start	Stop	REF	ALT	Zygosity	Type	GnomAD exome	GnomAD genome
P1	RET	10	43606712	43606712	A	C	heterozygous	missense	0	0
P1	RET	10	43609989	43609989	C	T	heterozygous	synonymous	9.83E-05	0
P2	GFRA1	10	118031541	118031541	T	A	heterozygous	startloss	0	0
P3	ZEB2	2	145157184	145157184	T	.	heterozygous	frameshift	0	0
P4	EDN3	20	57897443	57897443	.	A	heterozygous	frameshift	0.0022	0.0037
P4	EDNRB	13	78492383	78492383	C	T	heterozygous	missense	0	0

Patient	Gene	Variant	CADD	DANN	M-CAP	MetaSVM	MetaLR	Class
P1	RET	NM_020975.4: c.1321A>C, p.Lys441Gln	11.96	0.656	0.046	-0.891	0.296	VUS
P1	RET	NM_020975.4:c.1941C>T, p.=	8.026	.	.	.	.	Likely Benign
P2	GFRA1	NM_005264.4:c.1A>T, p.Met1?	23	0.935	.	-1.088	0.084	VUS
P3	ZEB2	NM_014795.3:c.1570delA, p.S524Vfs*20	34	.	.	.	.	Pathogenic
P4	EDN3	NM_207032.1:c.565dupA, p.T189Nfs*63	24.8	.	.	.	.	VUS
P4	EDNRB	NM_001201397.1:c.596G>A, p.C199Y	32	0.998	0.054	-0.515	0.256	VUS

clones were picked from each line in order to negate possible interference caused by the integration site of the lentiviral vector and two were used for each in further differentiations.

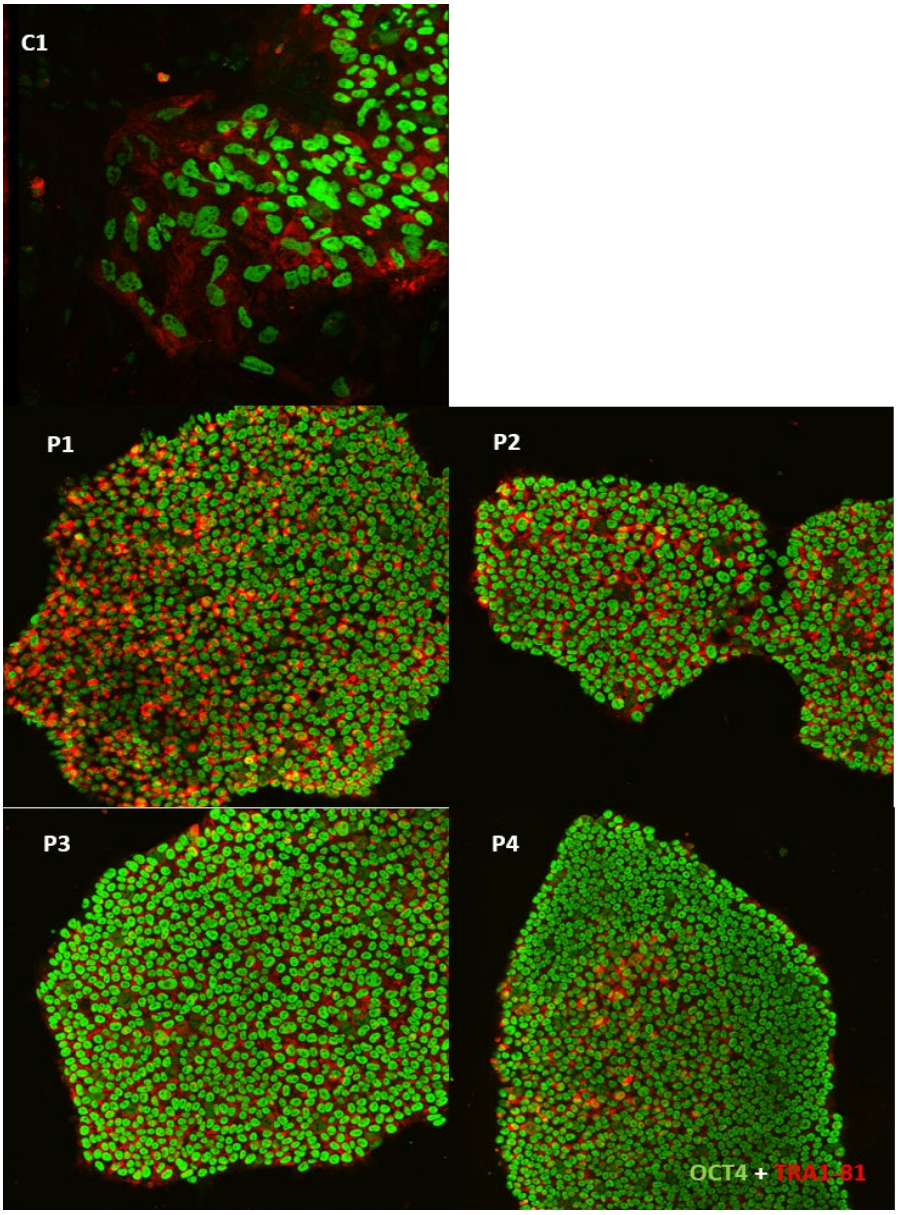
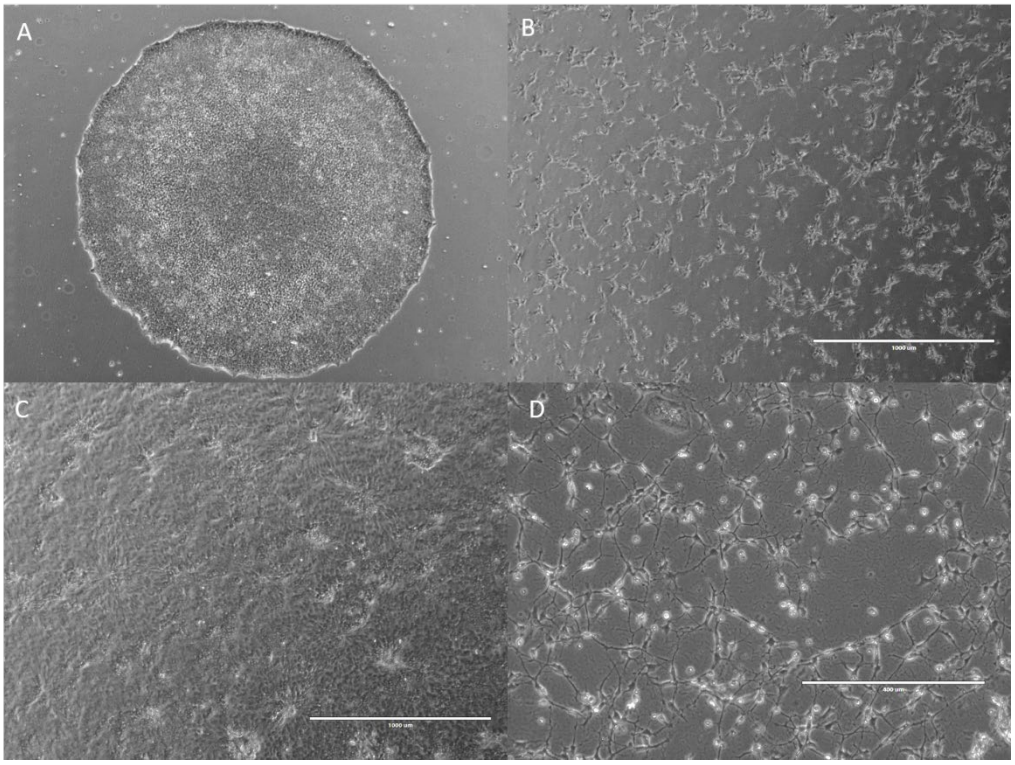


Figure 1. HSCR patient fibroblasts can be reprogrammed to iPSCs and express pluripotency markers OCT4 and TRA1-81.

Expression of NCC genes following differentiation

Using a variation of the protocol from Hackland, et al., all cell lines were differentiated towards a neural crest-like lineage, as assessed by expression of the neural crest cell marker HNK1 in FACS and further by expression of vagal neural crest specific *HOX* genes (B4 & B5). In the first 3 days of the differentiation period cells displayed more neural morphology with evidence of dendritic extensions (Figure 2B). By day (D)4 rapid proliferation led to formation of a dense monolayer (Figure 2C). Following D10 FACS plated cells showed a neural morphology with extensions towards other neighbouring cells (Figure 2D). Following differentiation and FACS, the expression of selected markers for vagal and enteric NCCs were initially measured to confirm either there was similar, higher or lower expression compared to controls (Figure 2E). The expression of the NCC genes measured were not statistically different from those of the controls in P1, 2 and 3 (Figure 2E).

The *RET* variant line, P1, showed no expression of *RET* RNA, indicating possible degradation, whereas other lines showed *RET* expression, indicative of vagal NC lineage. The expression of *P75*, *SOX10*, *CD49D*, *HNK1* and *SNAI2* remained similar to controls. *GFRA1* expression was reduced and *HOXB3* was not expressed, whereas *HOXB4* and *HOXB5* were very highly expressed, demonstrating a less cranial and more vagal NC lineage. The *GFRA1* variant line, P2, showed reduced expression of *GFRA1* and *CHD7*, but similar or increased expression of all other measured genes. The *ZEB2* variant line, P3, did not show reduced expression in any measured genes, but had highly increased expression of *SNAI2* and the *HOX* genes (B3, 4 and 5). The *EDNRB* and *EDN3* variant line, P4, was the exception, showing reduced expression of *P75, SOX10, CD49D, CHD7* and *GFRA1*, and no expression of the *HOX* genes.



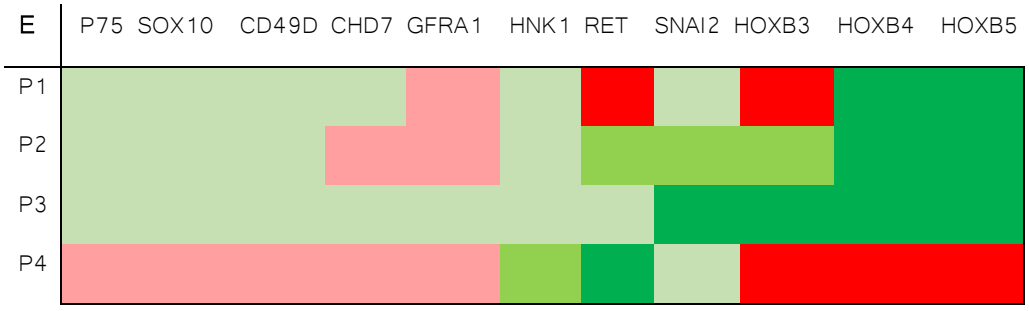


Figure 2. Representative bright field figures of vagal neural crest differentiation in Control 1. A) Day (D)-1 iPSC colony before dissociation; B) D0 single cells the day after seeding; C) D10 monolayer of differentiated cells before FACS; D) D12 48h after FACS showing neural-like phenotype and axonal growth. E) Expression of NCC genes measured by qPCR and PCR compared to control.
Red = not expressed; pink = reduced expression; light green = similar expression; green = increased expression; dark green = more than doubled expression.

Interestingly the expression of HNK1, which we used as a sorting marker, was increased and RET expression was very high in these cells. This could indicate that these cells are at an earlier NC development stage and the HOX genes have not been activated yet, or that there is another defect in differentiation along this lineage.

Cell-cell contact is an important factor in successful differentiation

The differentiation seeding density at D-1 was given as 10,000 cells/cm^2 [17], however using this density we noticed a lack of proliferation and a decreased yield of cells by D10. We increased seeding density 10-fold to 100,000 cells/cm^2 to determine whether cell-cell contact could increase cell yield. Cell survival and/or proliferation during differentiation increased, apparent as greater cell numbers at D10 FACS (data not shown). However, an increased seeding density also led to an overall increase in HNK1$^+$ cells at D10 FACS (Figure 3). The difference between low (10,000 cells/cm^2) and high (100,0000 cells/cm^2) was measured by 2-way ANOVA and found to be highly statistically significant (P= <0.0001) in all cell lines.

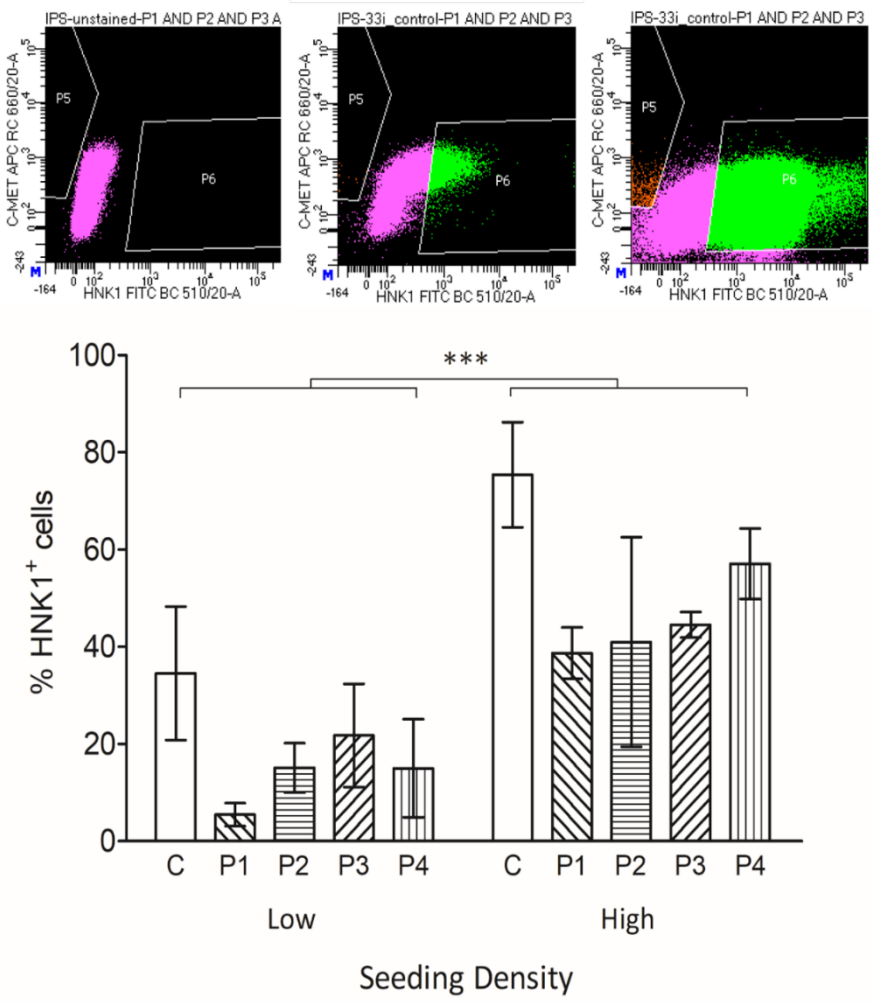


Figure 3. Representative FACS plots following vagal neural crest differentiation at low and high seeding densities. A) Low density seeding of P2 yields and average of ~15% HNK1 positive cells B) High density seeding of P2 yields and average of ~40% HNK1 positive cells C) Graph quantifying differences in differentiation for control and patient lines after seeding at low (10,000 cells/cm^2) and high (100,000 cells/cm^2) cell densities. Statistical significance was measured by 2-way ANOVA. Differences between patients and controls were not significant, but difference in cell seeding density was very highly significant (P=<0.0001).

Proliferation of differentiated cells

Proliferation was measured with a standard MTT assay following plating of FACS sorted HNK1+ cells at D10 of differentiation[25]. Proliferation rate relative to control is quantified in Figure 4. P1, 2 and 3 maintained similar proliferation rates to control when seeded at the same density following FACS. Regardless of seeding density the clones from P4 showed a marked increase in proliferation at all differentiation stages. However this did not correspond with an increased percentage of HNK1+ cells at FACS. P4 showed a highly significant difference from control, the proliferation rate was double that of control and other patient lines (P=<0.0001).

Together these data indicate that, with the exception of P4, the patient cell lines may not require genetic alteration before transplantation trials.

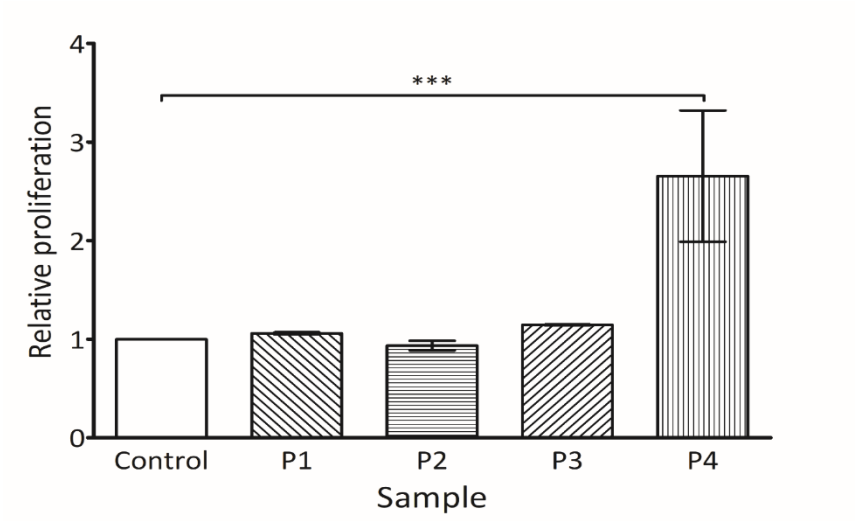


Figure 4. Proliferation and viability measured with an MTT assay after D10 of differentiation. Shown relative to controls set to 1. P1, 2 and P3 have similar proliferation levels to control where in P4 this is more than doubled. Statistical significance measured by 1-way ANOVA. The difference in P4 proliferation was very highly significant (P=<0.0001).

Discussion

We induced iPSCs with a standardised lentiviral protocol[24], creating multiple clones from each patient fibroblast line. Initially three published protocols for the induction of vagal neural crest were compared in a control line, P1 and P2 (data not shown). The first, from Fattahi, et al., uses BMP and TGFβ inhibition directly, adding GSK3 inhibition at D2 and Retinoic acid at D6 of the 11 day protocol[16]. The second protocol was the neural crest induction steps from Lai, et al., with BMP inhibition from D0-3, TGFβ inhibition from D0-4 and GSK3 inhibition from D2-10[19]. These protocols used BMP inhibition from the beginning, however it is known that BMP signalling is necessary for ENS development[26]. Due to animal-derived components in some media, BMP inhibition is required to minimise excess BMPs. As our cultures were feeder-free and used defined medium without animal-derived components, BMP inhibition may have been detrimental to this early specification. The third protocol, from Hackland, et al., addresses this and is simpler in terms of additions as the medium is constant for 10 days, using WNT activation coupled with SMAD, GSK3 and TGFβ inhibition, without inhibiting BMP signalling in order to establish an optimum level of BMPs for enteric NCC differentiation[17]. Overall, the percentages of HNK1$^+$ cells at FACS did not vary greatly, but the overall yield of cells was highest with the protocol from Hackland, et al. which was used for further experiments, although with an increased seeding density (100,000 cells/cm^2).

