The Genotype-Phenotype Correlation in Pompe Disease Studied from an Enzymatic and Molecular Perspective

Monica Y. Niño Martinez

The Genotype-Phenotype Correlation in Pompe Disease Studied from an Enzymatic and Molecular Perspective

THESIS

Monica Yasmin Niño Martínez

The research presented in this thesis was financially supported by grants from Colciencias, Colombia (Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias), Colombia (Administrative Department of Science, Technology and Innovation, Colombia. Thesis printing cost were sponsored by Erasmus University and the department of Clinical Genetics, Erasmus Medical Center, Rotterdam.

The studies described in this thesis were performed in the Center for Lysosomal and Metabolic Diseases, Department of Pediatrics and Department of Clinical Genetics of the Erasmus University Medical Center, Rotterdam, The Netherlands.

ISBN:	978-94-6402-290-2
Author:	Monica Yasmin Niño Martinez
Cover Design:	Monica Yasmin Niño Martinez
Layout:	Ilse Modder | www.ilsemodder.nl
Printed by:	GildePrint Eschede | www.gildeprint.nl

The Genotype-Phenotype Correlation in Pompe Disease Studied from an Enzymatic and Molecular Perspective

Is de klinische diversiteit van de ziekte van Pompe verklaarbaar op basis van enzymatische en moleculaire karakeristieken?

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 defense will be held on
7th of July 2020 at 15:30 hrs.

by
Monica Yasmin Niño Martínez
born in Tunja, Colombia

DOCTORAL COMMITTEE

Promotors:	Prof. dr. A.T. van der Ploeg
	Prof. dr. R.M.W. Hofstra
Other members:	Prof. dr. E.H.H.M. Rings
	Prof. dr. N.M. Verhoeven
	Prof. dr. P.A. van Doorn
Copromotor:	Dr. W.W.M.P. Pijnappel

TABLE OF CONTENTS

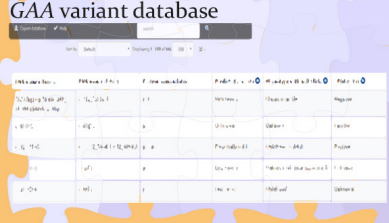
LIST OF ABBREVIATIONS

4MUG = 4-methylumbelliferyl-α-D-glucopyranoside
6MWT = 6-minute walk test
ACB = Acarbose
ACE = angiotensin converting enzyme
ACTN3 = alpha-actin-3
ALT = alanine aminotransferase
AMD = acid maltase deficiency
AON = antisense oligonucleotide
AST = aspartate aminotransferase
BBB = blood brain barrier
BGAl = β-galactosidase
CHO = Chinese hamster ovary
CK = creatine kinase
CNS = central nervous system
CRIM = cross reactive immunologic material
CT = computed tomography
DBS = dried blood spot
dbSNP = database for single nucleotide polymorphisms
DGEs = differentially expressed genes
DMD = Duchenne muscular dystrophy
ECG = electrocardiogram
EM = electron microscopy
ER = endoplasmic reticulum
ERAD = endoplasmic reticulum-associated protein degradation
ERT = enzyme replacement therapy
ESP = exome sequencing project
ExAC = exome aggregation consortium
FDA = Food and Drug Administration
FEV1 = forced expiratory volume
FSHD = facioscapulohumeral muscular dystrophy
GAA = acid α-glucosidase gene
GAA = acid α-glucosidase
GAA-S = (7-Benzoylamino-heptyl)-{2-[4-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro pyran-2-yloxy)-phenylcarbamoyl]-ethyl}-carbamic acid tert-butyl ester
Glc4 = tetrasaccharide 6-α-D-glucopyranosyl-maltotriose
GN = glycogen

GO = gene ontology
GoNL = genome of the Netherlands
GSDII = glycogen storage disease type II
HGVS = human genome variation society
HPLC = high performance liquid chromatography
IGF1 = insulin-like growth factor 1
IPA = ingenuity pathway analysis
iPSCs = induced pluripotent stem cells
IVS1 = c.-32-13T>G
LSD = lysosomal storage disorder
M6P = mannose 6-phosphate receptor
MAF = minor allele frequency
MDS = multidimensional scaling
MLPA = multiplex ligation-dependent probe amplification
NGS = next generation sequencing
NMD = neuromuscular disease
OOA = out of Africa
PANTHER = protein analysis through evolutionary relationships
PAS = periodic acid Schiff reagent
PD = Pompe disease
PMNs = polymorphonuclear granulocytes
PNS = peripheral nervous system
pre-Mrna = pre-messenger RNA
QMFT = quick motor function test
RER = rough endoplasmic reticulum
rhGAA = recombinant human GAA
RNA-seq = RNA sequencing
RT-qPCR = quantitative reverse transcription PCR
SDS-PAGE = sodium dodecyl sulfate polyacrylamide gel electrophoresis
SIFT = sorting intolerant from tolerant server
VC = vital capacity
WS = Walton and Gardner-Medwin score

Pompe disease

GAA variant database



CHAPTER I

SCOPE AND GENERAL INTRODUCTION

1 Scope and General introduction

SCOPE

This thesis is about Pompe disease. Pompe disease is characterized by skeletal muscle weakness leading to serious mobility and respiratory problems over time. The first symptoms can manifest at any age: from birth to late adulthood. In most severe form, babies suffer from generalized muscle weakness (so called 'floppy infants'), they have a grossly enlarged heart, and their respiratory function is compromised to such extend that they usually succumb in the first year of life. Affected adults often become wheelchair bound and may need respiratory support. Based on symptomatology, Pompe disease is listed among the neuromuscular diseases (NMD), but since the organ dysfunctions are caused by storage of glycogen inside lysosomes (one of several intra-cellular compartments) it is also listed as a glycogen storage disorder (GSD) and as a lysosomal storage disorder (LSD). Pompe disease is inherited in an autosomal recessive way and arises by disease causing variations in the genetic code of an enzyme called acid α-glucosidase (abbreviated 'GAA').

The research described in the various chapters of this thesis is focused on i) how to measure the GAA enzyme activity for diagnostic purposes, ii) investigations on genotype-phenotype relationships, iii) the role of the *ACE* polymorphism in disease severity and response to ERT among patients with the c.-32-13T>G / 'null' genotypes, and iv) gene expression profiling of patients with early as oposed to late presentation of symptoms. A General Introduction precedes the research chapters.

GENERAL INTRODUCTION

The name Pompe disease: historic overview

For a long time, 'Pompe disease' (OMIM # 232300) was not the commonly used name for this inherited disorder [1, 2]. The name refers to the Dutch pathologist Johannes Cassianus Pompe who published in 1932 a case report about 'idiopathic hypertrophy of the heart' as a prominent finding from an autopsy of a 7 month old girl. Dr. Pompe observed a defect in glycogen metabolism and described it as 'vacuolar deposit of glycogen' in the heart and many other tissues. He called this condition 'cardiomegalia

glycogenica diffusa', and made the analogy to the cases of kidney and liver enlargement described by von Gierke in 1929 [3, 4]. Two other articles were published in the same year by Drs Putschar and Bischoff who described very similar cases of children who had died in infancy with 'cardiomegaly' and severe muscle glycogen accumulation [5, 6]. This disorder became known as Pompe disease. Classification schemes in the 1950s delineated Pompe disease as type II glycogenosis, and the molecular defect (acid α-glucosidase or acid maltase deficiency) was deciphered in the 1960s as described in depth later [4, 6]. In 1950, Di Sant'Agnese and co-authors published a critical review of all similar cases reported up till then. The disease attracted the name 'Glycogen Storage Disease type II' (GSDII), as given to it by Dr GT Cori who had made an inventory of the several different glycogen storage diseases that had come to light by the year 1954 [7]. Nine years later, in 1963, Dr. Hers discovered that Pompe disease (GSDII) was caused by acid α-glucosidase (EC 3.2.1.20) deficiency, a lysosomal glycogen-degrading enzyme that splits also maltose into glucose. The name 'acid maltase deficiency' (AMD) settled when Dr. Engel described affected children and adults in 1973 [9]. Acid α-glucosidase deficiency became a specific diagnostic test for GSDII / Pompe disease / AMD and provided also new insights in disease pathogenesis. The publication by Hers was particularly interesting as it established the role of lysosomes, described in 1955 by DeDuve, as an organelle capable of degrading a great variety of macromolecular compounds at acidic pH (3.5-5.0) [8]. Soon after that ground-breaking discovery, publications appeared about patients with acid α-glucosidase/acid maltase deficiency and GSDII without cardiomyopathy, but with skeletal muscle weakness as main symptom [9-11]. When the gene encoding acid α-glucosidase / acid maltase was mapped to chromosome 17, the gene locus was abbreviated as *GAA*, for Glucosidase Acid Alpha [12-14], and the enzyme 'GAA', in connection with its deficiency, became yet another alternative to address the condition GSDII / Pompe disease / AMD. Mainly through the publication policy of researchers from Erasmus MC and the start of enzyme replacement therapy the name Pompe disease has gradually rooted. In the 2001 edition of the Metabolic and Molecular Bases of Inherited Disease the chapter on 'GSD II' was incorporated for the first time in the chapter series on Lysosomal Storage Disorders, while formerly it had been part of the chapters on Glycogen Storage Disorders of non-lysosomal origin [15]. In the latest online version of the Metabolic and Molecular Bases of Inherited Disease (OMMBID) the name Pompe Disease is called first in the Title and is currently broadly accepted [2]. The name Pompe disease is now generally accepted by the scientific community as the result of numerous publications using this name in the literature. However, especially in the older literature the alternative names AMD and GSDII have been exclusively used. This should be taken into account when searching the literature.