Maturity of iPSC-derived cells

Using the Hackland protocol, we created iPSC-NCCs from 4 HSCR patients and 2 controls. After the 10-day differentiation we noticed that

success, as measured by the percentage of HNK1+ cells with FACS, was variable between cultures. Cell density and cell-cell contact, i.e culture conditions, seemed to be more important factors than the underlying genetic pathology. Altering the seeding density from relatively low (10,000 cells/cm^2) to high (100,000 cells/cm^2) significantly improved the percentages of HNK1+ cells at D10 of differentiation (Figure 3). Intracellular signalling and excretion of factors after initial induction may make the difference between a successful and unsuccessful culture.

Although we did not notice a difference between the patient derived ENCCs it is possible that enteric neural subtype specifications of the iPSC-NCCs may have been affected by the genetic variants. If this is the case single-cell RNA sequencing could be utilized to find differences in neural subpopulations between patient lines and controls. It should be noted that the stage of differentiation that we accomplished with these protocols is unlikely to show large differences in subtype specification as further maturation may be necessary.

One of the major challenges remaining in the use of pluripotent cells, for both transplantation studies and disease modelling, is the differentiation to more mature cells. Morphological immaturity, lack of expression of markers for mature somatic cell types and the maintenance of potent proliferation in iPSC-derived cells has been documented in multiple cell types[27,28]. There are drives to create protocols for the maturation of these cells, though this may only be necessary for disease modelling purposes. Initial transplantation studies from the Tsakiridis lab, who developed the Hackland et al., protocol, have shown that when cells are more terminally differentiated towards enteric NCCs they do not integrate as well as the more immature vagal

NCCs. This may be due to the higher proliferative and migratory potential in more immature cells, which can then recognize their niche and integrate more effectively (Unpublished, Frith, McCann & Tsakiridis).

Our iPSC-NCCs were used to determine possible proliferation differences induced by the disease-causing mutations present in the different mutated cell lines. Interestingly, the *EDNRB/EDN3* line, P4, had the highest proliferation rate (Figure 4) while seemingly differentiating worst (Figure 2E), although HNK1 expression was a poor marker for this as FACS results were not indicative of decreased NCC differentiation (Figure 2C). The difference in differentiation of these cells does not seem to correlate with the clinical outcomes of the patient (Table 1). This patient had a classical length, rectosigmoid, segment and one episode of HSCR associated enterocolitis (HAEC) post-surgery. They are otherwise growing well on supplemental nutrition, though still require laxatives and rectal irrigation, as do the other patients. It may be that the function of the NCCs is not negatively affected by the decrease in expression of the genes noted in Table 2, and the increase in *RET* may be a compensatory mechanism in this case. Although P4 was the line with the highest proliferation rate, the survival of these cells after FACS did not differ from the other lines. Further culture is needed to test if apoptosis is increased. While *EDNRB* deficiency has been shown to reduce proliferation in early post-natal rat brain[29], Waardenburg-Shah syndrome patients with *EDNRB* or *EDN3* variants do not tend to have associated CNS defects beyond mild mental retardation. More frequently peripheral nervous system defects and pigment abnormalities are associated[30,31]. The increased proliferation rate together with the expression patterns in P4 suggests that these cells

remain at a more immature stage of differentiation. It is possible that more time in culture could further mature this line.

Cell Replacement Therapy

The current treatment for HSCR is the surgical removal of the aganglionic region and anastomosis of the remaining ganglionic gut with the rectum. This leaves a small area at the rectum and anal sphincter that remains aganglionic[32]. Although life-saving, this treatment is far from perfect for quality of life and long-term functional outcomes for patients, with both faecal incontinence and severe constipation frequently affecting patients post-surgery[33,34]. In working towards a possible cell replacement therapy for HSCR, there is debate as to whether therapeutic cells derived from patients with disease-causing variations may require gene therapy, using lentiviral, adenoviral or CRISPR/Cas strategies, to express a healthy copy or repair the gene(s) with a pathogenic variation[28]. In the majority of HSCR cases there is a relatively small region of colon that is aganglionic, whereas the rest of the GI tract is ganglionic and functions fairly normally[35]. This means that the developing ENCCs were effective in migrating, proliferating, differentiating and surviving in these ganglionic regions (comprising the vast majority of the GI tract), which implies that the pathogenic defect in the cells has not completely impaired their ENS-forming capacity. It is therefore likely that autologous treatment for HSCR patients, in the majority of cases, is an option without the requirement of gene therapy. It is, however, possible that in some cases there are differences in enteric neural subtype specification and/or a hypoganglionosis in the remaining gut, which could account for differences in gut function following corrective surgery[34,36,37]. One known change in enteric

neuronal subtype is an increase in nitric oxide synthase (NOS) neurons[38] and NOS-interacting protein[39] in HSCR patient colon.

Genetic stability of the therapeutic cell source should be assessed to ensure that karyotypically normal cells are used for transplantation as pluripotent cells have a tendency to accumulate chromosomal abnormalities[40]. Therefore, although iPSCs provide the theoretical ideal cell source for transplantation options, there remain characterisation and safety hurdles to overcome before the goal of clinical application can be achieved. In our study we noted that iPSCs maintained in culture for 30 passages showed no large chromosomal abnormalities. Microarray data will determine whether differentiation protocols introduce genetic alterations (ongoing).

iPSCs for disease modelling

In order to test patient cell lines prior to transplantation it would be beneficial to use standardised assays to measure changes in function and differentiation capacity of these cells. This could include migration, proliferation and subtype specification of the cells and give more information on pathogenic mechanisms[41]. The development of sensitive and replicable assays will accelerate efficacy testing of cells and bring the possibility of cell transplantation therapy to the clinic faster.

To determine whether genes or variants within these genes cause or contribute to HSCR *in vitro* and *in vivo* models have been explored. Mouse models for aganglionosis have been instrumental in confirming pathogenicity of HSCR genes[42,43], however time and cost make these models inconvenient for large gene screenings. The use of zebrafish as a developmental model for HSCR is quicker and cheaper and has been informative in investigating candidate genes for HSCR given the ability

to visualise changes in ENS cell number and basic gut motility using various reporters[44-46]. However, along with ethical considerations for minimisation of animal models, the genetic orthologs in zebrafish, as well as the ENS complexity and subtype specification, cannot always be counted on for comparison with the human ENS.

A human model would circumvent most of these considerations. This could be done via a human *in vitro* ENS model. Human enteric NCCs could be used to make such a model and these cells can be cultured from gut biopsies[11,47,48]. However, the required numbers of cells for multiple assays or for transplantation trials are difficult to achieve with current culture protocols. This drawback could be overcome by using iPSC-NCCs. Ultimately, the development of a 3D microenvironment, such as an organ-on-chip system would be the ideal controlled test for comparison of patient-derived cells. An organ-on-chip system should also incorporate other cell lineages from the same patient to assess intercellular and extracellular signals from smooth muscle, mucosa and extracellular matrix components. Such a system could also act as a functional test before cell transplantation into the gut of patients.

Conclusions

The current standard of care for HSCR is surgical removal of the aganglionic segment. Further GI symptoms are managed with rectal-irrigation, over-the-counter laxatives and/or anti-diarrheal drugs, electrical stimulation of the anal sphincter or injection of thickening agents to aid anal sphincter closure and prevent anal leakage[49]. The aim of future therapies should be to enable patients to regain intrinsic sphincter control and reduce lifelong reliance on these management options. Our study shows that patient derived cells are not vastly

different from control cells, and may be suitable for transplantation trials.

Three of the four patient iPSC-NCCs differentiated successfully to give similar expression profiles to that of control iPSC-NCCs. It is likely that with transplantation in this immature differentiation stage their integration would be sufficient to colonise the anal sphincter region without the requirement of genetic alteration. Further work will focus on differentiation of these cells to other GI tract cell types for creation of a 3D co-culture system and transplantation of these cells to aganglionic mouse colon to test this hypothesis.

References

1. Anderson RB, Newgreen DF, Young HM. Neural Crest and the Development of the Enteric Nervous System. In: Saint-Jeannet J-P, ed. *Neural Crest Induction and Differentiation*. Boston, MA: Springer US; 2006:181-196.
2. Neural Crest and the Development of the Enteric Nervous System. Landes Bioscience; 2013. https://www.ncbi.nlm.nih.gov/books/NBK6273/.
3. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.
4. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental Biology.* 2016;417(2):198-208.
5. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer variants synergistically drive dysfunction of a gene regulatory network in Hirschsprung disease. *Cell.* 2016;167(2):355-368.e310.
6. Burns AJ, Goldstein AM, Newgreen DF, et al. White paper on guidelines concerning enteric nervous system stem cell therapy for enteric neuropathies. *Developmental biology.* 2016.
7. Almond S, Lindley RM, Kenny SE, Connell MG, Edgar DH. Characterisation and transplantation of enteric nervous system progenitor cells. *Gut.* 2007;56(4):489-496.
8. Metzger M, Bareiss PM, Danker T, et al. Expansion and differentiation of neural progenitors derived from the human adult enteric nervous system. *Gastroenterology.* 2009;137(6):2063-2073. e2064.
9. Hotta R, Stamp LA, Foong JPP, et al. Transplanted progenitors generate functional enteric neurons in the postnatal colon. *The Journal of clinical investigation.* 2013;123(3):1182-1191.
10. Nishikawa R, Hotta R, Shimojima N, et al. Migration and differentiation of transplanted enteric neural crest-derived cells in murine model of Hirschsprung 's disease. *Cytotechnology.* 2014:1-10.
11. Cooper JE, McCann CJ, Natarajan D, et al. In Vivo Transplantation of Enteric Neural Crest Cells into Mouse Gut; Engraftment, Functional Integration and Long-Term Safety. *PLoS ONE.* 2016;11(1):e0147989.
12. Cooper JE, Natarajan D, McCann CJ, et al. In vivo transplantation of fetal human gut-derived enteric neural crest cells. *Neurogastroenterology & Motility.* 2016.
13. Wilkinson DJ, Bethell GS, Shukla R, Kenny SE, Edgar DH. Isolation of Enteric Nervous System Progenitor Cells from the Aganglionic Gut of Patients with Hirschsprung's Disease. 2015.
14. Zhang D, Rollo BN, Nagy N, Stamp L, Newgreen DF. The enteric neural crest progressively loses capacity to form enteric nervous system. *Developmental Biology.* 2018.
15. Takahashi K, Tanabe K, Ohnuki M, et al. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *cell.* 2007;131(5):861-872.
16. Fattahi F, Steinbeck JA, Kriks S, et al. Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. *Nature.* 2016;531(7592):105-109.

17. Hackland JOS, Frith TJR, Thompson O, et al. Top-Down Inhibition of BMP Signaling Enables Robust Induction of hPSCs Into Neural Crest in Fully Defined, Xeno-free Conditions. *Stem Cell Reports.* 2017;9(4):1043-1052.
18. Li W, Huang L, Zeng J, et al. Characterization and transplantation of enteric neural crest cells from human induced pluripotent stem cells. *Molecular psychiatry.* 2016.
19. Lai FP, Lau ST, Wong JK, et al. Correction of Hirschsprung-Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function. *Gastroenterology.* 2017;153(1):139-153 e138.
20. Tang CS-m, Li P, Lai FP-L, et al. Identification of Genes Associated with Hirschsprung Disease, Based on Whole-genome Sequence Analysis, and Potential Effects on Enteric Nervous System Development. *Gastroenterology.* 2018.
21. Wang Y, Lai X, Huang L, et al. Establishment of an induced pluripotent stem cell model of Hirschsrpung disease, a congenital condition of the enteric nervous system, from a patient carrying a novel RET mutation. *NeuroReport.* 2018;29(12):975-980.
22. Brouwer RWW, van den Hout M, Kockx CEM, et al. Nimbus: A design-driven analyses suite for amplicon based NGS data. *Bioinformatics (Oxford, England).* 2018.
23. Alves MM, Sribudiani Y, Brouwer RW, et al. Contribution of rare and common variants determine complex diseases-Hirschsprung disease as a model. *Dev Biol.* 2013;382(1):320-329.
24. Warlich E, Kuehle J, Cantz T, et al. Lentiviral Vector Design and Imaging Approaches to Visualize the Early Stages of Cellular Reprogramming. *Molecular Therapy.* 2011;19(4):782-789.
25. Mosmann T. Rapid colorimetric assay for cellular growth and survival: Application to proliferation and cytotoxicity assays. *Journal of Immunological Methods.* 1983;65(1):55-63.
26. Goldstein AM, Brewer KC, Doyle AM, Nagy N, Roberts DJ. BMP signaling is necessary for neural crest cell migration and ganglion formation in the enteric nervous system. *Mechanisms of Development.* 2005;122(6):821-833.
27. Veerman CC, Kosmidis G, Mummery CL, Casini S, Verkerk AO, Bellin M. Immaturity of human stem-cell-derived cardiomyocytes in culture: fatal flaw or soluble problem? *Stem cells and development.* 2015;24(9):1035-1052.
28. Rathod R, Surendran H, Battu R, Desai J, Pal R. Induced pluripotent stem cells (iPSC)-derived retinal cells in disease modeling and regenerative medicine. *Journal of chemical neuroanatomy.* 2018.
29. Vidovic M, Chen M-M, Lu Q-Y, et al. Deficiency in Endothelin Receptor B Reduces Proliferation of Neuronal Progenitors and Increases Apoptosis in Postnatal Rat Cerebellum. *Cellular and Molecular Neurobiology.* 2008;28(8):1129.
30. Tüysüz B, Collin A, Arapoğlu M, Suyugül N. Clinical variability of Waardenburg–Shah syndrome in patients with proximal 13q deletion syndrome including the endothelin-B receptor locus. *American Journal of Medical Genetics Part A.* 2009;149A(10):2290-2295.
31. Mccallion AS, Chakravarti A. EDNRB/EDN3 and Hirschsprung disease type II. *Pigment cell research.* 2001;14(3):161-169.

32. Goldstein AM, Thapar N, Karunaratne TB, De Giorgio R. Clinical aspects of neurointestinal disease: Pathophysiology, diagnosis, and treatment. *Developmental Biology.* 2016;417(2):217-228.
33. Neuvonen MI, Kyrklund K, Rintala RJ, Pakarinen MP. Bowel Function and Quality of Life After Transanal Endorectal Pull-through for Hirschsprung Disease. *Annals of surgery.* 2017;265(3):622-629.
34. Onishi S, Nakame K, Kaji T, et al. The bowel function and quality of life of Hirschsprung disease patients who have reached 18 years of age or older – the long-term outcomes after undergoing the transabdominal soave procedure. *Journal of Pediatric Surgery.* 2017;52(12):2001-2005.
35. Holschneider AM, Puri P. *Hirschsprung's disease and allied disorders.* Springer Science & Business Media; 2007.
36. Meinds RJ, van der Steeg AFW, Sloots CEJ, et al. Long-term functional outcomes and quality of life in patients with Hirschsprung's disease. *BJS.*0(0).
37. Tani G, Tomuschat C, O'Donnell AM, Coyle D, Puri P. Increased population of immature enteric glial cells in the resected proximal ganglionic bowel of Hirschsprung's disease patients. *Journal of Surgical Research.* 2017;218(Supplement C):150-155.
38. Tomuschat C, O'Donnell A-M, Coyle D, Dreher N, Kelly D, Puri P. NOS-interacting protein (NOSIP) is increased in the colon of patients with Hirschsprungs's disease. *Journal of pediatric surgery.* 2017;52(5):772-777.
39. Cheng LS, Schwartz DM, Hotta R, Graham HK, Goldstein AM. Bowel dysfunction following pullthrough surgery is associated with an overabundance of nitrergic neurons in Hirschsprung disease. *Journal of Pediatric Surgery.* 2016;51(11):1834-1838.
40. Taapken SM, Nisler BS, Newton MA, et al. Karyotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nature biotechnology.* 2011;29(4):313.
41. Workman MJ, Mahe MM, Trisno S, et al. Engineered human pluripotent-stem-cell-derived intestinal tissues with a functional enteric nervous system. *Nature medicine.* 2017;23(1):49.
42. Zimmer J, Puri P. Knockout mouse models of Hirschsprung's disease. *Pediatr Surg Int.* 2015:1-8.
43. Bondurand N, Southard-Smith EM. Mouse models of Hirschsprung disease and other developmental disorders of the enteric nervous system: Old and new players. *Developmental Biology.* 2016.
44. Gui H, Schriemer D, Cheng WW, et al. Whole exome sequencing coupled with unbiased functional analysis reveals new Hirschsprung disease genes. *Genome biology.* 2017;18(1):48.
45. Sribudiani Y, Chauhan RK, Alves MM, et al. Identification of Variants in RET and IHH Pathway Members in a Large Family With History of Hirschsprung Disease. *Gastroenterology.* 2018.
46. Schill EM, Lake JI, Tusheva OA, et al. Ibuprofen slows migration and inhibits bowel colonization by enteric nervous system precursors in zebrafish, chick and mouse. *Developmental Biology.* 2016;409(2):473-488.
47. Burns AJ. Migration of neural crest-derived enteric nervous system precursor cells to and within the gastrointestinal tract. *International Journal of Developmental Biology.* 2003;49(2-3):143-150.

48. Burns AJ, Roberts RR, Bornstein JC, Young HM. Development of the enteric nervous system and its role in intestinal motility during fetal and early postnatal stages. *Seminars in Pediatric Surgery.* 2009;18(4):196-205.
49. Westfal ML, Goldstein AM. Pediatric enteric neuropathies: diagnosis and current management. *Current opinion in pediatrics.* 2017;29(3):347-353.


5

Supplementary Table 1

Protein	Antibody	Dilution	Company
OCT4	goat anti human	1:100	Santa Cruz Biotechnology
TRA1 81	mouse anti human	1:100	Santa Cruz Biotechnology


6

Chapter 6

Treatment avenues for Hirschsprung disease and a novel treatment option for faecal incontinence

Katherine C. MacKenzie[1], Martijn Timmermans[2], Leo Jessen[2], Sonja Blaauw[3], René M. Wijnen[4] & Robert M.W. Hofstra[1]

[1] Department of Clinical Genetics, Erasmus University Medical Centre, Rotterdam, Netherlands
[2] V.O. Patents & Trademarks, Den Haag, Netherlands
[3] Technology Transfer Office, Erasmus University Medical centre, Rotterdam, Netherlands
[4] Department of Paediatric Surgery, Erasmus University Medical Centre, Rotterdam, Netherlands

Patent filed

Introduction

Hirschsprung disease (HSCR) is a rare congenital neuropathy affecting approximately 1 in 5000 live births. It presents as a lack of enteric nervous system (ENS) ganglia in a variable length of the distal segment of the colon which results in tonic constriction of the smooth muscles around the gut lumen and a functional obstruction[1]. The current treatment for HSCR is surgical resection of the affected region. Although this removes the majority of aganglionic bowel, the internal anal sphincter (IAS) and the external anal sphincter (EAS) lack an ENS and the complex can be damaged by the surgical procedure[2-4]. Consequently, the current standards of care for HSCR patients leave many with poor anal sphincter control and reports of both chronic constipation and anal leakage following surgery[5,6]. Treatments or therapies for HSCR therefore need to aim for sphincter control, for both opening and closing. Current surgical solutions for this problem focus on closure, with no solution that will address both contraction and relaxation of the muscles[7].