The disease is rare, but has gained world-wide interest since 2006 mostly due to the introduction of enzyme replacement therapy (ERT), which has substantially contributed to adoption of the name Pompe disease for all presentations across the clinical spectrum of acid α-glucosidase / acid maltase / GAA deficiency. The fact that glycogen stores in lysosomes makes Pompe disease both a glycogen storage disease (GSDII), of which there are more than 10 different types, as well as a lysosomal storage disorder (LSD), of which there are over 70 [16, 17].

Incidence and geographic distribution of Pompe disease

Pompe disease is caused by a set of two disease-associated variants in the *GAA* gene; one in each allele, and is thereby an autosomal recessive disorder. Some of the disease-associated *GAA* variants clearly demonstrate a founder effect, which combined with population migration largely determines the incidence of the disease in specific populations and geographic areas [2, 18-23]. Based on pathogenic sequence detection in the Dutch population, the overall incidence of Pompe disease in the Netherlands was estimated to be around 1 in 40.000 newborns. Further, by the type of sequence variant it was estimated that classic infantile Pompe disease has an incidence of 1 in 138.000, whereas childhood and adult Pompe disease has an incidence of 1 in 57.000 [19, 24]. These figures are in line with estimates based on the number of diagnosed cases over a long period [24, 25]. A similar approach was taken in Australia and Portugal where the number of diagnosed cases was evaluated over a period of 16 years (1980-1996) and 19 years (1982-2001), respectively, and a much lower prevalence was registered (1:146.000 and 1:600.000), respectively [22, 23]. Martiniuk et al., based their estimate of the incidence of Pompe disease in the United States on the carrier frequency of 7, then known, rather common *GAA* sequence variants among inhabitants of New York and derived from their study an incidence 1:40.000 [20]. However, after the most common disease-associated *GAA* variant among African-Americans was identified the data of Martiniuk was recalculated and it was argued that the prevalence of Pompe disease among African-Americans could be as high as 1:14.000 [2, 26]. A similarly high prevalence may also count for Taiwan where the carrier frequency was initially estimated to be 1:100 to 1:200, but later adapted based on new born screening (NBS). The overall incidence was 1 in 17.000, and the incidence for classic infantile was 1 in 52.000 and for other types was 1 in 25.000 [21, 27, 28]. A recent report of NBS in Japan reported that the incidence in that country is 1 in 34.000. In this study 103.204 newborns were screened; three newborns were diagnosed with putative late-onset Pompe disease and no classic infantile Pompe disease was detected [18]. A pilot newborn screening study conducted in Austria using bloodspots identified 4 samples with low GAA enzyme activity and two disease-associated *GAA* variants in each case. If all donors of

these samples would, indeed, develop Pompe disease the prevalence in Austria would be 1:8.700 [29]. Pompe disease seems to be very rare in Finland. The extrapolation from the assumed incidences, estimates that the prevalence worldwide could be 5.000-10.000 people of both genders, different ages and ethnic groups [30]. The incidence of Pompe disease in its different clinical presentations and in different populations is shown in Figure 1 and summarized in Table 1.

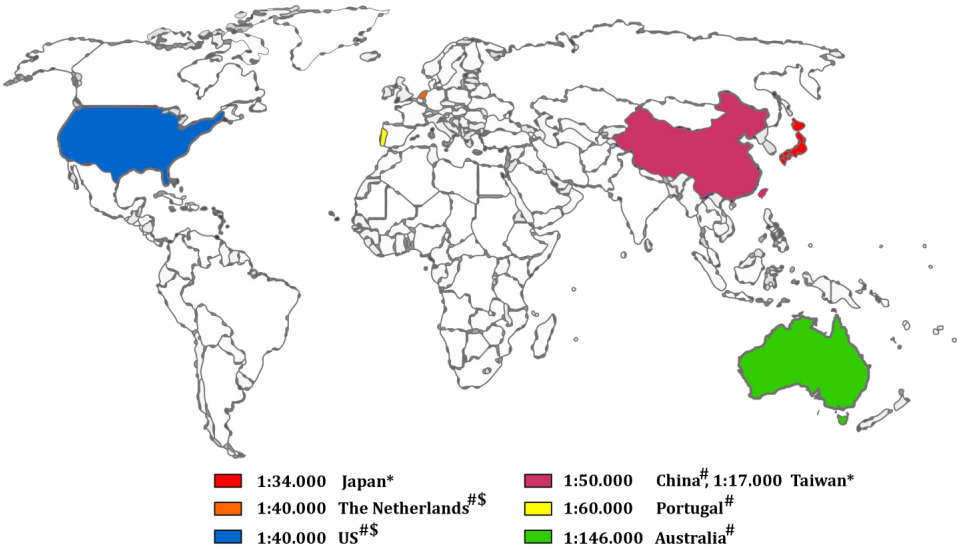


Figure 1. Reported incidence of Pompe disease. The incidence of Pompe disease is not exactly known, but seems to differ by geographic region and population. The estimates are based on the number of diagnostic cases over a given period of time#, the observed carrier frequency based on *GAA* sequence analysis$, or on the number of patients identified in nationwide newborn screening programs*. Based on [18-24, 27, 28].

Table 1. Reported Incidence of Pompe Disease in Different Populations

Population	Reported Incidence	Reference
African American	1:14.000 (Classic infantile)	Reuser et al 2018, Hopkins et al 2015
Netherlands	1:40.000 (Combined)[1] 1:138.000 (Classic infantile) 1:57.000 (Late onset)	Ausems et al 1999, Poorthuis et al 1999
US	1:40.000 (Combined)	Martiniuk et al 1998
South China Taiwan	1:50.000 (Classic infantile) 1:17.000 (Combined) 1:52.000 (Classic infantile) 1:25.000 (Late onset)	Lin et al 1987, Chiang et al 2012, Chien et al 2015
Japan	1:34.000 (Late onset)	Momosaki et al 2019
European descent	1:100.000 (Classic infantile) 1:60.000 (Late onset)	Martiniuk et al 1998
Australia	1:146.000	Meikle et al 1999
Portugal	1:600.000	Pinto et al 2004

[1]Combined: all Pompe disease phenotypes

More accurate estimates of the prevalence of Pompe disease are not available at present, but so far it is clear that the disease is not restricted to specific populations or geographic areas. The world-wide distribution of Pompe disease seems mainly determined by the natural occurrence of disease-associated variants in the *GAA* gene that are carried over from parents to children. The finding of *de novo* variants in affected offspring is extremely rare [31]. Migration has definitely played a role in the spreading of Pompe disease. The origin and spreading of the African-American variant c.2560C>T p.(Arg854*) has been documented best. It came from a North African tribe, members whereof migrated Westwards to the Atlantic coast from where they were shipped to North- and South-America as slaves [32-34]. The high incidence of this variant among American blacks, but also its finding among Colombian patients testifies of this event [35]. Other members of the tribe migrated southwards along the African West-coast and settled in Namibia. The same variant was also identified in a family with African-Arabian roots now living in the UK [2]. Likewise, the high frequency of c.1935C>A, p.(Asp645Glu), in the Taiwanese population must have arisen by a local founder effect, but was also encountered in patients with Asian ancestry from China, Japan, Thailand, and other countries [2, 36-39]. The most common disease-associated variant in Pompe disease is c.-32-13T>G p.(=),p.(0) (IVS1). This variant is widely spread among Caucasians and extremely rare in other ethnicities. The incidence is typically high in European countries, but the variant occurs also in countries where Europeans moved to like Canada, the USA, Australia, New Zealand, South-America and others [35, 40-43]. Other frequently encountered disease-associated *GAA* variants in the Caucasian population are the very small deletion c.525del, p.(Glu176Argfs*45), and the much larger deletion c.2481+102_2646+31del, p.(Gly828_Asn882del), which are both associated with the classic infantile phenotype when combined with a 'null' allele [44]. The most common disease-associated *GAA* variants in particular ethnic groups are listed in Table 2 [2].