Given the low prevalence of HSCR, the demographic for patients that have continued anal sphincter malfunction throughout their lives is therefore not sufficiently large to have prompted commercial investment in novel treatments. However, considering anal incontinence through other causes, the numbers of affected individuals collectively increases. The overall prevalence of anal incontinence in adults is estimated to be 11-15% and increases with age[8]. A review of pelvic floor imaging techniques reported that approximately one third of people living in retirement homes or similar institutions suffer from age related incontinence[9]. In the USA and Europe prevalence of anal incontinence in elderly nursing homes

is approximately 45%, and this is likely to be underestimated due to underreporting[10,11]. There are also patients suffering from anal incontinence due to lower body paralysis, caused by spinal cord injury[12,13] or other muscular damage, that would benefit from additional treatment options[14].

Some treatment options currently available for anal incontinence are discussed briefly below, however any currently available prosthetic option is focussed on constriction and closure of the sphincter, with no solutions allowing for sphincter relaxation, which is necessary in the case of HSCR. Here, we propose a physiologically relevant prosthetic solution for anal incontinence that could be suitable for multiple causes, and affect relaxation as well as contraction of the anal sphincter. The proposed solution of an electronic prosthetic anal sphincter is a multi-disciplinary approach, combining surgical techniques with microelectronics and physiological function in order to provide patients with a modern and convenient solution to anatomical malfunction.

The anal sphincter complex

The anal sphincter complex comprises the internal anal sphincter (IAS) and the external anal sphincter (EAS), the combined function of these allows for normal physiological defecation[15]. The IAS is partially a continuation of the circular muscle fibres of the colonic walls and its function is entirely involuntary, controlled by parasympathetic nerve fibres and the myenteric plexus of the ENS[16]. Conversely, the EAS is mainly innervated by the pudendal nerve and is under voluntary control[17]. During defecation, the increased pressure and stretch from bowel contents on the anal canal signals for relaxation. The opening of

the IAS and the pressure transfer to the EAS is picked up by sensory nerves more distally. These combined signals let us know that we need to defecate and the voluntary control of the EAS allows us to choose when that can take place[15]. The lack of this control is the primary cause of adult faecal incontinence, and a number of pathologies can contribute to this[18].

Common treatment options

Following basic exams for incontinence, conservative treatment is preferred. Patient training with pelvic floor muscle exercises, dietetic management and regular colonic irrigation for stool evacuation is a common treatment for incontinence, however muscle training can be difficult to explain to younger patients, such as those with HSCR or anorectal malformation[6,19]. Conservative treatments are favoured by doctors for treatment of older patients, given in combination with electrical nerve stimulation. Electrical nerve stimulation, in various forms, has been reported to have extremely variable success rates, 0-100%, the majority between ~50-80%, however this differs little from regular muscle training[20]. For all patients with chronic issues, additional drugs are recommended. A mix of laxative drugs, when constipated, and anti-diarrheal drugs, when frequent leakage occurs, will allow the patient to regain some control over their bowel movements. However, the frequent use of these drugs has a negative impact on natural gut motility, absorption and overall function[19,20].

Prosthetics

If a patient is not responsive to conservative therapies, there are currently no routinely offered options for a prosthetic sphincter. There have been clinical trials of a fluid-pumped anal sphincter

(US5593443A) that works on the basis of a fluid-filled ring that is surgically fitted around the outside of the distal colon and can be manually pumped-up to close the rectum or let out to allow the rectum to release on its own. The uptake for available prosthetics is low. Due to the manual nature of the prosthetics, the issue remains of knowing when you have the real need to empty your bowel. A stoma is the only surgical alternative for frequent issues with anal incontinence when a patient is unresponsive to non-surgical therapies[7].

An extensive review by Fattorini, et al., detailed research into alternative prosthetic options for anal incontinence[7]. All of these options are based on a cuff-like system, where the device will surround the colon/rectum wall, and therefore require open surgery for placement. No device listed is intended for implantation into the muscle of surrounding fat of the anal sphincters. Additionally, none of these devices function in a way that is close to physiologically and anatomically "normal".

Electronic anal sphincter

The device proposed in our patent could be surgically inserted either into the muscle or the adjacent fat of the anal sphincters. As this is not a cuff device it would not function for the constriction of the colon, but specifically target the region of the anal sphincters for closure and opening, rather than allowing passive opening. This is beneficial for patients with HSCR where tonic contraction of the muscle is unable to release due to lack of enteric neural signals[21]. Improvements on a stoma are evident as this would not be a detectible or visible device

and could be operated in a minimally noticeable and minimally invasive way, and would not require manual cleaning/emptying.

The device proposed will be composed of two separate ring structures that would be electronically linked. The first would be dependent on a pressure-sensor that will trigger opening, "relaxation" , when the pressure reaches the normal physiological range for requiring defecation. This release and opening of the first structure will trigger an alert to the user, either linked to a remote-control, a beeper/pager alert and/or a phone application that will let the user know they will need to defecate soon. The second structure can then be released at the user's convenience. The device will be implantable at each of the anatomical anal sphincters.

Given the safety and other challenges with the use of cell therapy that are yet to be overcome, such an approach is unlikely to reach the clinic for a number of years, and could have highly variable success rates within patients. This could be appropriate for any issue of anal incontinence, and possibly other neuro-muscular issues at other GI junctions (e.g. oesophageal sphincter, pyloric sphincter, ileocecal sphincter) and possibly other organs that require a similar mechanism. Patients that could benefit from this device are those with sphincter control issues due to HSCR, Inflammatory Bowel disease, nerve damage (for example from childbirth, injury or surgery), multiple sclerosis and/or spinal cord damage as well as age-related incontinence.

Further detailed descriptions and illustrations of the proposed device are specified in the following patent document.

The Electronic Anal Sphincter

Patent number - P122362NL00

Title: Electronic anal sphincter

Field of the invention

The invention relates to an artificial sphincter.

Abstract

A human body is provided with a plurality of sphincters – a circular muscle for constricting and/or opening a body lumen around which it is provided. An implantable device such as an artificial sphincter may be required when the naturally present sphincter is not functioning as desired. The implantable device may be provided with a lumen mechanism and an engagement section which are arranged for interfacing with the body lumen such that constriction and/or opening forces may be coupled between the body lumen and the implantable device. In a preferred embodiment, the lumen mechanism is arranged as a foldable planar linkage, of which a radial expansion and/or contraction is directly coupled to a distance and/or angle between different joints and/or links comprised by the linkage. By providing an actuator and/or sensor between different links and/or joints of the linkage, a sphincter-like coupling may be obtained between the device and the body lumen.

Background

Hirschsprung disease is an enteric neuropathological disease, which affects the nervous system of the gastrointestinal (GI) tract, the enteric

nervous system (ENS). The ENS normally regulates the peristaltic movement of, and defecation of faeces from, the gastrointestinal tract, and is not present in the distal portion of the colon in a patient with Hirschsprung disease. This results in a portion of colon that is constricted and surgical intervention is required to correct the defect. However, although this surgery removes most of the constricted portion of gut, the anal sphincters of the patient remain without innervation.

Patients that have these defective sphincters experience incontinence problems with some requiring frequent enemas, the implantation of a stoma and/or requiring the use of diapers in the years post-surgery.

US5593444A discloses a liquid-pumped anal sphincter that works on the basis of a liquid-filled ring that is surgically fitted around the outside of the distal colon and can be manually pumped-up to close the rectum or let out to allow the rectum to release on its own. A stoma is an alternative to cope with issues with anal incontinence.

Summary

Known artificial sphincter devices are provided around the sphincter, and are only arranged for forcefully closing the sphincter. Passive opening is not possible due to automatically contracted muscle that is not able to release due to lack of enteric neural signals. Drawbacks of a stoma are that is an easily detectible or visible device and it requires manual cleaning or emptying.

It is preferred to provide an improved artificial sphincter.

A first aspect provides an implantable device for engaging with a body lumen, comprising of a flexible housing, comprising of a housing lumen arranged to be radially expanded and retracted between an opened and a closed position. A lumen mechanism provided in the housing, comprising a mechanism lumen which is substantially aligned with the housing lumen and arranged to be radially expanded and retracted between an opened position, corresponding to the opened position of the housing, and a closed position corresponding to the closed position of the housing. An engagement section for engaging with the body lumen, for providing a coupling between radial expansion and retraction activity of the body lumen and the implantable device.

A lumen is defined as a passage for a fluid and/or solid, which may be water, blood, faeces, urine, any other bodily fluid, solid, emulsion, suspension, mixture, any other bodily solid, or any combination thereof. A lumen has a certain flow-through area for the passage, which may be increased or decreased. A lumen may thus be regarded as a cross-section of an elongated, e.g. tube-like, organ arranged for transporting a fluid and/or solid there through.

In a human body, a body lumen may be surrounded by a sphincter muscle. Sphincter muscles may be voluntary or involuntary controlled; i.e. actively and consciously controllable by a human or animal or not. Voluntary sphincter muscles may be used to control a flow-through area of a body lumen, whilst involuntary sphincter muscles may be used as sensors to sense whether a body lumen is open, closed, or somewhere in between. An implantable device according to the first

aspect may thus be employed as an artificial sphincter muscle, seeing as it may deliver similar properties as a sphincter muscle.

Examples of sphincter muscles for which the implantable device such as an artificial sphincter may be used are the internal anal sphincter, external anal sphincter, ileocecal sphincter, urethral sphincter, pyloric sphincter, oesophageal sphincters and the urethral sphincter.

With engaging with a body lumen, it is implied that forces may be exchanged between the engagement section and the body lumen, in two directions. Hence, forces from the body lumen may open and close the implantable device via the engagement section, and forces from the implantable device may open and close the body lumen via the engagement section. As such, the mechanism lumen and the body lumen may open and close simultaneously.

The housing lumen provides a passage through the housing, wherein said passage has a certain flow-through area. With the radial expansion of the housing lumen, this flow-through area increases, and with the radial retraction, the flow-through area decreases. The flow-through area may, in a retracted state, be substantially zero, thus substantially blocking any flow through the housing lumen.

The housing may comprise materials which are compatible to be inserted into a human body, such as titanium, collagen, biocompatible polymer, other compatible materials, or any combination thereof.

The housing may be substantially doughnut-shaped, i.e. the housing comprises a non-zero inner radius defining the housing lumen as an inner boundary, and an outer radius defining an outer boundary, wherein the outer radius is larger than the inner radius.

The lumen mechanism is provided inside the housing, and as such at least partially defines the shape of the housing. The housing is provided at least substantially around the lumen mechanism, and may as such provide a barrier between the lumen mechanism and its moving parts, and the body in which the device is implanted.

Because the engagement section is arranged to provide a coupling between both radial retraction and expansion activity, not only can the body lumen be forcefully closed, but can also be forcefully opened.

The radial expansion and retraction activity of the body lumen and the implantable device may comprise a radial expansion of the body lumen, which by virtue of the coupling provided by the engagement section, may result in a radial expansion of the implantable device, and more in particular the lumen mechanism. Similarly, may the radial expansion and retraction activity of the body lumen and the implantable device comprise a radial retraction of the body lumen which by virtue of the coupling, provided by the engagement section, may result in a radial retraction of the implantable device.

The radial expansion and retraction activity of the body lumen and the implantable device may comprise a radial expansion of the implantable device, and more in particular the lumen mechanism, which by virtue of the coupling, provided by the engagement section, may result in a radial expansion of the body lumen. Similarly, may the radial expansion and retraction activity of the body lumen and the implantable device comprise a radial retraction of the implantable device, and more in particular the lumen mechanism, which by virtue of the coupling, provided by the engagement section, may result in a radial retraction of the body lumen.

The lumen mechanism may comprise a foldable planar linkage comprising a plurality of links connected by a plurality of joints, wherein the linkage encloses an internal area substantially corresponding to the mechanism lumen and wherein the linkage is foldable between a folded position and an unfolded position, wherein in the folded position the internal area enclosed by the linkage is smaller than in the unfolded position. Typically, such a linkage is over constrained. A degree of freedom may however be obtained by choosing appropriate link lengths.

The surface area of the inner area may be controlled by folding and unfolding the linkage, wherein folding the linkage decreases the inner area and unfolding the linkage increases the inner area. With the inner area decreasing, a radial distance between two joints which substantially lie on the same radially extending line increases. Furthermore, with the inner area decreasing, a tangential distance between two other joints which are provided substantially equiradially relative to a centre point of the linkage decreases.

This relationship between the surface area of the inner area and distances between different joints may be used both for forcefully opening and closing a sphincter, as well as detecting a pressure in the sphincter. In the first case of opening and closing the sphincter, one or more actuators may be provided to control one or more distances between joints for controlling the inner area of the linkage when the linkage is provided around the sphincter. In the latter case of detecting a pressure in the sphincter, when the linkage is provided around the sphincter, the sphincter will exert a pressure to the inner joints of the

linkage, which may be measured by measuring a force between different joints of the linkage.

As such, the use of a foldable planar linkages in an artificial sphincter with an actuator and/or a sensor may be regarded as a single innovative concept, using the relationship between the inner area delimited by the linkage and the distance between different joints. Since this distance may be measured and/or controlled, it is possible to control and/or measure the inner area of the linkage, as well as to measure a force or pressure exerted on the inner area of the linkage.

An example of a foldable planar linkage is a Hoberman mechanism, which is a one degree-of-freedom (DOF) mechanism comprising a plurality of linkages. The specific hinging connections of the linkages, and their lengths, allow a coupling between a circumferential motion and a radial motion. Hoberman mechanisms have been used for example in the space industry and in toys. Use of other foldable planar linkages with any number of links and any number of joints is also envisioned.

A foldable planar linkage is often substantially axisymmetric, in folded and unfolded state. As a result thereof, a substantially equal radial force may be obtained across the entire circumference of the linkage which may prevent pressure points on the body lumen.

When the implantable device is to be used for at least partially controlling the flow-through area of a body lumen, the device may comprise of an actuator for manipulating the lumen mechanism such that the mechanism lumen can be positioned between a closed position and an opened position, wherein the engagement section is

arranged to transfer retraction and expansion forces from the actuator to the body lumen to which the implantable device is engaged. As such, the implantable device may be employed as an artificial voluntary sphincter. As the mechanism is to be provided around tissue or within tissue defining the body lumen, the closed position of the lumen mechanism and the closed position of the housing do not necessarily define positions in which there is no central passage defined in the centre of the lumen mechanism and/or the housing. Rather, in the opened position, this passage has an area larger than in the closed position.

The actuator may comprise two ends, and the actuator may be arranged to manipulate an actuator distance between the two ends. Next, the actuator may at a first end be connected to a first joint, which first joint connects a first link and a second link, at a second end connected to a second joint, which second joint connects a third link and a fourth link, wherein manipulation of the actuator distance folds the planar linkage between the folded position and the unfolded position.

Next to a relation between a distance between two joints, there is also a relation between an angle between two adjacent links and the flow-through area of the mechanism lumen. As such, an angular actuator may be provided connecting two adjacent links, and arranged to manipulate the angle between the two adjacent links.

Next to or instead of being used as an actuating device, embodiments of the implantable device may comprise a sensor for determining lumen data of the mechanism lumen, wherein the engagement section is arranged to transfer retraction and expansion forces from the body

lumen to which the implantable device is engaged to the sensor. As such, the implantable device may be employed as an artificial involuntary sphincter.

Lumen data of the mechanism lumen may comprise data on a flow-through area through the mechanism lumen at a certain point in time. Such data may relate to an absolute value, e.g. a diameter measured in millimetres or an area measured in millimetres squared. Such data may alternatively or additionally relate to a relative value relating to a state of the mechanism lumen, e.g. closed, fully opened, or partially opened.

Lumen data of the mechanism lumen may further comprise, as an option or additionally to the data on the flow-through area, data on a pressure exerted on the lumen mechanism. Such a pressure may be caused by material passing through the body lumen which cause the body lumen to expand radially. E.g. faeces passing through part of a colon as a body lumen may cause the colon to expand radially by exerting a pressure on the colon. Pressure may also be exerted on the lumen mechanism by a sphincter muscle or artificial sphincter provided adjacent to it.

The sensor for determining lumen data of the mechanism lumen may comprise two ends. The sensor may be arranged to obtain data on a distance and/or pressure between the two ends, and the sensor may be at a first end connected to a first joint, which first joint connects a first link and a second link, at a second end connected to a second joint, which second joint connects a third link and a fourth link.

When the lumen mechanism is arranged as a planar linkage mechanism, this sensor may determine lumen data based on the measured distance and/or pressure between the two ends.

The sensor for determining lumen data of the mechanism lumen may be connected to a first link and a second link provided adjacent to the first link. In such a case, the sensor may be arranged to obtain data on an angle and/or force between the first link and the second link.

When the lumen mechanism is arranged as a planar linkage mechanism, a first subset of joints may be provided on a first radius from a centre point of the linkage, a second subset of joints may be provided on a second radius from the centre point of the linkage, and a third subset of joints may be provided on a third radius from the centre point of the linkage, and wherein the links may be provided between joints from the first subset and the second subset, and between the second subset and the third subset.

The implantable device may be provided with an input module for receiving a control signal, and a control module for operating the actuator in accordance with the control signal. The control signal may comprise data on a preferred state for the implantable device. Such a preferred state may relate to the state of the mechanism lumen, e.g. open or closed, or to a particular flow-through area for the mechanism lumen. The control signal may be provided by a user, a sensor, or both.

The sensor may be arranged to output an output signal comprising distance/pressure data. The output signal may be provided to a user, e.g. to a portable user device with a graphical interface. As such, the

user may be made aware of a state of the sensor, and thus a state of the implantable device.

A second aspect provides a kit of parts comprising a first implantable device according to any of the embodiments in which the first implantable device comprises an actuator for manipulating the lumen mechanism and a second implantable device in which the second implantable device comprises a sensor for determining lumen data on the mechanism lumen of the second implantable device. As such, a situation comparable to the human anatomy, in which the colon is provided with a first, involuntary, sphincter muscle and a second, voluntary, sphincter muscle may be approximated.

A third aspect provides a method for opening and closing an implantable device, comprising of a user controlled device, receiving distance/pressure data from a distance/pressure sensor comprised by a second implantable device, by the user control device, outputting a user signal related to the received distance/pressure data, by the user control device, receiving a user input signal comprising actuation data, by the user control device, sending an actuation signal comprising actuation data to a second implantable device comprising a lumen actuator, and actuating the lumen actuator according to the received actuation data.

A fourth aspect provides another method for opening and closing an implantable device, comprising detecting, with a sensor provided by the implantable device, a change in body lumen opening state of a body lumen around which the implantable device is provided, and manipulating, with an actuator provided by the implantable device, a mechanism lumen in accordance with the change in body lumen

opening state. With such a method, a sphincter muscle may be assisted in its operation.

Brief description of the figures

The various aspects and embodiments thereof will now be discussed in conjunction with drawings. In the drawings:

Fig. 1A shows a schematic view of an implantable device provided around a body lumen;

Fig. 1B and Fig. 1C show the implantable device in a side view around the body lumen;

Fig. 2 shows a schematic view of another embodiment of the implantable device provided around a body lumen;

Fig. 3A and Fig. 3B show yet another embodiment of the implantable device

Fig. 4A and Fig. 4B show part of a planar linkage mechanism as a lumen mechanism of an implantable device.

Fig. 5A and Fig. 5B show proposed linkage arrangement for the implantable device, planar and joined as a ring.

Fig. 6 shows proposed housing lumen that would act as a casing for the implantable device.

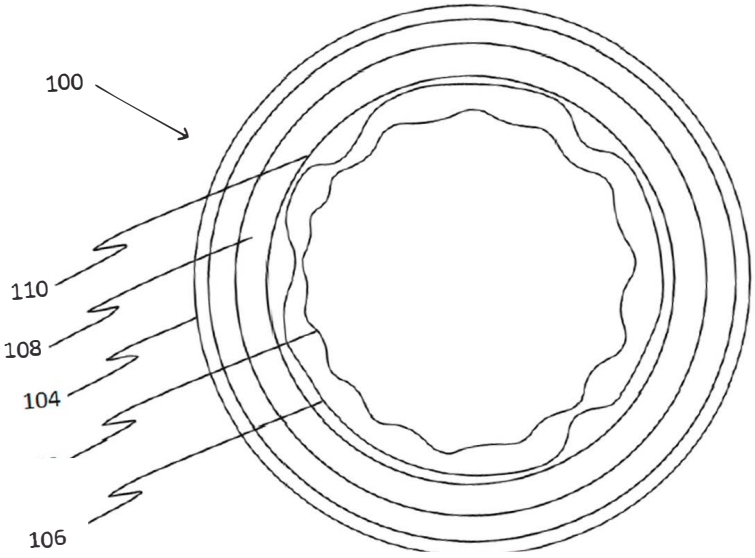


FIG 1A

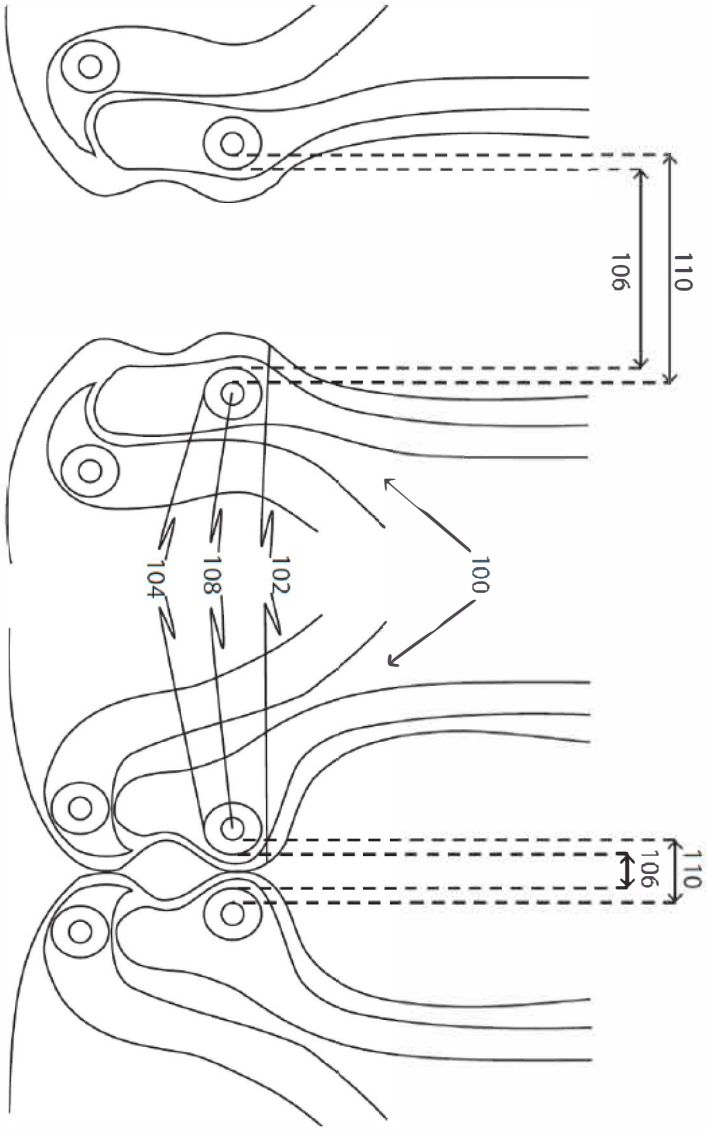


FIG 1B

FIG 1C

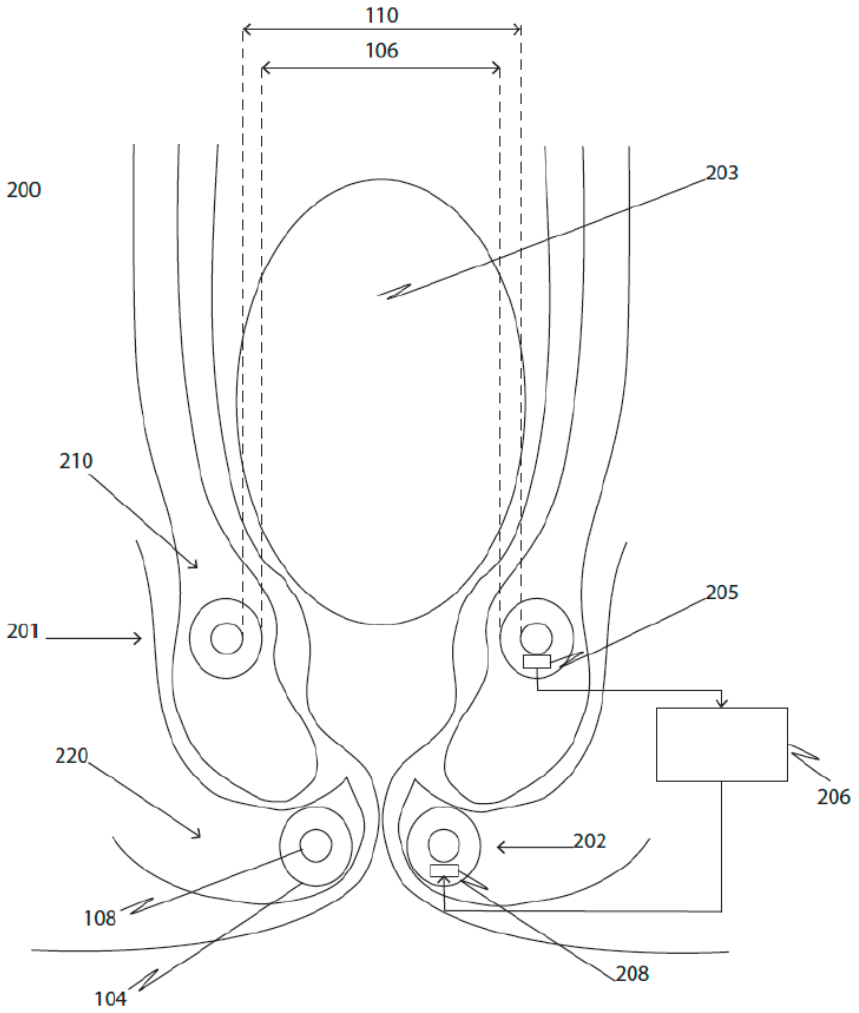


FIG 2

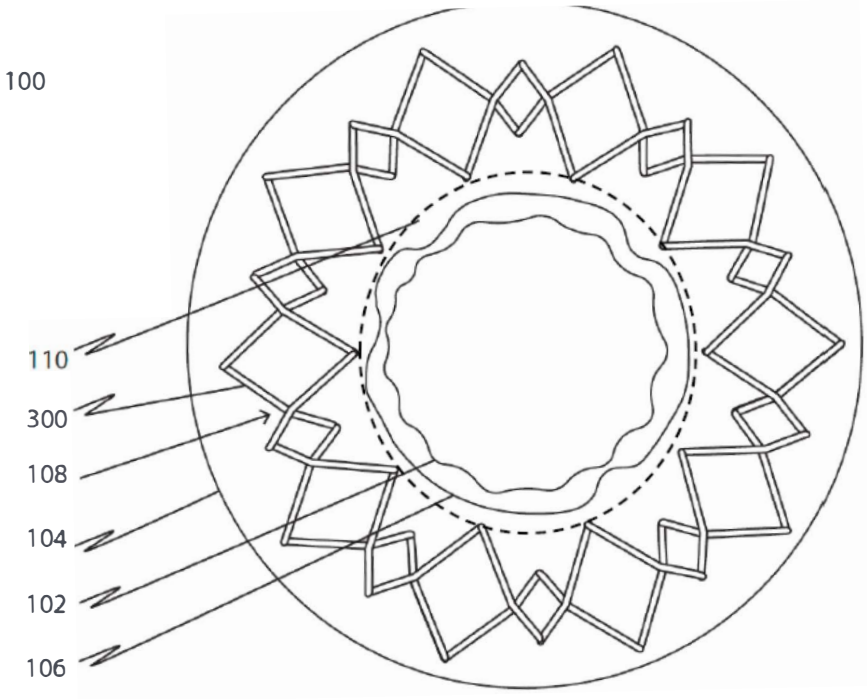


FIG 3A

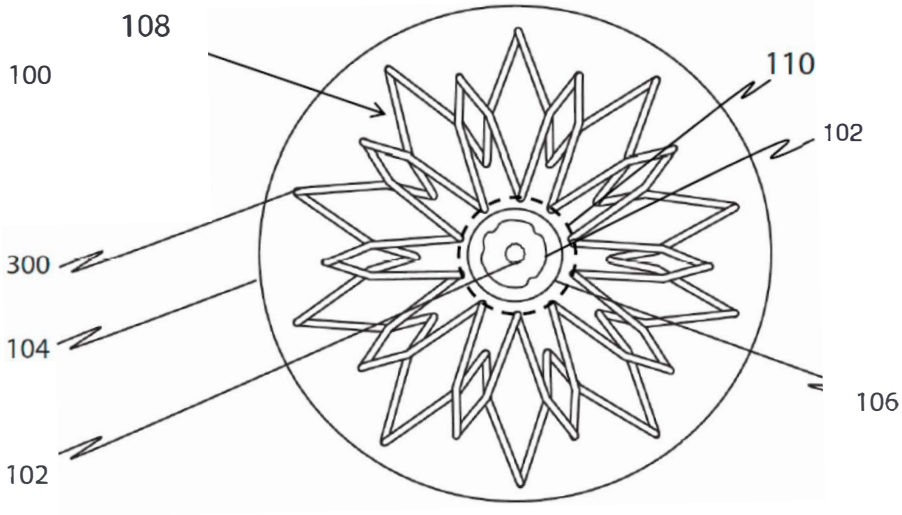


FIG 3B

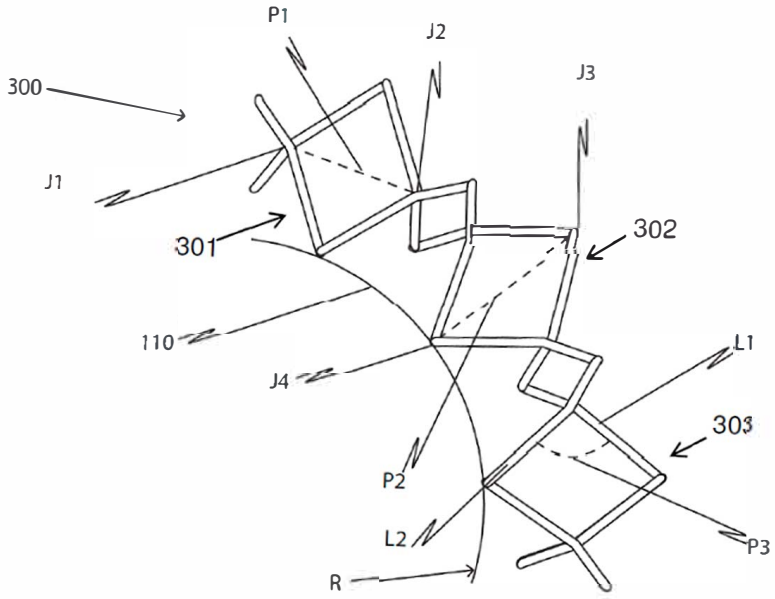

P1
J2
J3
300
J1
301
302
110
L1
J4
303
P2
P3
L2
R

FIG 4A

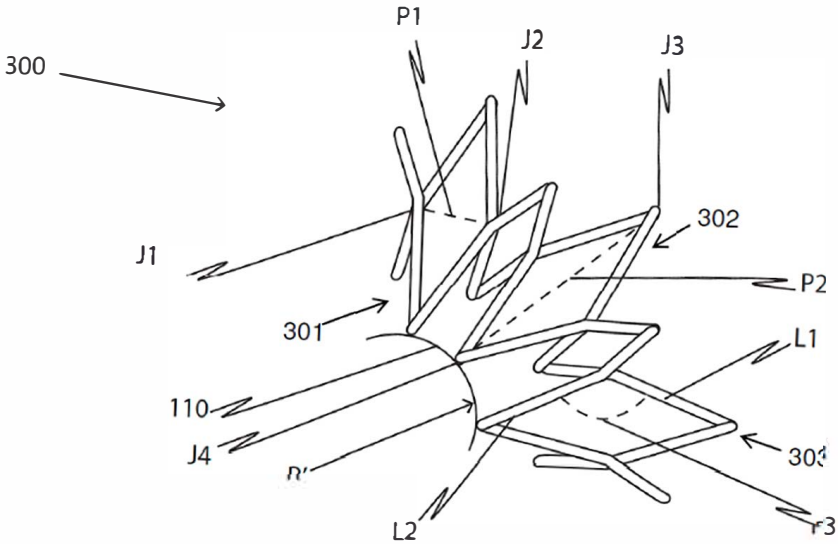

P1
J2
J3
300
302
J1
P2
301
L1
110
J4
303
L2

FIG 4B

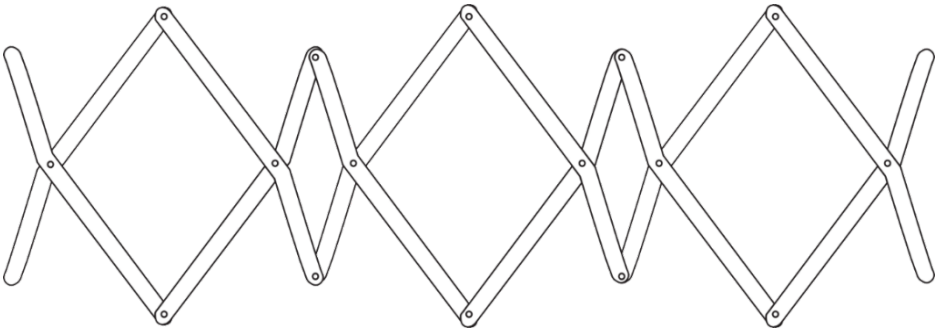
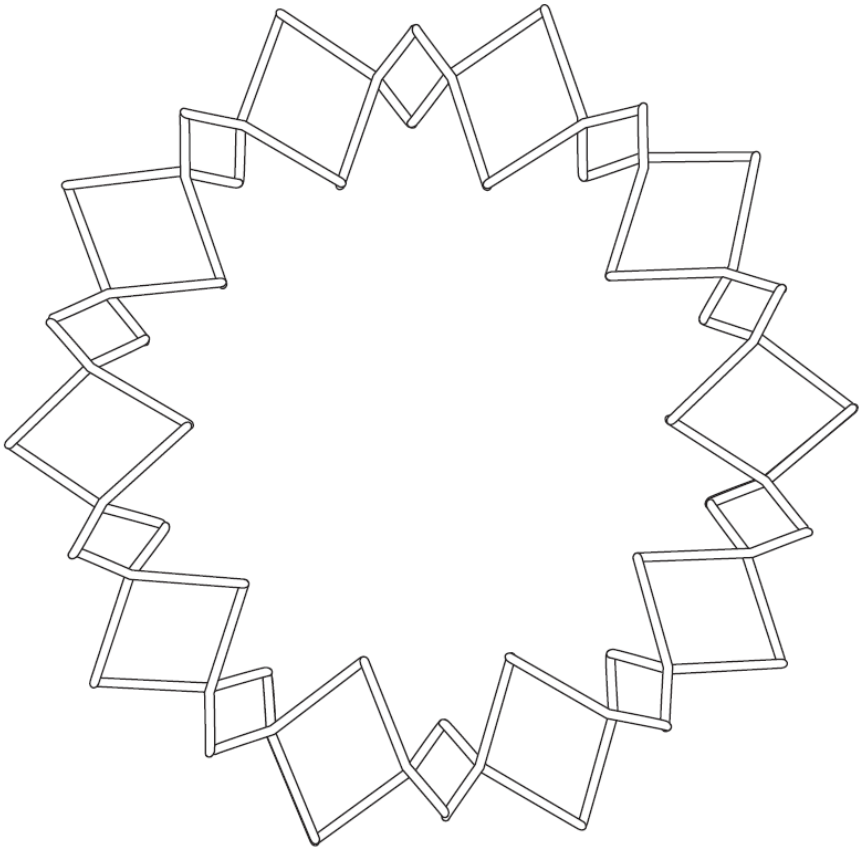


FIG 5A

FIG 5B



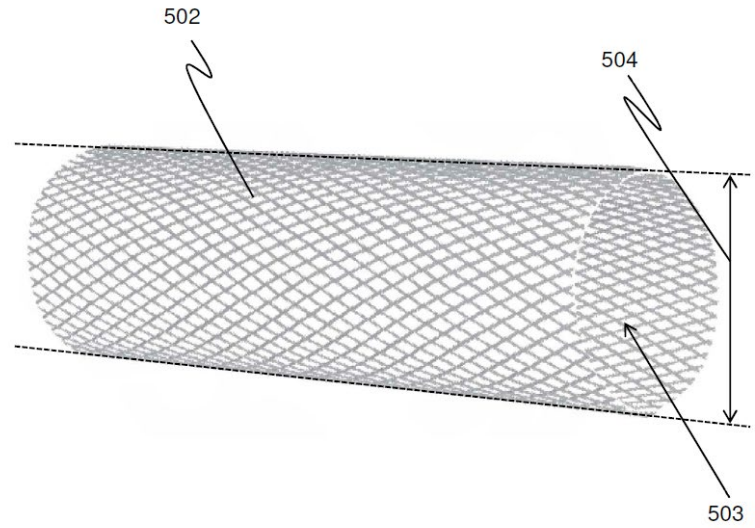


FIG 6

Detailed description of the figures

Fig. 1A shows a schematic view of a body with a rectum 102 as a body lumen in a cross-sectional view. Provided around the rectum 102 is an implantable device 100 for engaging with a body lumen, comprising a housing 104 with a housing lumen 106. Provided in the housing 104 is a lumen mechanism 108 with a mechanism lumen 110.