Table 2. Common variants in particular ethnic groups

Pathogenic Variant (DNA)	Pathogenic Variant (Protein)	Ethnic Group	Allele Frequency	Phenotype
c.1935C>A	p.(Asp645Glu)	Taiwan	0.8	Classic infantile
c.2560C>T	p.(Arg854*)	Afro-American	0.5	Classic infantile
c.525del	p.(Glu176Argfs*45)	Dutch	0.34	Classic infantile
c.2481+102_2646+31del	p.(Gly828_Asn882del)	Dutch	0.23	Classic infantile
c.-32-13T>G	p.(=), p.(0)	Caucasian	0.4-0.9	Childhood/ Adult
c.2238G>C	p.(Trp746Cys)	Chinese/ Taiwanese	0.27	Childhood/ Adult

(adapted from Reuser et al, 2018, OMMBID, Chapter 135)

Signs and clinical spectrum of Pompe disease

The clinical expression of Pompe disease is dominated by skeletal muscle weakness leading to mobility problems. Affected babies are 'floppy'. Arm and leg muscles are weak, and they may be difficult to feed due to 'facial' muscle weakness [1]. In affected adults, a difference between proximal and distal muscle weakness is evident. The mobility problems are mainly caused by muscle weakness in the upper legs and the limb girdle [45]. Weakness of the para-spinal muscles plays a role too and can cause scoliosis, lordosis and kyphosis in affected young children. The reason why scoliosis is more frequent in childhood than adult patients is likely reflecting the fact that neuromuscular scoliosis typically develops over time with increased vulnerability to abnormal forces in a growing and developing musculoskeletal system [46]. Respiratory problems due to weakness of the diaphragm are a less visible but equally serious problem [45]. While the disease progresses, the vital capacity reduces, and this is most notable in supine position [47]. Apart from skeletal muscle, other vital systems are affected [48]. The floppy babies typically have cardiomegaly (a big heart) associated with hypertrophic cardiomyopathy. The wall of the left ventricle and the cardiac septum thicken initially, but then weaken causing cardiac dilatation and cardiac failure [1, 44, 49]. The picture described here for affected infants is typically called classic-infantile Pompe disease as it addresses the phenotype originally described by Drs Pompe and Putschar [3, 5, 50]. There is no evidence that cardiomyopathy is a prominent symptom in any other form of Pompe disease; from childhood to adult cases. Though cardiac involvement in children and adults with Pompe disease with the common c.-32-13T>G variant has been reported it is considered very rare [51, 52].

Apart from skeletal muscle weakness in all forms of Pompe disease and additional cardiomyopathy in the most severe infantile form, there are other organs involved leading to less visible symptoms. It concerns in part problems relating to smooth muscle pathology such as esophageal, stomach, intestinal, and bladder dysfunction, and weakness of the blood vessel wall [53-55]. White matter changes were recently reported as novel feature of Pompe disease in infants who survived far beyond the natural expectation thanks to enzyme replacement therapy [56, 57].

In short, Pompe disease presents as a spectrum of phenotypes differing mainly by 'yes or no' cardiac involvement, age at onset, rate of disease progression, and severity of symptoms [19, 50]. Tables 3A and 3B provide an overview of symptoms that have been reported in infants, children and adults with Pompe disease [2, 44, 58, 59].