The mechanism lumen 110 may define a passage there through for a fluid and/or solid. As such, the mechanism lumen 110 may have a certain flow-through area for the passage, which may be increased or decreased by respectively expanding or retracting the mechanism lumen 110.

Figs. 1B and 1C show the implantable device 100 in a side view, wherein in Fig. 1B the device 100 is in an expanded state and in Fig. 1C the device 100 is in a contracted state, wherein the housing lumen 106 and mechanism lumen 110 are smaller than their respective counterparts in Fig. 1B.

When provided with an actuator, the implantable device 100 may provide a force to the rectum 102 such that the rectum 102 deforms from the expanded state of Fig. 1C into the contracted state of Fig. 1B. Alternatively or additionally, the actuator of the implantable device 100 may provide a force to the rectum 102 such that the rectum 102 deforms from the contracted state of Fig. 1B to the expanded state of Fig. 1C.

For exchanging forces between the body lumen 102 and the implantable device 100, the implantable device 100 comprises an engagement section. Through this engagement section, radial

expansion and retraction force may be transferred between the body lumen and the implantable device.

Examples of possible engagement sections are, when the housing 104 is at the housing lumen 106 connected to the body lumen 102, the part of the housing 104 connected to the body lumen 102. Such a connection may be constituted by a glue, stitching, by allowing the body to grow tissue between the implantable device 100 and the body lumen 102, by implanting the device 100 in a sphincter muscle or the surrounding fat, or any other method of connecting the implantable device 100 to the body lumen. The engagement section need thus not be a separate part of the implantable device 100, but may be formed by one of the other parts of the implantable device 100.

Fig. 2 shows a schematic view of a body with the rectum 102 as the body lumen in a front view. Implanted in the body is an implantable device combination 200 of a first implantable device 210 at a first location 201 and a second implantable device 220 at a second location 202, both implantable devices comprising a housing 104 with provided therein a lumen mechanism 108. In the implementation of Fig. 2, the second location 202 is downstream of the first location.

A human rectum is provided with an internal sphincter and an external sphincter, of which the first one is an involuntary sensor-like sphincter, and the latter one is a voluntary actuator-like sphincter. The combination of the two implantable devices may be used to mimic the workings of the two rectal sphincters.

The first implantable device 210 may be provided with a sensor 205 for determining lumen data on the mechanism lumen 110. As such, it may

detect the open state of the body lumen at a first location 201 due to faeces 203 passing through the body lumen.

Upon such a detection, the sensor 205 may send a signal to a user device 206. In response to receiving the signal, the user device 206 may provide an alert to the user making him aware of the open state of the body lumen at the first location 201.

A user device 206 may for example be a smartphone, pager, or dedicated portable device, and the device 206 may comprise signalling means using for example an audible and/or visual signal for alerting the user.

The second implantable device 220 is provided with an actuator 208, arranged to manipulate the mechanism lumen of the lumen mechanism 108. In the schematic Fig. 2, the second device 220 has contracted the rectum at the second location 202 by virtue of the actuator 208.

The actuator 208 may be provided with an input module for receiving a control signal, e.g. from the user device 206. Such a control signal may be used by a control module to operate the actuator 208 in accordance with the control signal. The user may, via the user device 206, provide a choice of when to activate opening of the second implantable device 220.

With the combination of the first implantable device 210, the second implantable device 220 and the user device 206, it becomes possible to allow a user to know when there is a need to empty their bowel, and to control when this should happen.

In an alternative embodiment, contrary to what is shown in Fig. 2, an implantable device may be provided with a single housing and a single lumen mechanism, and this lumen mechanism may be provided both with an actuator and a sensor. As such, it may be possible to first detect a state of the body lumen with the sensor, and operate the actuator in response to the detected state.

For example, when the sensor detects that the body lumen is moving from a contracted state to an expanded state, the actuator may assist in opening the body lumen by opening the mechanism lumen.

Fig. 3A shows an embodiment of the implantable device, wherein the lumen mechanism 108 comprises a foldable planar linkage 300 comprising a plurality of links connected by a plurality of joints. The linkage 300 encloses an internal area substantially corresponding to the mechanism lumen 110, indicated as the dashed circle. The linkage 300 is foldable between an unfolded position, as shown in Fig. 3A, and a folded position, as shown in Fig. 3B. The internal area enclosed by the planar linkage 300 is in the folded position of Fig. 3B smaller than in the unfolded position of Fig. 3A, and may be substantially zero in the folded position Fig. 3B such that substantially no passage of fluid, solid, and/or gas is possible through the body lumen 102.

Fig. 4A shows part of a planar linkage mechanism 300 as a lumen mechanism with a mechanism lumen 110 for an implantable device, and will be used to elaborate on some envisioned locations and manner of operation for actuators and/or sensors. Of the planar linkage mechanism 300, a first submechanism 301, a second submechanism 302, and a third submechanism are shown. Each submechanism comprises four joints and four links, wherein the links

are at ends connected to adjacent links with a joint. The three submechanism define a mechanism lumen 110 with a particular radius R.

Now looking at the first submechanism 301, as a dotted line a first location P1 for a sensor and/or actuator is shown. In this orientation, the sensor and/or actuator is oriented substantially tangentially, and spans between two joints referenced to as J1 and J2. Since the planar linkage mechanism 300 only has a single degree of freedom due to the specifically chosen link lengths, there is a direct relation between the distance between J1 and J2 and the radius R of the mechanism lumen 110.

If an actuator is provided at the first location P1, such an actuator may be used to control a distance between joints J1 and J2. For example, the distance between J1 and J2 may be decreased to obtain the situation as sketched in Fig. 4B. Here, the radius R' of the mechanism lumen 110 has decreased by virtue of the decrease in distance between J1 and J2. As such, an actuator controlling the distance between joints J1 and J2 at the first location P1 may be used to control the flow-through area through the mechanism lumen 110.

Next to being a possible location for an actuator, at the first location P1 also a sensor may be provided, wherein the sensor may be connected to joints J1 and J2. Connected in such a way, the sensor may be used to determine a distance and/or a force between the joints J1 and J2. Such a distance may be used to determine a state of opening of the lumen mechanism, and when engaged with the body lumen, a state of opening of the body lumen. Such a force may be used to determine a radial pressure on the lumen mechanism, a

pressure which may originate from a pressure exerted by the body lumen on the engagement section of the implantable device.

Now referring back to Fig. 4A, a second location P2 is shown at the second submechanism 302 indicating a second option for a location for a sensor and/or actuator. A sensor and/or actuator provided at the second location P2 may be spanned between two joints J3 and J4, in an orientation substantially radially relative to the mechanism lumen 110. The distance between J3 and J4 is directly related to radius R of the mechanism lumen 110, a relationship which may be used for controlling the radius R with an actuator or determining the radius R with a sensor.

If an actuator is used to increase the distance between joints J3 and J4 relative to the distance shown in Fig. 4A, a situation as sketched out in Fig. 4B may be obtained, wherein the radius R' has decrease relative to the radius R by virtue of the increase in distance between joints J3 and J4.

Similar to the first location P1, at the second location P2 a sensor may be provided. Now between joints J3 and J4, a sensor may be provided for measuring a distance and/or force between joints J3 and J4.

Again, referring back to Fig. 4A, a third location P3 is indicated by a dotted line as a location where a sensor and/or an actuator may be provided. The dotted line P3 refers to the angle between links L1 and L2, which are provided at a certain angle relative to one another. The angle between links L1 and L2 is directly related to the radius R of the mechanism lumen 110, and this relationship may thus be used for a sensor and/or an actuator.

If an actuator is provided at the third location P3, the angle between links L1 and L2 may become controllable, and with that the radius R may become controllable. If the actuator would enlarge the angle between links L1 and L2, a situation as sketched out in Fig. 4B may be obtained, wherein the radius R has shrunk to radius R' due to the increase in angle between links L1 and L2. At the third location P3, also a sensor may be employed for determining data on an angle and/or torque between links L1 and L2.

While figures 4A and 4B have been used to show some example locations for actuators and/or sensors, any other location is also envisioned where a actuator or sensor may be placed to control and/or detect radius R of the mechanism lumen 110.

When a distance is to be controlled, any type of actuator may be used arranged to manipulate a distance, such as linear actuators, which may be magnetically-driven, piston rod actuators, spindle drives, any other type of actuator or any combination thereof.

When an angle is to be controlled, the same types of actuators as described above may be used. Alternatively, a rotational actuator may be used for manipulating the angle between two adjacent links which are connected by the same joint.

For implanting an implantable device according to any of the embodiments as described herein, a method may be used wherein the implantable device is first manipulated into an opened state, figure 5A, wherein, instead of the housing comprising a housing lumen, figure 6, and the lumen mechanism comprising a mechanism lumen, both the housing and the lumen mechanism are opened up to form an opened,

for example C-like, shape. As such, the implantable device may be more easily provided around the body lumen.

Once the device is implanted, the ends of the C-shape are connected, figure 5B, to surround the body lumen. Several options are available for providing this connection. In one option, the ends of the C-shape of the housing are closed and connected. This means that ends of the lumen mechanism are not connected. In such implementation, the lumen mechanism is provided to maintain its C-shape even with the ends of the C-not connected. In another implementation, the ends of the C-shape of the housing are connected such that a toroid shape volume is provided (popularly known as a doughnut shaped volume). In this volume, the ends of the C-shaped lumen mechanism may be connected forming an O-shape, or not.

When provided around the body lumen, the housing and the lumen mechanism may be closed up to respectively form the housing lumen, figure 6, and the mechanism lumen, figure 5B. For example, when the lumen mechanism comprises the foldable planar linkage, one or more links may be disconnected at one or more joints such that the linkage may be provided around the body lumen. When the implantable device is then provided around the body lumen, the disconnected links may be reconnected to form the mechanism lumen and the housing may be closed up again to form the housing lumen.

For creating an appropriate space around the body lumen to position the implantable device in, it may be required to make space in the tissue surrounding the body lumen. For example, may a hooked needle be used for opening up tissue, such as fat tissue of muscle tissue, surrounding the body lumen.

After the implantable device has been implanted, i.e. it has been provided around the body lumen, tissue may grow back around the housing lumen. With this healed tissue, a force coupling may be established between the implantable device and the body lumen such that radial expansion and retraction forces may be coupled between the body lumen and the implantable device. Parts of the implantable device, such as the housing, may comprise materials which promote growth of body tissue on these parts of the implantable device which comprise such a material.

Fig. 6 shows a sheath 500 comprising a woven tube 502, which tube 502 defines an inner space 503 with a certain diameter 504. The tube 502 has a substantially cylindrical shape, and the inner space 503 is sufficiently large for accommodating one or more implantable devices. The tube 502 is woven such that it provides a certain degree of flexibility to allow it to be contracted – i.e. the diameter 504 decreases – and/or extended – i.e. the diameter 504 increases. A force required for the extension and/or contraction may be provided by an implantable device as discussed above.

The sheath 500 comprises a biocompatible material, and may as such be implanted in a human body together with the implantable device. The material comprised by the sheath 500, and the shape of the woven tube 502 may be chosen such to stimulate tissue growth on the sheath 500.

The sheath may be implanted around the implantable device and/or surrounding tissue, such that tissue may heal around the sheath 500. When tissue has grown around the sheath 500, and the implantable device has been provided within the inner space 503 of the sheath

500, the implantable device is enclosed in tissue and as such the enclosing tissue may be contracted and/or extended along with the device.

Claims

1. An implantable device for engaging with a body lumen, comprising:

- a flexible housing comprising a housing lumen arranged to be radially expanded and retracted between an opened and a closed position;

- a lumen mechanism provided in the housing, comprising a mechanism lumen which is substantially aligned with the housing lumen and arranged to be radially expanded and retracted between an opened position corresponding to the opened position of the housing and a closed position corresponding to the closed position of the housing; and

- an engagement section for engaging with the body lumen for providing a coupling between radial expansion and retraction activity of the body lumen and the implantable device.

2. Implantable device according to claim 1, wherein the lumen mechanism comprises a foldable planar linkage comprising a plurality of links connected by a plurality of joints, wherein the linkage encloses an internal area substantially corresponding to the mechanism lumen and wherein the linkage is foldable between a folded position and an unfolded position, wherein in the folded position the internal area enclosed by the linkage is smaller than in the unfolded position.

3. Implantable device according to claim 1 or 2, further comprising:

- an actuator for manipulating the lumen mechanism such that the mechanism lumen can be positioned between a closed position and an opened position, wherein the engagement section is arranged to transfer retraction and expansion forces from the actuator to the body lumen to which the implantable device is engaged.

4. Implantable device according to claim 3, wherein the actuator comprises two ends, the actuator is arranged to manipulate an actuator distance between the two ends and the actuator is:

   - at a first end connected to a first joint, which first joint connects a first link and a second link;

   - at a second end connected to a second joint, which second joint connects a third link and a fourth link;

   wherein manipulation of the actuator distance folds the planar linkage between the folded position and the unfolded position.

5. Implantable device according to claim 3, wherein the actuator is connected to a first link and a second link, wherein the first link is provided adjacent to the second link and the actuator is arranged for manipulating an angle between the first link and the second link.

6. Implantable device according to any of claims 3-5, further comprising an input module for receiving a control signal, and a control module for operating the actuator in accordance with the control signal.

7. Implantable device according to 1 or 2 further comprising:

- a sensor for determining lumen data on the mechanism lumen, wherein the engagement section is arranged to transfer retraction and expansion forces from the body lumen to which the implantable device is engaged to the sensor.

8. Implantable device according to claim 7, to the extent dependent on claim 2, wherein the sensor comprises two ends, and the sensor is arranged to obtain data on a distance and/or pressure between the two ends, and the sensor is:

- at a first end connected to a first joint, which first joint connects a first link and a second link;

- at a second end connected to a second joint, which second joint connects a third link and a fourth link.

9. Implantable device according to claim 7, to the extent dependent on claim 2, wherein the sensor is connected to a first link and a second link, wherein the first link is provided adjacent to the second link and the sensor is arranged to obtain data on an angle and/or force between the first link and the second link.

10. Implantable device according to any of the claims 7-9, wherein the sensor is arranged to output an output signal comprising distance/pressure data.

11. Kit of parts, comprising:

- a first implantable device according to any of the claims 3-6; and

- a second implantable device according to any of the claims 7-10.

12. Method for opening and closing an implantable device, comprising:

- by a user control device, receiving distance/pressure data from a distance/pressure sensor comprised by a first implantable device;
- by the user control device, outputting a user signal related to the received distance/pressure data;
- by the user control device, receiving a user input signal comprising actuation data;
- by the user control device, sending an actuation signal comprising actuation data to a second implantable device comprising a lumen actuator; and
- actuating the lumen actuator according to the received actuation data.

13. Method for opening and closing an implantable device, comprising:

- detecting, with a sensor provided by the implantable device, a change in body lumen opening state of a body lumen around which the implantable device is provided; and
- manipulating, with an actuator provided by the implantable device, a mechanism lumen in accordance with the change in body lumen opening state.

Conclusions

Current treatment options for anal incontinence leave much to be desired in terms of long-term quality of life. Incontinence due to HSCR, anorectal malformation, spinal cord injury or other neuromuscular sphincter malfunction can affect young patients throughout their lives. Additionally, age-related incontinence affects a substantial proportion of patients over 65. Here, we propose an electronic anal sphincter prosthesis as a possible treatment option that may be suitable to a large number of patients and could mimic anatomical function. The development of this device and surgical techniques involved to deliver it safely to patients is likely to be more cost-effective, less labour-intensive and suitable to a wider range of patients than the development of allogenic cell-transplantation therapy.

References

1. Holschneider AM, Puri P. *Hirschsprung's disease and allied disorders.* Springer Science & Business Media; 2007.
2. De la Torre L, Ortega A. Transanal versus open endorectal pull-through for Hirschsprung's disease. *Journal of Pediatric Surgery.* 2000;35(11):1630-1632.
3. Neuvonen MI, Kyrklund K, Rintala RJ, Pakarinen MP. Bowel Function and Quality of Life After Transanal Endorectal Pull-through for Hirschsprung Disease. *Ann Surg.* 2017;265(3):622-629.
4. Tannuri ACA, Tannuri U, Romão RLP. Transanal endorectal pull-through in children with Hirschsprung's disease—technical refinements and comparison of results with the Duhamel procedure. *Journal of Pediatric Surgery.* 2009;44(4):767-772.
5. Cheng LS, Schwartz DM, Hotta R, Graham HK, Goldstein AM. Bowel dysfunction following pullthrough surgery is associated with an overabundance of nitrergic neurons in Hirschsprung disease. *Journal of Pediatric Surgery.* 2016;51(11):1834-1838.
6. Onishi S, Nakame K, Kaji T, et al. The bowel function and quality of life of Hirschsprung disease patients who have reached 18 years of age or older – the long-term outcomes after undergoing the transabdominal soave procedure. *Journal of Pediatric Surgery.* 2017;52(12):2001-2005.
7. Fattorini E, Brusa T, Gingert C, et al. Artificial Muscle Devices: Innovations and Prospects for Fecal Incontinence Treatment. *Ann Biomed Eng.* 2016;44(5):1355-1369.
8. Macmillan AK, Merrie AE, Marshall RJ, Parry BR. The prevalence of fecal incontinence in community-dwelling adults: a systematic review of the literature. *Dis Colon Rectum.* 2004;47(9):1341-1349.
9. Stoker J, Halligan S, Bartram CI. Pelvic floor imaging. *Radiology.* 2001;218(3):621-641.
10. Dey AN. *Characteristics of elderly nursing home residents: data from the 1995 National Nursing Home Survey.* US Department of Health and Human Services, Centers for Disease Control and ...; 1997.
11. Saga S, Vinsnes AG, Mørkved S, Norton C, Seim A. Prevalence and correlates of fecal incontinence among nursing home residents: a population-based cross-sectional study. *BMC Geriatr.* 2013;13(1):87.
12. Glickman S, Kamm MA. Bowel dysfunction in spinal-cord-injury patients. *The Lancet.* 1996;347(9016):1651-1653.
13. Singh A, Tetreault L, Kalsi-Ryan S, Nouri A, Fehlings MG. Global prevalence and incidence of traumatic spinal cord injury. *Clin Epidemiol.* 2014;6:309-331.

14. Krassioukov A, Eng JJ, Claxton G, Sakakibara BM, Shum S. Neurogenic bowel management after spinal cord injury: a systematic review of the evidence. *Spinal Cord.* 2010;48(10):718-733.
15. Fritsch H, Brenner E, Lienemann A, Ludwikowski B. Anal Sphincter Complex. *Dis Colon Rectum.* 2002;45(2):188-194.
16. Costa M, Brookes SJH, Hennig GW. Anatomy and physiology of the enteric nervous system. *Gut.* 2000;47(suppl 4):iv15-iv19.
17. Snooks S, Swash M. The innervation of the muscles of continence. *Ann R Coll Surg Engl.* 1986;68(1):45.
18. Rao SS. Pathophysiology of adult fecal incontinence. *Gastroenterology.* 2004;126:S14-S22.
19. Peña A, Guardino K, Tovilla J, Levitt M, Rodriguez G, Torres R. Bowel management for fecal incontinence in patients with anorectal malformations. *Journal of pediatric surgery.* 1998;33(1):133-137.
20. Norton C, Whitehead W, Bliss D, Harari D, Lang J. Management of fecal incontinence in adults. *Neurourology and Urodynamics: Official Journal of the International Continence Society.* 2010;29(1):199-206.
21. Meinds RJ, van der Steeg AFW, Sloots CEJ, et al. Long-term functional outcomes and quality of life in patients with Hirschsprung's disease. *Br J Surg.* 2019;106(4):499-507.


7

Chapter 7

General Discussion

Hirschsprung disease (HSCR) is a rare congenital malformation of the enteric nervous system (ENS) characterised by an aganglionosis in a variable length of the distal gastrointestinal (GI) tract. Aganglionosis results in constriction of the smooth muscle in the affected region, leading to a functional obstruction[1]. During development, the cells that form the ENS bud from the neural crest and rapidly proliferate and migrate along the developing gut tube. These neural crest cells (NCCs) eventually form the enteric ganglia and differentiate into neuronal subtypes and glial cells located in the submucosal and myenteric plexuses of the ENS. A disturbance to proliferative, migratory, differentiative and/or survival functions in these cells could contribute to the pathogenesis of HSCR[2]. The current treatment for HSCR consists of surgical resection of the affected gut region, however most patients continue to have prolonged GI tract complications following surgery[3,4]. Understanding the development and pathogenesis of HSCR is vital to provide improved treatment options for these patients.

Factors involved in the development of HSCR

The disruption of ENS development is thought to be broadly under genetic control. A number of key NCC regulating genes have been implicated in HSCR pathogenesis. However, pathogenic variants in these genes only explain ~50% of HSCR cases[2]. It is therefore likely that other strong pathogenic variants in yet unknown genes, and/or combinations of pathogenic variants and weaker modifying variants, contribute to the aetiology and pathogenesis of HSCR[5]. Of the genes known to be involved in HSCR pathology, *RET* is understood to be the major influencer, acting as a fulcrum in the balance between aganglionosis and hyperganglionosis[5]. If there is a strong *RET* or other causal variant then the presence or absence of risk alleles[6-8] at common

SNP locations is unlikely to influence disease development substantially. The *RET* balance is therefore more representative in cases where a combination of variants and other risk factors have the cumulative effect of a HSCR phenotype.

If a patient presents with a multi-feature syndromic form of HSCR then the risk alleles may slightly influence presence or absence of HSCR, but more likely large copy number variations and chromosomal displacement affect its presence or absence. It could also be that combinations of missense variants and modifying SNPs have a cumulative influence on pathology in both non-syndromic and syndromic forms of HSCR.

Genetics of isolated HSCR

In non-syndromic isolated HSCR cases *RET* is the major genetic risk factor. However, the inheritance in sporadic non-syndromic HSCR cases is considered complex[2]. In this group, many genes other than *RET* have been identified. The identified genes are primarily involved in either the *RET* signalling pathway or the endothelin signalling pathway, and it is likely that other up- or down-stream influencers of these are yet to be identified. Due to the frequency of sporadic isolated cases and the unexplained genetic origin in many familial cases, it can be safely assumed that there are further disease genes yet to be identified, and/or factors other than the genomic DNA sequence of a patient that may play a role in disease pathogenesis[9,10]. Susceptibility to HSCR could be further influenced by stochastic effects on gene expression, additional variants in other genes that influence regulatory elements, and environmental effects on gene expression[6]. Additionally, the effects of

variants may be subject to epigenetic factors and changes in methylation patterns, which could also affect the complex heritability.

Genetics of unexplained syndromic HSCR

Approximately 12% of HSCR cases are associated with chromosomal anomalies, and approximately 18% of cases present in combination with other defects or features[11]. In contrast to non-syndromic HSCR, in syndromic HSCR we assumed that large copy number variations (CNVs) could explain part of the missing heritability. In **chapter 3** we discuss CNV and how large CNVs may influence HSCR pathology if they overlap dose-sensitive genes that affect ENS development. HSCR can present as one symptom in patients with multiple associated anomalies, these can be recognised syndromes or newly presenting multi-feature presentations[11,12]. In multi-feature patients with known strong pathogenic variants CNVs have little influence on HSCR development. We demonstrate that CNVs in multi-feature patients, without a known pathogenic variant, tend to be longer and affect regions that contain genes expressed in the developing ENS. This data needs to be replicated in independent cohorts to confirm if candidate genes within large CNVs are seen in multiple patients. Investigations in zebrafish models, to knock down these genes or express multiple copies, may help to confirm new candidate genes present within CN loss/gain regions.

Missense variations in known syndromic HSCR

Missense variants can go unnoticed and be present in the healthy population, but still disturb protein folding, binding or other functions. They can have the same pathogenicity as a loss of function variant, specifically in recessive diseases in which carriership does not result in a disease phenotype. In **chapter 4** the identification of missense variants

in Goldberg-Shprintzen syndrome (GOSHS), as well as the presence or absence of HSCR as a variable clinical feature in GOSHS is discussed. Truncating variants in the KIF1 binding protein gene (*KIF1BP*) are known to cause GOSHS[13]. The presentation of two patients with missense variants in *KIF1BP* was interesting given the lack of knowledge of protein folding and interactions. There is no crystal structure available of KIF1BP and the interactive sites are not well understood. Given that loss-of-function variants have been reported in all 7 exons, it is likely that there are many regions, even in the terminal regions of exon 7 that are vital to RNA or protein stability. The two patients presented with different phenotypes and the finding of missense *KIF1BP* variants in patient NL1 was surprising, given their clinical presentation. The homozygosity of the missense variant in patient CYP3, or the region affected, may have been a tipping factor in the development of classical GOSHS with HSCR. The expression of KIF1BP in CYP3 is lower than in NL1, despite the compound heterozygous variants in exons 1 and 7. A CNV analysis in these patients, particularly patient NL1, may be informative to determine if there are other candidate regions that have influenced neural development. As no other pathogenic variant was identified in diagnostic screening of patient NL1, and given their syndromic features, it may be that CNV of dose dependent genes contributes to their neurological phenotypes. This is the case in HSCR patients with associated syndromic features and would be interesting to investigate in other multi-feature patients.

Missing heritability in HSCR

As previously mentioned, in many HSCR cases a genetic cause for the disease cannot yet be identified. This can be partly explained by the complexity of inheritance patterns, presence of low-penetrant non-

coding variants and the influence of risk alleles[14]. In **chapter 2** we discussed somatic mosaicism as an influencer of HSCR, and the difficulty of identifying this phenomenon in relation to a HSCR phenotype. The developmental patterning of the ENS is such that a variant originating early has a high chance of being out-competed by other "healthy" cells. Unless this somatic variant gives a competitive advantage, such as higher proliferation or migration rate. This could lead to these cells reaching the distal colon and being unable to differentiate to the correct lineages and/or survive. We conclude that it is therefore extremely difficult to prove that true somatic variants contribute to HSCR aetiology, but it cannot totally be ruled out as a mode of missing heritability or a cause of some sporadic cases.

As somatic mutations appear unlikely to play a major role in HSCR, the question of what might explain these unsolved cases remains. A possible problem could be the genes that are selected as being candidates for HSCR. Current filtering criteria selects genes that have a clear role in neuronal development, are expressed in the developing central nervous system (CNS) or ENS, or expressed in NCCs[15]. It can be assumed that there are HSCR causing variants that are not expressed in ENCCs themselves, but in other developing GI tract tissues. Variants in genes expressed in smooth muscle or connective tissue could change the local gut environment through which ENCCs migrate[16]. These changes could leave the distal colonic segment unable to support the migration, incorporation or survival of the ENCCs, an example of this is variation in *EDN3*[17,18].

Moreover, the timing of cell-cycle exit has been shown to be of importance to ENS subtype specification[19], the current filtering criteria

could also exclude cell cycle genes that may influence ENS differentiation. In order to fully understand the enteric neural subtypes derived from ENCCs it would be beneficial to utilise single-cell RNA sequencing technologies to build a control database from healthy gut at various developmental stages as well as postnatally. This could be built of the ENS as well as other GI tract cell types. Once a baseline of expression at various points along the GI tract is established then comparisons can be made with different disease states. This could also help with the identification of causal genes and further understand links in the developmental pathways that are disrupted in HSCR development.

Non-genetic influencing factors

Other than inherited and somatic alterations, external factors can also influence ENS development by changing the epigenetic landscape[9,20]. Using a similar strategy to that mentioned above, it would be beneficial to create a control database for the methylation state of various regions and developmental stages of the GI tract. Initial efforts could focus on isolation of ENCCs from control gut in order to establish methylation patterns of the ENS. Eventually other cell types and full gut sections could be included to be able to compare HSCR patient tissue and find large methylation changes. The creation of expression maps of the ENS, both with RNA and methylation patterns, would be pivotal to linking known pathways together and finding new players in the network of ENS development.

The ability of clinical geneticists to give reliable genetic counselling to the families of patients relies on an understanding of the heritability of the condition. This is dependent on a knowledge of the genetic

background of HSCR pathogenesis as well as the non-genetic influencing factors. However, it may not be possible to explore every avenue, and there will still be sporadic cases that are difficult to explain.

Modelling HSCR

Proliferation, migration, differentiation and survival of ENCCs is a major focus of research and deviations of these processes are shown to cause aganglionosis in animal models of HSCR[21]. There are cases where the causative variant is known to be disruptive in these mechanisms and is shown to have expression in ENCCs specifically. However, for many patients the remaining ganglionic gut functions to a manageable degree and no other physiological problems are reported other than malfunction of the anal sphincter region. Animal models for developmental disease can only truly be useful if the genetic variant is known, and there is a practical limit to the number of genes that can be investigated in one model. The use of patient-specific cells to create a model circumvents this issue as the genetic background does not need to be fully known to functionally test for defects.

In **chapter 5** the creation of induced pluripotent stem cells (iPSCs) from HSCR patients was presented. These iPSC lines each harbour variants for known causative HSCR genes and were investigated for differences in function that may influence their ability to rescue the ENS using cell transplantation strategies. These iPSCs could also be utilised for disease modelling at the cellular level.

The development of a reliable *in vitro* model for HSCR might help to dissect small cellular changes in ENCCs and/or the local gut environment. A 3D organ-on-chip system for GI tract development and function would be ideal to investigate the interplay of the different cell-

types in the gut that are necessary for normal ENS development. Initial establishment of such a system with iPSC lines from healthy controls will allow the optimisation of cell-type ratios and flow of nutrients. Such a model, and generation of iPSCs from patients, will allow for the investigation of individual patient gut and help to determine the functions that may be disrupted during development. Moreover, it will also be help to determine whether ENCCs, smooth muscle, mucosal interaction or extracellular matrix (ECM) components were altered. Although gut organoids also offer a method of creating a 3D gut-like environment for cell transplantations, the spheroid structure is random which will influence intercellular signalling[22]. Gut organoids form a self-organising lumen system which creates multiple signalling gradients that are difficult to compare between organoids. An organ-on-chip system would offer a more organised and replicable model and is already being used with mucosal models to create an epithelial layer that can support a microbiota[23].

Cell transplantation therapy

As discussed in **chapter 5**, the possibilities for cell replacement therapy in HSCR have been explored for a number of years. The ability of transplanted cells to find the appropriate positioning, into either submucosal or myenteric plexus, and to form functional connections, both with other transplanted cells and with target cells within the gut, is vital to transplantation success. The appropriate cell type for this purpose, and an optimal transplantation strategy still need to be determined when scaling up from mouse to human gut[24]. No human trials have yet been attempted, but initial transplants into HSCR mouse models have been encouraging. Both injection of postnatal enteric neural crest cells (ENCCs)[25,26] and of more proliferative pluripotent stem

cells (PSCs) differentiated towards neural crest[27] have shown an integration and spread of transplanted cells.

Initial transplantation studies established the potential of postnatally derived ENCCs to integrate and survive in explanted colonic segments[28,29]. *In vivo* transplantation of ENCCs of embryonic and postnatal origin has been shown to lead to the engraftment of donor-derived cells within recipient colon[25,30]. Additionally, it was shown that ENCC-derived neurons adopt the appropriate localisation within the gut and can give rise to various enteric neurons, including the main subtypes for excitation (ChAT, VAChT, Calretinin and Calbin-din) and relaxation (nNOS and VIP)[25,30]. The transplanted cellular networks were also shown to closely localise with the endogenous ENS, suggesting functional integration of the transplanted neurons

Methods of transplantation

Current protocols for transplantation of cells into *in vivo* gut involves injection of cells in suspension with saline or matrigel[26,27]. Injection of ENCCs has proven safe in longer term follow-up and no migration of cells to ectopic sites was observed. Using PSCs, it is uncertain how many of the injected cells remain at the injection site and which other locations cells may reach. It is a recognised pattern following injection of PSCs to other organs that, although beneficial effects may be seen in the target tissue, cells are found in other organs[31]. It is also likely that many injected cells will die before making cell-cell contact[32]. Therefore, the already high numbers of cells required with this method, given the size of the target organ, may be even higher than anticipated when accounting for cell viability. Cellular scaffolds, injectable gels or other devices may make the environment more amenable to cell invasion,

lead to easier introduction of the cells, increase cell survival and avoid the migration of cells beyond the desired location[33,34].

Over-invasion and mass migration and proliferation is a known issue in PSC transplantation[31]. The study by Fattahi et al., considered as a landmark paper, shows transplantation of PSC-derived NCCs to wild-type and Ednrb$^{-/-}$ mouse models[27]. The study shows a promising start to integration of PSC-derived NCCs and migration of transplanted cells along the GI tract. However, it has yet to be determined whether these cells also migrate to other regions in the body, especially connecting abdominal organs and this was not investigated with the reported transplantations. The extensive migration that they present is contrary to ENCC transplants reported in literature, which show a more modest migration to form small ENS-like plexuses in mouse gut[25,26,28,30,35,36]. The interesting factor in these studies is how many injection sites would be necessary in the human gut in order to effectively form a functional ENS. The highly proliferative nature of the PSC-derived cells could circumvent this, although it is vital to ensure that all transplanted cells are adequately differentiated to at least a multi-potent single germ lineage progenitor state as opposed to retaining ability to form other lineages. The slowing of the cell cycle to a point of normal turnover for tissue maintenance is required after the desired integration of cells, otherwise the tumour-risk from the transplanted cells is higher[37].

Numbers matter

It has been shown that in development critical threshold numbers of NCCs are required for full colonisation of the GI tract[38]. It is also known that cells benefit in culture from contact with other cells producing "friendly cytokines"[39]. A similar phenomenon may translate in

transplantation where both adequate cell number and density are vital to transplantation success. A disadvantage of the significant cell expansion needed to create large numbers of cells is the propensity of cells to acquire genetic and epigenetic changes upon long-term culture and expansion[40-42]. Such changes may reduce the efficacy of generating specific cell derivatives, or could potentially compromise safety, for example by promoting tumour growth. ENCCs offer less tumorigenic risk than pluripotent alternatives[37], and a more primed ability to form these enteric neural subtypes directly[25]. However, they may not be proliferative enough for expansion to required numbers[43]. Current culture methods do not generate enough cells for both characterisation and transplantation. Pluripotent cell types expand more quickly in culture than postnatal stem cells, which would decrease expansion time for generating sufficient numbers.

Safety of Cell Transplantation

Human PSCs are becoming more popular as a therapeutic tool and are currently being investigated in clinical trials for a number of conditions, including macular degeneration, spinal cord injury, diabetes, heart disease and Parkinson's disease[44]. However, before these cells can safely be used in routine therapies, a better understanding of their behaviour, and understanding the possible genetic changes that may have occurred during their processing, is required.

Safety discussions to date have focused mainly on possibility of teratoma formation from transplanted cells, migration of cells beyond the tissue of interest and the occurrence of genetic variation arising during cell culture. To mitigate the possibility of teratoma development from transplanted cells the appropriate differentiation stage would

need to be reached. Cell therapy should consist of administering progenitors or differentiated derivatives rather than undifferentiated stem cells, and the accidental transplantation of undifferentiated cells should be avoided[37]. With the appropriate checks and characterisation of cells the chance of teratomas would be minimal. Another safety issue is the spread of the transplanted cells outside the tissue of interest, as discussed briefly above. When cells integrate into non-target tissues, possibly in combination with (epi)genetic changes, this could have profound consequences[31]. Monitoring the spread of cells is therefore crucial and the investigation into methods to ensure that transplanted cells remain at the transplantation site is necessary.

The primary focus of attention should be on the potential (epi)genetic changes that may have arisen during creation and culture of PSCs. It is these somatic (epi)genetic changes that may have a substantial impact on the behaviour of the PSCs and may even lead to malignant transformation of the mutated cells[45]. Monitoring the PSCs for such genetic changes is therefore crucial. Discussions are still ongoing on how to screen cells and how to interpret the results in order to evaluate their significance for the safety of therapeutic applications[46]. Due to the risks associated with significant cell expansion *in vitro*, the time in culture and number of passages of cells should be kept as low as possible. Cells should be checked genetically, preferably by exome sequencing, before use. As long as the effects of individual genetic variants on the PSCs or differentiated cell types are uncertain, and while cells are likely to spread from the target tissue, the introduction of a conditional suicide gene could provide a fail-safe strategy for eliminating cells after transplantation if a problem were to arise[46].

Other challenges to overcome

Eventually, with a greater understanding of the development of enteric neural subtypes at all levels of GI tract development, reliable induction of ENCC-specific differentiation will be possible with iPSC lines. A non-integrating viral transduction system to iPSCs will further reduce the inter-line variability and allow for the patient specific defects to be more readily compared. The therapeutic potential of iPSCs is great, they are arguably the most valuable tool for personalised regenerative medicine, but for their full potential to be realised it will be necessary to recognise and correct for their disadvantages.

As the gut receiving transplantation has developed without ganglia the ECM and cellular environment of the tissue may be less receptive to ENS cell transplantation. As mentioned above, the creation of a suitable transplant agent or cell scaffold system may create a more receptive environment for the transplanted cells[33,47]. It is yet to be established if cytotoxicity from inflammation or fibrotic regions of scar formation at the anastomotic region in HSCR could make tissues less permissive to cell transplant invasion. In spinal cord injury the fibrotic tissue is known to be an issue for transplantation and scar ablation is common prior administration of cells[48].