Table 3A. Common findings at presentation of Classic Infantile Pompe disease

Clinical signs	Frequency (%)
Left ventricular hypertrophy	83-100
Cardiomegaly	92-100
Cardiomyopathy	88
Congestive heart failure	50
Hypotonia/muscle weakness	52-96
Hepatomegaly	29-90
Respiratory distress	41-78
Macroglosia	29-62
Feeding difficulties	57
Failure to thrive	53

Reuser et al, 2018, OMMBID, Chapter 135, van den Hout et al, 2003, Kishnani et al 2006

Table 3B. Common findings in Childhood and Adult Pompe disease

System	Symptoms and manifestations
Muscular	Proximal muscles weaker than distal Lower extremities weaker that upper Weakness common in: o Pelvic girdle/lower extremity musculature o Trunk muscles o Shoulder girdle musculature and scapular stabilizers o Neck flexors typically weaker than neck extensors o Respiratory muscles o Facial muscles Fatigue Exercise intolerance Physical limitations and functional loss of: o Ambulation o Independent transitions between positions o Independent breathing Compensatory movements / deviations including: o Beevor sign (abnormal upward umbilicus movement upon truncal flexion while the patient is in a supine position) o Positive Trendelenburg's sign o Waddling gait o Use of modified Gower's maneuver Contractures
Respiratory	Sleep apnea Decreased vital capacity Diaphragm weakness Dyspnea
Musculoskeletal	Osteoporosis, osteopenia, low bone mineral density (BMD) Scoliosis Excessive kyphosis and/or lordosis Scapular winging Rigid spine syndrome (RSS)
Neurological	Sensorineural hearing loss (HL)

Table 3B continued.

System	Symptoms and manifestations
Vascular*	Cerebral/intracranial aneurysms, aneurysmal thrombosis
	Lacunar encephalopathy
	Dolichoectasia of basilar artery
	Microhemorrhages/intraparenchymal hemorrhage
	Aortic abnormalities
Cardiac	Cardiac hypertrophy (rare cardiomyopathy)
	Abnormal heart rhythm such as: o Supraventricular tachycardia (SVT) o Wolff-Parkinson-White (WPW) syndrome
Gastrointestinal/ Genitourinary	Hepatomegaly
	Abdominal discomfort
	Constipation
	Chronic diarrhea
	Poor weight gain/overweight
	Urinary and bowel incontinence due to weakness of anal sphincter
Pain and fatigue	Muscle pain, cramps (Lower back, shoulders, upper legs/thighs)
	Mental/physical tiredness from muscle weakness
	Exhaustion (fragmented sleep, daytime sleepiness)

Chan J. et al, 2017. *This maybe the case, but is currently under investigation.

From disease-associated *GAA* variants to clinical signs

How to explain the clinical heterogeneity of Pompe disease? The very first and wrong hypothesis was that classic-infantile and 'adult' Pompe disease -initially called 'the muscular form of GSDII'- were two separate genetic diseases [9, 60]. However, later studies provided indirect evidence that all forms of Pompe disease are related to disease-associated variants in one and the same gene coding for acid α-glucosidase (*GAA*) [16, 61, 62]. The *GAA* gene was cloned and sequence variants were identified that differentiated, indeed, between very severe and less severe forms of Pompe disease (discussed below) [12, 13, 63-65]. Cloning of the *GAA* gene also allowed deeper insight in GAA protein expression, post-translational processing including proteolytic cleavage, modification wih mannose 6-phosphate groups, and intra-cellular routing to the lysosomes, as investigated with immuno-biochemical approaches. Disease-associated variants have been found to affect each of these processes, as well as basic gene expression processes. For instance, a stretch of *GAA* sequence may be missing whereby transcription of the gene into a stable messenger (pre-mRNA) becomes impossible. Also, the introduction of a pre-mature stop codon through a mutational event can lead to what is called 'non-sense mediated mRNA decay'. Defects in splicing also can play a role. The normal pre-mRNA transcript is spliced, whereby the introns are removed and the exons linked. Defects in this mechanism, caused by base pair changes, can lead to mRNA decay and total or partial GAA deficiency [66]. The majority of disease-associated variants concerns missense mutations (discussed below). In

those cases the correctly spliced mRNA is translated into GAA protein, but amino acid substitutions can lead to endoplasmic-reticulum-associated protein degradation (ERAD), or affect the fine structure of GAA so that its function is -partially- lost [67].