Multi-disciplinary Treatment for HSC

Current standards of care for HSCR patients leave many with poor anal sphincter control[4,49], and this has been highlighted as a target region for initial cell therapy[50]. Given the safety and other challenges with the use of cell therapy that are yet to be overcome, such an approach is unlikely to reach the clinic for a number of years, and could have variable success rates within patients. In **chapter 6**, we introduce a novel treatment

possibility with the description of a device that could mimic the physiological function of the anal sphincter. In HSCR patients the constriction of the sphincter can lead to chronic constipation, and the damage to the musculature from surgery can lead to continued incontinence and anal leakage. There is need for a solution to allow for sphincter control for both opening and closing. Current solutions for this problem focus on closure, with no solution that will address both contraction and relaxation of the muscles[51].

The anal sphincter complex is comprised of the internal anal sphincter (IAS) and the external anal sphincter (EAS). It is the combined function of these sphincters that allows normal physiological defecation[52]. Function of the IAS is entirely involuntary and controlled by parasympathetic nerve fibres and the myenteric plexus whereas the EAS is mainly innervated by the pudendal nerve and is under voluntary control[53]. During defecation the increased pressure and stretch from bowel contents on the internal sphincter signals for it to relax, which is picked up by sensory nerves more distally and by the EAS. The pressure is sensed and these combined signals let us know that we need to defecate. The voluntary control of the EAS allows us to choose when to defecate. The lack of this control is the primary cause of adult faecal incontinence, and a number of pathologies can contribute to this[54]. The proposed solution of an electronic prosthetic anal sphincter is a multi-disciplinary approach, combining surgical techniques with microelectronics and physiological function in order to provide patients with a modern and convenient solution to anatomical malfunction.

Conclusions

HSCR development is complex and multifaceted. Although the overarching *RET* pathway can account for much of HSCR pathophysiology, there remains much to be learned from genetic studies. Collection of patient material and inclusion of parents and other family members in sequencing and functional cell-based research approaches will be instrumental in discovering new candidate genes and pathways that contribute to HSCR aetiology. The advancing of technologies to reliably sequence small amounts of DNA, and identification of methylation and other epigenetic marks will help to answer many questions that remain concerning the missing heritability observed in familial and sporadic HSCR cases. This, together with the understanding of HSCR as a variable symptom in multi-feature and syndromic cases, will enable more informative genetic counselling to patients and their families.

Treatment options for HSCR have been stagnant for the past decades, and ongoing research is necessary before novel cell therapy approaches can be applied in a clinical setting. The differentiation of PSCs to appropriate lineages is progressing quickly and, given a concerted effort in safety and efficacy trials, clinical application is approaching. A further possible option that may be suitable to more patients is the use of an electronic prosthetic that could mimic anatomical function. In the future this could be used in combination with cell transplantation therapy and allow for the training of surrounding tissue and conditioning of transplanted cells to function within the anal sphincter complex.

References

1. Holschneider AM, Puri P. *Hirschsprung's disease and allied disorders.* Springer Science & Business Media; 2007.
2. Butler Tjaden NE, Trainor PA. The developmental etiology and pathogenesis of Hirschsprung disease. *Translational research : the journal of laboratory and clinical medicine.* 2013;162(1):1-15.
3. Neuvonen MI, Kyrklund K, Rintala RJ, Pakarinen MP. Bowel Function and Quality of Life After Transanal Endorectal Pull-through for Hirschsprung Disease. *Ann Surg.* 2017;265(3):622-629.
4. Onishi S, Nakame K, Kaji T, et al. The bowel function and quality of life of Hirschsprung disease patients who have reached 18 years of age or older – the long-term outcomes after undergoing the transabdominal soave procedure. *Journal of Pediatric Surgery.* 2017;52(12):2001-2005.
5. Brosens E, Burns AJ, Brooks AS, et al. Genetics of enteric neuropathies. *Developmental Biology.* 2016;417(2):198-208.
6. Chatterjee S, Kapoor A, Akiyama JA, et al. Enhancer variants synergistically drive dysfunction of a gene regulatory network in Hirschsprung disease. *Cell.* 2016;167(2):355-368.e310.
7. Emison ES, McCallion AS, Kashuk CS, et al. A common sex-dependent mutation in a RET enhancer underlies Hirschsprung disease risk. *Nature.* 2005;434:857.
8. Sancandi M, Griseri P, Pesce B, et al. Single nucleotide polymorphic alleles in the 5′ region of the RET proto-oncogene define a risk haplotype in Hirschsprung's disease. *Journal of Medical Genetics.* 2003;40(9):714-718.
9. Torroglosa A, Alves MM, Fernández RM, Antiñolo G, Hofstra RM, Borrego S. Epigenetics in ENS development and Hirschsprung disease. *Developmental Biology.* 2016;417(2):209-216.
10. D'Gama AM, Walsh CA. Somatic mosaicism and neurodevelopmental disease. *Nature Neuroscience.* 2018:1.
11. Amiel J, Sproat-Emison E, Garcia-Barcelo M, et al. Hirschsprung disease, associated syndromes and genetics: a review. *Journal of Medical Genetics.* 2008;45(1):1-14.
12. Brooks A. *Genetics of syndromic and non-syndromic Hirschsprung disease*. Enschede: Clinical Genetics, Erasmus University Rotterdam; 2005.
13. Brooks AS, Bertoli-Avella AM, Burzynski GM, et al. Homozygous Nonsense Mutations in *KIAA1279* Are Associated with Malformations of the Central and Enteric Nervous Systems. *The American Journal of Human Genetics.* 2005;77(1):120-126.
14. Bergeron KF, Silversides DW, Pilon N. The developmental genetics of Hirschsprung's disease. *Clinical genetics.* 2013;83(1):15-22.
15. Schriemer D, Sribudiani Y, Ijpma A, et al. Regulators of gene expression in Enteric Neural Crest Cells are putative Hirschsprung disease genes. *Developmental biology.* 2016.
16. Raghavan S, Bitar KN. The influence of extracellular matrix composition on the differentiation of neuronal subtypes in tissue engineered innervated intestinal smooth muscle sheets. *Biomaterials.* 2014;35(26):7429-7440.
17. Watanabe Y, Stanchina L, Lecerf L, et al. Differentiation of Mouse Enteric Nervous System Progenitor Cells Is Controlled by Endothelin 3 and Requires

Regulation of Ednrb by SOX10 and ZEB2. *Gastroenterology.* 2017;152(5):1139-1150.e1134.
18. Puffenberger EG, Hosoda K, Washington SS, et al. A missense mutation of the endothelin-B receptor gene in multigenic hirschsprung's disease. *Cell.* 1994;79(7):1257-1266.
19. Bergner AJ, Stamp LA, Gonsalvez DG, et al. Birthdating of myenteric neuron subtypes in the small intestine of the mouse. *J Comp Neurol.* 2014;522(3):514-527.
20. Tang W, Li B, Tang J, et al. Methylation analysis of EDNRB in human colon tissues of Hirschsprung's disease. *Pediatric Surgery International.* 2013;29(7):683-688.
21. Bondurand N, Southard-Smith EM. Mouse models of Hirschsprung disease and other developmental disorders of the enteric nervous system: Old and new players. *Developmental Biology.* 2016.
22. McCracken KW, Cata EM, Crawford CM, et al. Modelling human development and disease in pluripotent stem-cell-derived gastric organoids. *Nature* 2014;advance online publication:1-19.
23. Kim HJ, Huh D, Hamilton G, Ingber DE. Human gut-on-a-chip inhabited by microbial flora that experiences intestinal peristalsis-like motions and flow. *Lab on a Chip.* 2012;12(12):2165-2174.
24. McCann CJ, Borrelli O, Thapar N. Stem cell therapy in severe pediatric motility disorders. *Current Opinion in Pharmacology.* 2018;43:145-149.
25. Cooper JE, McCann CJ, Natarajan D, et al. In Vivo Transplantation of Enteric Neural Crest Cells into Mouse Gut; Engraftment, Functional Integration and Long-Term Safety. *PLoS ONE.* 2016;11(1):e0147989.
26. Cooper JE, Natarajan D, McCann CJ, et al. In vivo transplantation of fetal human gut-derived enteric neural crest cells. *Neurogastroenterology & Motility.* 2016.
27. Fattahi F, Steinbeck JA, Kriks S, et al. Deriving human ENS lineages for cell therapy and drug discovery in Hirschsprung disease. *Nature.* 2016;531(7592):105-109.
28. Almond S, Lindley RM, Kenny SE, Connell MG, Edgar DH. Characterisation and transplantation of enteric nervous system progenitor cells. *Gut.* 2007;56(4):489-496.
29. Lindley RM, Hawcutt DB, Connell MG, et al. Human and mouse enteric nervous system neurosphere transplants regulate the function of aganglionic embryonic distal colon. *Gastroenterology.* 2008;135(1):205-216. e206.
30. Hotta R, Stamp LA, Foong JPP, et al. Transplanted progenitors generate functional enteric neurons in the postnatal colon. *The Journal of clinical investigation.* 2013;123(3):1182-1191.
31. van den Akker F, Feyen DA, van den Hoogen P, et al. Intramyocardial stem cell injection: go (ne) with the flow. *Eur Heart J.* 2017;38(3):184-186.
32. Aguado BA, Mulyasasmita W, Su J, Lampe KJ, Heilshorn SC. Improving viability of stem cells during syringe needle flow through the design of hydrogel cell carriers. *Tissue engineering Part A.* 2012;18(7-8):806-815.
33. Drury JL, Mooney DJ. Hydrogels for tissue engineering: scaffold design variables and applications. *Biomaterials.* 2003;24(24):4337-4351.
34. O'Cearbhaill ED, Ng KS, Karp JM. Emerging Medical Devices for Minimally Invasive Cell Therapy. *Mayo Clin Proc.* 2014;89(2):259-273.

35. Hetz S, Acikgoez A, Voss U, et al. In Vivo Transplantation of Neurosphere-Like Bodies Derived from the Human Postnatal and Adult Enteric Nervous System: A Pilot Study. *PLoS ONE.* 2014;9(4):e93605.
36. Li W, Huang L, Zeng J, et al. Characterization and transplantation of enteric neural crest cells from human induced pluripotent stem cells. *Mol Psychiatry.* 2016.
37. Ben-David U, Benvenisty N. The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nature Reviews Cancer.* 2011;11(4):268.
38. Barlow AJ, Wallace AS, Thapar N, Burns AJ. Critical numbers of neural crest cells are required in the pathways from the neural tube to the foregut to ensure complete enteric nervous system formation. *Development.* 2008;135(9):1681-1691.
39. Hague A, Hicks D, Bracey T, Paraskeva C. Cell-cell contact and specific cytokines inhibit apoptosis of colonic epithelial cells: growth factors protect against c-myc-independent apoptosis. *Br J Cancer.* 1997;75(7):960.
40. Baker DE, Harrison NJ, Maltby E, et al. Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nat Biotechnol.* 2007;25(2):207.
41. Amps K, Andrews PW, Anyfantis G, et al. Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. *Nat Biotechnol.* 2011;29(12):1132.
42. Draper JS, Smith K, Gokhale P, et al. Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nat Biotechnol.* 2004;22(1):53.
43. Metzger M, Bareiss PM, Danker T, et al. Expansion and differentiation of neural progenitors derived from the human adult enteric nervous system. *Gastroenterology.* 2009;137(6):2063-2073. e2064.
44. Trounson A, DeWitt ND. Pluripotent stem cells progressing to the clinic. *Nature Reviews Molecular Cell Biology.* 2016;17:194.
45. Goldring CE, Duffy PA, Benvenisty N, et al. Assessing the safety of stem cell therapeutics. *Cell stem cell.* 2011;8(6):618-628.
46. Andrews PW, Ben-David U, Benvenisty N, et al. Assessing the safety of human pluripotent stem cells and their derivatives for clinical applications. *Stem cell reports.* 2017;9(1):1-4.
47. Tate CC, Shear DA, Tate MC, Archer DR, Stein DG, LaPlaca MC. Laminin and fibronectin scaffolds enhance neural stem cell transplantation into the injured brain. *J Tissue Eng Regen Med.* 2009;3(3):208-217.
48. Li J, Lepski G. Cell transplantation for spinal cord injury: a systematic review. *BioMed research international.* 2013;2013.
49. Cheng LS, Schwartz DM, Hotta R, Graham HK, Goldstein AM. Bowel dysfunction following pullthrough surgery is associated with an overabundance of nitrergic neurons in Hirschsprung disease. *Journal of Pediatric Surgery.* 2016;51(11):1834-1838.
50. Burns AJ, Goldstein AM, Newgreen DF, et al. White paper on guidelines concerning enteric nervous system stem cell therapy for enteric neuropathies. *Developmental biology.* 2016.
51. Fattorini E, Brusa T, Gingert C, et al. Artificial Muscle Devices: Innovations and Prospects for Fecal Incontinence Treatment. *Ann Biomed Eng.* 2016;44(5):1355-1369.

52. Fritsch H, Brenner E, Lienemann A, Ludwikowski B. Anal Sphincter Complex. *Dis Colon Rectum.* 2002;45(2):188-194.
53. Snooks S, Swash M. The innervation of the muscles of continence. *Ann R Coll Surg Engl.* 1986;68(1):45.
54. Rao SS. Pathophysiology of adult fecal incontinence. *Gastroenterology.* 2004;126:S14-S22.


7


A

Appendix

Summary

Hirschsprung disease (HSCR) is a disease of the intestines. It is characterised by an absence of the enteric nervous system (ENS) in a distal portion of the colon causing a contraction of the muscles and a functional obstruction. This results from defects in the differentiation, proliferation, migration and/or survival of ENS progenitors during development. HSCR can be an isolated trait or be part of a multi-feature syndrome. There are multiple developmental pathways which contribute to these defects and the genetic background of HSCR is complex, ~70% of cases cannot be explained by known genes.

The possibilities of somatic variations contributing to the development of, and accounting for the missing heritability in, HSCR are discussed in chapters 2.1 and 2.2. In **chapter 2.1** we outline the need for appropriate distinction between inherited parental mosaicism and true somatic mosaicism, proposing an experimental design to differentiate between the two in HSCR patient tissue. In **chapter 2.2** this experimental design is utilised to look for ENCC specific variation in patient tissue. Although somatic variants were present in all included patients, somatic variants in HSCR related genes were not. Due to the nature of ENS development it is likely that somatic variants could not be identified in the distal colon. If damaging somatic variants were to occur in ENCCs, these cells would likely be out-competed to ENS niches before reaching these distal regions.

In syndromic cases, HSCR can be a variable feature, the presence of common *RET* variants as well as the type of causative variant may influence HSCR development. The presence of HSCR in cases of Goldberg-Shprintzen syndrome (GOSHS) is discussed in **chapter 3**.

Pathogenic variations in KIF1 binding protein (*KIF1BP*) lead to GOSHS. A number of patients have been reported in the literature and truncating variations in *KIF1BP* have been found to be causative in all sequenced cases. We report nine new patients with *KIF1BP* variations, and functionally investigate three new missense variants which were suspected to be pathogenic in two patients with differing phenotypes. The three missense variants were found to result in a decrease in KIF1BP expression. The variant resulting in the lowest expression was present in the patient with the classical GOSHS and HSCR. Common predisposing HSCR SNPs were not found to have correlation with the presence or absence of HSCR in GOSHS patients.

Another possible genetic factor for HSCR in syndromic cases is the presence of copy number variations (CNVs) that affect dosage sensitive HSCR loci. In **chapter 4** we compare the size and number of CNVs between syndromic and non-syndromic cases of HSCR to find new candidate genes/loci. Syndromic HSCR patients with an unknown genetic aetiology have more and larger CNVs than isolated HSCR cases with a known pathogenic variant. These large CNVs overlap with dose-sensitive genes which may help to identify candidate genes for HSCR.

To further look into the underlying pathogenesis of HSCR development we created iPSC lines from four patients with different pathogenic variations. These cells offer possibilities for disease modelling, functional investigation of variants as well opening the door to future iPSC-enteric neuron transplantation options. In **chapter 5** we present the characterisation of these patient-derived iPSCs and explore variations in their function compared to iPSCs generated from healthy controls. Three out of four of the patient-derived lines differentiated

effectively towards a vagal neural crest lineage and expressed appropriate markers. The less successful line also showed a markedly higher proliferation rate suggesting immaturity in differentiation. High seeding density was an important factor in successful differentiation, implicating cell-cell contact as vital for neural crest formation. Together our data suggest that for many HSCR patients, gene correction may not be necessary before transplantation trials.

In order to bring an expedient option for the treatment of HSCR patients who continue to suffer from gastrointestinal problems following the current surgical standard of care, we may have to think beyond biological interventions. The technological world has arguably been able to advance at a faster rate than the development of purely biological treatment options. With the advances in microelectronics and prosthetic technologies incorporating sensory input, the opportunities for developing transplantable devices may be provide a more elegant solution than the more primitive prostheses currently available. These ideas are discussed in **chapter 6,** together with a patent proposal for an artificial prosthetic sphincter with an anatomically relevant mechanism and design which could provide therapeutic options for HSCR patients and others suffering from faecal incontinence or loss of anal sphincter control.

In conclusion, the work of this thesis investigates the development of, and treatment options for, HSCR. We explore the missing heritability that is seen in HSCR, modes of development and differentiation of the progenitors of the ENS, suitability of various cell sources for transplantation therapy, and explore other possible treatment avenues for current and future HSCR patients.

Samenvatting

De ziekte van Hirschsprung (HSCR) is een aangeboren neurologische aandoening van het maagdarmkanaal. De ziekte wordt gekenmerkt door de afwezigheid van ganglia in het uiteinde van de darm. Ganglia zijn groepjes zenuwcellen en gliacellen die behoren tot het enterische zenuwstelsel (engels: ENS). Daar waar de ganglia ontbreken trekken de spieren samen en veroorzaken zo een verstopping. Het ontbreken van ganglia kan het gevolg zijn van fouten in de differentiatie, proliferatie, migratie en/of overleving van ENS voorlopercellen tijdens hun ontwikkeling. De aanleg / ontwikkeling van het enterische zenuwstelsel is complex en dat kan meestal ook gezegd worden van het ontstaan van HSCR. De belangrijkste oorzaak voor het ontstaan van HSCR zijn fouten (mutaties) in het erfelijk materiaal. Er is één gen dat het meest gemuteerd voorkomt en dat is het *RET*-gen. We vinden mutaties van het *RET*-gen die genoeg veranderingen kunnen veroorzaken om de ziekte doen ontstaan, maar, veel vaker vinden we ook variaties die de kans op de ziekte verhogen, maar niet veroorzaken. Echter, in ~70% van de gevallen kan de volledige oorzaak nog niet worden verklaard door mutaties.

HSCR kan als een op zichzelf staande ziekte voorkomen, maar kan ook in combinatie met andere afwijkingen worden gediagnosticeerd; dan noemen we het een syndroom.

Zoals gezegd is de oorzaak van de ziekte veelal erfelijk, en mutaties in een groot aantal genen zijn al gevonden. Deze mutaties erft de patiënt vaak over van één van de ouders. Soms ontstaan de mutaties in de geslachtscellen (dit noemen we ook wel kiembaan mozaïcisme) en heeft alleen het kind de genetische afwijking. Maar zelfs als we alle genen

screenen vinden we niet altijd een duidelijke verklaring. Een hypothese die we hebben onderzocht was of er mutaties voorkomen alleen in het ENS. Die mutaties moeten dan tijdens de ontwikkeling van het kind ontstaan. We noemen dergelijke mutaties somatische mutaties of variaties. Deze hypothese, dat somatische varianten bijdragen aan de ontwikkeling van HSCR, wordt besproken in de hoofdstukken 2.1 en 2.2. In hoofdstuk 2.1 schetsen we de noodzaak voor een duidelijk onderscheid tussen kiembaan mozaïcisme en werkelijk somatisch mozaïcisme, en stellen we een toepasbaar experimenteel plan voor om in weefsel van HSCR patiënten dit onderscheid ook daadwerkelijk te kunnen maken. In hoofdstuk 2.2 wordt van dit experimentele plan gebruik gemaakt om te onderzoeken of er in het ENS van patiënten inderdaad sprake is van specifieke somatische variaties in de zenuwcellen. Alhoewel somatische varianten aanwezig waren in alle onderzochte patiënten, vonden we er geen in de genen waarvan bekend was dat ze HSCR kunnen veroorzaken. Ook beschrijven we dat, gezien de manier waarop het ENS wordt aangelegd, de kans op het vinden van somatische varianten niet heel waarschijnlijk is; we denken namelijk dat als somatische varianten zouden voorkomen in het ENS, deze cellen vermoedelijk weg zouden worden geconcurreerd door gezonde ENS specifieke cellen voordat ze deze plek bereiken.

Zoals gezegd, kan HSCR onderdeel zijn van een syndroom. Eén van de bekende syndromen is Goldberg-Shprintzen syndroom (GOSHS). We bespreken dit syndroom in hoofdstuk 3. Pathogene varianten in het KIF1 bindingseiwit (KIF1BP) veroorzaken GOSHS, met HSCR als een variabel kenmerk. Alhoewel HSCR geen criterium is voor de diagnose GOSHS, zijn er mogelijk wel factoren die er voor zorgen dat deze patiënten HSCR

ontwikkelen. Een aantal patiënten zijn beschreven in de literatuur en in al die patiënten waarbij DNA geanalyseerd werd, werden varianten gevonden die een verkort KIF1BP eiwit tot gevolg hebben. We beschrijven negen nieuwe patiënten met varianten in KIF1BP. Ook hebben we functioneel onderzoek gedaan naar drie nieuwe missense varianten (missense: vervanging van een aminozuur door een ander aminozuur), t.w. varianten waarvan we denken dat ze ziektes kunnen veroorzaken. We hebben deze varianten gevonden in twee patiënten met verschillende fenotypes. De drie missense varianten bleken een verlaagde KIF1BP expressie tot gevolg te hebben. De variant met de laagste expressie werd gevonden in de patiënt met klassiek GOSHS en HSCR. We hebben geen veelvoorkomende veranderingen gevonden die gecorreleerd zijn aan de aan- of afwezigheid van HSCR in GOSHS patiënten.

Een andere mogelijk genetische factor voor HSCR in syndromale gevallen, is de aanwezigheid van grote variaties in het DNA, d.w.z. grote stukken van een chromosoom die extra of juist minder aanwezig zijn. In hoofdstuk 4 vergelijken we de grootte en het aantal chromosoomafwijkingen (CNVs) tussen syndromale en niet-syndromale HSCR patiënten om nieuwe kandidaat genen/loci te vinden. Syndromale HSCR patiënten met een onbekende genetische oorzaak hebben meer en grotere CNVs dan patiënten met alleen HSCR en een bekende ziekteverwekkende mutatie. Deze grote CNVs overlappen met dosis-gevoelige genen. Deze bevindingen kunnen helpen bij het identificeren van nieuwe kandidaat genen voor HSCR.

Om de onderliggende pathogenese van HSCR verder te onderzoeken, hebben we cellijnen gemaakt van geïnduceerde pluripotente stamcellen

(engels: iPSCs) van vier patiënten met mutaties in verschillende genen. Deze cellen bieden de mogelijkheid om het ziektebeeld te modelleren en de varianten functioneel te onderzoeken. Dit type onderzoek opent de deur naar toekomstige therapie. We denken dat daar waar de neuronen ontbreken deze cellen mogelijk gebruikt kunnen worden voor transplantatie in de darm. In hoofdstuk 5 laten we de karakterisering, van deze van de patiënt afgeleide cellijnen zien. Ook gaan we na of de variaties die we hebben gevonden effect hebben op de cellijnen (we vergelijken de patiënten cellijnen met controle cellijnen). Drie van de vier van de patiënt afgeleide lijnen lieten geen echte verschillen zien. Eén cellijn was anders. De cellijn vermeerderde zich aanmerkelijk sneller dan de rest, hetgeen duidt op een onrijpe differentiatie. Een belangrijke factor voor het goed groeien is een hoge celdichtheid bij het opgroeien, wat er op duidt dat cel-cel contact van vitaal belang is bij de vorming van deze cellen. Omdat de meeste cellen geen groot verschil lieten zien lijkt het erop dat deze cellijnen geschikt zouden moeten zijn voor therapie.

Om met een doeltreffend alternatief te komen voor de behandeling van HSCR patiënten die last blijven houden van gastro-intestinale problemen na de huidige standaard operatieve behandeling, moeten we verder denken dan de biologische interventies. In de technische wereld is de vooruitgang sneller gegaan dan de ontwikkeling op het gebied van puur biologische behandelmogelijkheden. Met name de vooruitgang in de micro-elektronica en de prothetische technologieën die gevoelssensoren weten in te bouwen, bieden mogelijkheden om transplanteerbare hulpmiddelen te ontwikkelen. Deze ideeën worden besproken in hoofdstuk 6, samen met een patent protocol voor een kunstmatige sluitspierprothese met een anatomisch toepasbaar

mechanisme en ontwerp. Deze zou kunnen voorzien in therapeutische behandelmogelijkheden voor HSCR en andere patiënten die leiden onder fecale incontinentie of het gebrek aan controle over de sluitspier.

Samengevat, het werk gepresenteerd in dit proefschrift, onderzoekt de ontwikkeling van, en de behandelmogelijkheden voor, HSCR. We hebben gezocht naar erfelijke factoren in HSCR, de wijze van ontwikkeling en differentiatie van de ENS voorlopercellen, de geschiktheid van cellen met verschillende origine voor transplantatie therapieën én andere behandelmogelijkheden voor huidige en toekomstige HSCR patiënten.

Curriculum vitae

Personal Information

Name:	Katherine Christa MacKenzie
Date of Birth:	31-05-1990
Place of Birth:	Alexandria, Scotland

Professional Experience

2014-2019	PhD student with Prof. Dr. Robert Hofstra, in the Department of Clinical Genetics, Erasmus MC, Rotterdam

Education

2013-2014	Master of Science *Stem Cell Technology* University of Nottingham, England
2008-2013	Bachelor of Science with Honours *Anatomical Sciences* University of Dundee, Scotland

Internships

2013	Nurture Fertility Clinic, Nottingham, England
2011-2012	Defence Science and Technology Laboratory, Kent, England
2011	Scottish Centre for Regenerative Medicine, Edinburgh, Scotland

PhD Portfolio

Courses		
Genetics course	2014	**3**
Laboratory Animal Science (Artikel 9)	2014	**3**
Research Integrity	2015	**0.3**
OIC Functional Imaging	2015	**1.5**
Safely Working in the Laboratory	2016	**0.3**
Special Topics course: Epigenetics	2016	**3**
Special Topics course: CRISPR	2017	**3**
Biomedical English Writing	2017	**2**

Seminars & Workshops		
Sophia Research Days	2014-18	**1**
Clinical Genetics Meetings	2014-18	**2**
Clinical Genetics Seminars	2014-18	**1**
Mouse Models course	2015	**2**
MGC workshop: Maastricht	2015	**1**
MGC workshop: Dortmund	2016	**1**
Journal Club	2016-18	**1**
MGC workshop: Texel	2018	**2**

Conferences & Symposia		
MGC Symposia	2014-18	**2**
International Symposium on Development of the ENS: Rotterdam	2015	**1**
Mouse Models Symposium: Leiden	2015	**0.5**
Stem Cell Symposium: Utrecht	2016	**0.5**
European Society of Human Genetics: Copenhagen	2017	**1**
International Symposium on Development of the ENS: Boston	2018	**2**

Teaching		
Supervisor Bachelor student	2016	**1**
Teaching Assistant Journal Club	2016-17	**2**
Teaching Assistant Philosophy & ethics	2016-18	**3**
Total		**40**

Acknowledgements

Oh dear. It's finally done. Phew! There're a lot of people that I need to thank for helping me get to the stage of actually having a thesis, this might end up as the longest chapter!

Firstly, it's meant so much to this work, and many other projects within our department, that patients and their families have given their permission to use their material and data to better understand genetic diseases. For all the patients with Hirschsprung disease and their families, I hope that this work has pushed the boundaries of our knowledge a little wider and taken us a step closer to treatment.

To the members of my thesis committee, thank you for reading and accepting this book and being on the panel. I hope you found it interesting.

Dear Robert, when we first met in London, I had no idea that it would lead me to Rotterdam. I'm so very grateful to you for bringing me here and for supporting me through the last 4.5 years. It's definitely been a challenge for me to make it to here, but it's all come together, just like you always said. Thank you for everything.

Dear Alan, I really enjoyed working with you, both in Rotterdam and my visit to London. Thank you for introducing me to Robert and for supporting me in the various cell problems over the years. Best of luck with everything in the US and your role at Takeda.

Dear René, thank you for the collaboration we had with you and the rest of Paediatric Surgery. The support of you, Pim, Conny, and Ivo's group in Nijmegen has been so important to the work our group has done for the last years and has been vital to this thesis. I hope that the collaboration continues and that this work has laid a foundation more!

Maria, I'm so happy to have worked with you, I couldn't have finished this thesis without your help and encouragement. Having you as a friend as well as a supervisor means a lot to me and I know we will still see a lot of each other, I need to play with Arya more before she starts school! All the biggest hugs to you, Kaushal and Arya.

Erwin, thank you so much for not strangling me over the CNV chapter. Your knowledge of genetics and genetic analysis still baffles me. You've been such

a great support and it's been really nice working with you. I hope that Suze inherits your great sense of humour and that you can team up to out-vote Neeltje on the whole mini-pig situation!

Darling **Bianca**, you are a star! I owe you so much for all your help and support and couldn't be happier to have had you as a colleague and a friend these last few years. Our trips to Seville and Copenhagen were so much fun, and the concerts and dinners we shared. Your kindness, honesty and humour has made my time here feel more like home and I'm so grateful to you for that. Thank you for being my paranymph, it means a lot to me that you are going to be by my side!

Jonathan, didn't I tell you that it'd be hard! I'll be waiting to hear all your whining and complaining about cells and FACS and sequences not working! JK, I can't think of a better person to take over the patient projects. It's been really nice working with you and laughing with you this last year and you're going to be an amazing surgeon in the near future.

Almira, it's been a pleasure teaching you the tricks (that I'm aware of) for the iPS cells and helping you get to grips with the task ahead of you. It'll be overwhelming and challenging and I hope what I did these last few years will give you a head-start with your projects. Thanks for continuing this work and good luck with your thesis, you're going to do great!

Of course, the other members of the GI Genetics group, thanks for our discussions and all the input and help, especially towards the end. **Rhiana**, I've missed you a lot since you left. I'm really grateful to you for so much. I wish all the best for you and your beautiful children! **William**, it was so sad to see you go. The talks we had about science and academia in general and your honesty, advice and guidance always helped a lot. **Alice**, thanks for your support in the KBP project, it's been lovely working with you. **Veerle**, your visits from Maastricht were always a nice highlight to the month. I enjoyed our talks about yoga and food a little more than those about my projects! **Jeannette**, thanks so much for all that you have done for me and the whole group. You're the organiser of the chaos and have been so helpful the whole way through my time here. **Rajendra**, it was fun to work with you and I am so happy that you've found a job you enjoy and you **Chaitali** & Rudransh are happy in Rotterdam. **Maria E**, it feels like no time at all has passed since you were here in Rotterdam, but so much has happened! I enjoyed working with you so much and wish it could have lasted longer. I hope you and Alan are super happy in Seville and I'll see

you there very soon! Nathalie, my ray of lab-sunshine. Every time you visited here you were so enthusiastic and happy about your research that it almost made me happy about mine! Thanks for all the fun! Danny, thank you for all our talks during our CAM work, you helped a lot in the start of my PhD. Best of luck to you and your family in Indonesia. Yuying, it's been so lovely getting to know you. I hope that when you're back in China the work you've done will help you in your surgical training, good luck for everything and keep in touch! Chantal, you're so organised and driven that I can't see you having any problems that you won't be able to handle in finishing your thesis! It's been lovely having you as a part-time office buddy and good luck with everything to come. And the students that we've had over these last few years, Zakia, Tim, Jeffrey, Jagoda and Sander – you guys worked really hard and I hope you enjoy the rest of your scientific journey!

I want to thank my colleagues for the support in the lab, and also for all the social events we enjoyed together. Nynke, I always admired you as a scientist and I'm so glad that you got the position you wanted in Groningen. I have so many fun memories with you from all our activities. Thanks for everything! Laura K, it was really lovely sharing an office with you, I'm still amazed that you could concentrate through my chat! I'm glad to have spent these years working with you. Good luck with your future career! Atze, thanks so much for being a great friend and being so patient waiting for your scarf! You've been one of the people that's made the Netherlands feel like home for me and I'm really grateful. I can't wait to see more of you, Maaike and your beautiful family! Mike, you're still a ****, but you're a good one. Thanks for helping me whenever I needed it and for being honest when I was being an idiot. I had a great time arguing and conspiring with you. Good luck with finishing this shit! Erik, I did it, just like you said I would! Thanks for your encouragement and for your help with the iPS and all the chats we had to try and distract Nynke and Laura. Qiushi, I miss you and your hugs. It was lovely sharing an office with you and I hope I'll be able to see you sometime in the future in China or somewhere in the world! Elena, Eva, Soheil, and Wim Q, for the short time that we shared an office it was really nice! I hope all of your PhDs go amazingly and you don't have too many huge problems, good luck!

Fenne, thanks for all the support you gave when things were going crappy and for all the fun we had outside the lab! Saif, thank you for always saying the right things to cheer me up, I always remember the note you left

with the croissant when I had that shitty day. Don't let set-backs get to you, you're going to do amazingly. **Michael**, thanks for getting sweets/beer with me when our experiments were going crappy, you knew the way to my heart. In the end I'll actually defend before you! HAH! Good luck, buddy! **Tom**, thanks for the many, many hours of sorting and complaining together. You and the kabouters made the job much more fun. **Stijn**, even though it took me a while to break your social-shell with over-enthusiastic smiling, I got there in the end! Thanks for being a great friend. **Tsung Wai**, thank you so much for all the help with our horrible cells! Your willingness to stay a bit late to help us out with the patient samples was really appreciated! **Mehmaz**, thank you so much for your work on the iPSCs and for all your support when I had questions. **Tjakko**, **Pim**, **Vincenzo**, **Rob**, **Renate**, **Annelies**, **Stefan**, **Joost**, **Grazia** & **Adriana**, thank you all for your input at our research meetings and for always being happy to answer questions or give help whenever I've asked.

My fellow PhDs and the members of the MGC, thanks for having so many fun (and alcohol-fuelled) times with me! **Pablo**, **Rodrigo**, **Alessandro**, **Fabio**, **Kyra**, **Natasha**, **Wojtek**, **Douglas**, **Monica**, **Judith**, **Ruben**, **Irene**, **Daphne**, good luck with finishing and thanks for being great! And other colleagues and students that have been in and out of the department, **Bep**, **Rachel N**, **Isa Q**, **Wim M**, **Gerben**, **Rob V**, **Liesanne**, **Guido**, **Helen**, **Rachel**, **Michelle**, **Herma**, **Joon**, **Esmay**, **Esmee**, **Roy**, **Ronald**, **Shimriet**, **Simone**, **Marianne**, **Max**, **Mariana**, **Anita**, **Darija**, **Matthijs**, **Lucas**, **Kirke**, **Robin**, **Woutje**, **Leontine**, **Martina**, **Isa Gordaliza**, **Andrea Arrizibita**, it was really nice working with you all! To the other staff of Clinical Genetics on the 20th and 24th floors, thanks for the fun lab days out and for the support and help whenever asked. Also within the Erasmus MC, thanks to **Claire Wyman** and **Martine Jaegle** for letting me help you with the bachelor modules, I really enjoyed it!

To our collaborators in the UK, **Nikhil**, **Dipa**, **Connor**, **Ben**, **Julie**, **Tom**, **Jim** and **Peter Andrews**, thanks for all your help when I visited and over email, it was great to work with you. To **Nayia**, thanks for your help and collaboration with the *KIF1BP* project.

I think after my second year I would have been driven over the edge if it weren't for my yoga practice! I switched from doing it at home to practicing at Balanzs studio, and it had such a positive impact on my life. So, I'd like to thank all my teachers and fellow yogis at **Balanzs**, especially **Erlijn**, Fi, **Cecilio**, **Bente**, **Benjamin**, **Tim G**, **Charlotte** and **Luca**, you are inspirations to me and I want to thank you for everything you've given me.

My fellow losers, Hidde, Daan, Tedo, Chenise & Glenn, Dungeons & Dragons night was another tool for science-escapeism. Thanks for being great adventurers and even better friends. You all get 3 inspiration points to never use! To my lab-sisters ♥, Demy, Laura V, Ana and Shami, you guys made me laugh, saw me cry and helped me deal with the shitty things. You're all so wonderful, and you know how much I appreciate you, thank you! To my amazing friends, Rosie, Debbie, Lotte, Claire, Jennifer, Sheona, Natalia, Emma, Hoksha, Lucy, Calum, Euan, Cat, Niall, Mini, Reuben, Scott, and my lovely boyfriend André, thanks for listening when I spewed stress at you and for just generally being wonderful humans. I love you all!

And now I'm going to get all soppy, because that's what the last paragraphs are for ☺. Erudini & Martyna, I couldn't ask for two more incredible friends. You guys have seen all my sides and you've stuck with me for my most stressful and low times. You, Figaro & Meeko are my Rotterdam family and I love you like sisters, I couldn't have gotten through these last few years without you and can't thank you enough for everything that you are!

And most importantly, the MacKenzies! My darling Mammy, you're always there for me to cry to and laugh with, you're hilarious and loving and just the best mum anyone could ask for. Pops, you're the reason I got into science and there's no way I could have motivated myself to finish this without your support through every stage of my studies. Lizzie, Donald and Rachel, you guys are my rocks, you are my favourite people, I'm so lucky that my brother and sisters are my best friends. I can't wait to celebrate finishing this with you guys!!

WOO HOO!

Katherine x

Hrschsprung Disease Development & Treatment Avenues

Katherine C. MacKenzie