Essential steps in the biosynthesis of GAA are the co-translational transport of the protein into the lumen of the endoplasmic reticulum (ER), the folding and glycosylation at 7 positions that are encoded by the amino acid sequence Asn-Xaa-Ser/Thr, the post-translation modification of the carbohydrate chains enabling binding to the mannose 6-phosphate receptor and securing transport to the lysosomes [68, 69], and the proteolytic processing of GAA, which enhances the enzyme's capacity to degrade lysosomal glycogen. The different steps of GAA biosynthesis are depicted in Figure 2.

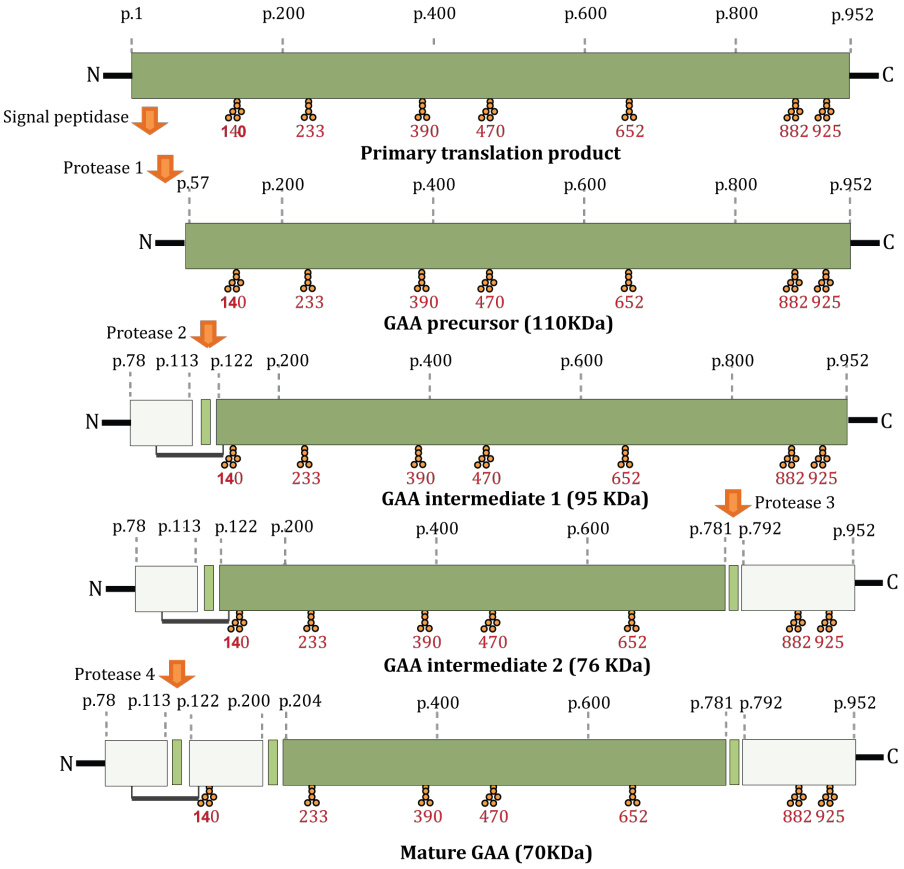


Figure 2. Maturation of GAA. GAA is synthesized as a precursor protein that matures by glycosylation and proteolytic processing. The posttranslational glycosylation followed by carbohydrate modification enables binding to the mannose 6-phosphate receptor, securing transport to the lysosomes. The proteolytic processing enhances the enzyme's capacity to degrade lysosomal glycogen. Adapted from [74].

GAA function

The fraction of glycogen that is degraded by GAA is imported from the cytoplasm where the main glycogen metabolism takes place. It is not known what signals are driving the lysosomal import of glycogen, but the mechanism of import is through autophagy. Cytoplasmic glycogen particles are sequestered from the bulk by a membrane that wraps around them. The vesicle (autophagosome) moves to the lysosomes where after the membranes of autophagosome and lysosome fuse. Glycogen enters the lysosome where it meets GAA that by itself is brought to the lysosomes by vesicular transport from Golgi-complex to lysosomes. Glucose, the degradation product of glycogen, is small enough to cross the lysosomal membrane and go back to the cytoplasm (Figure 3A), but glycogen is too big. Thus, glycogen will start to accumulate inside the lysosomes if the GAA activity is too low to degrade it all within a certain time frame (Figure 3B).

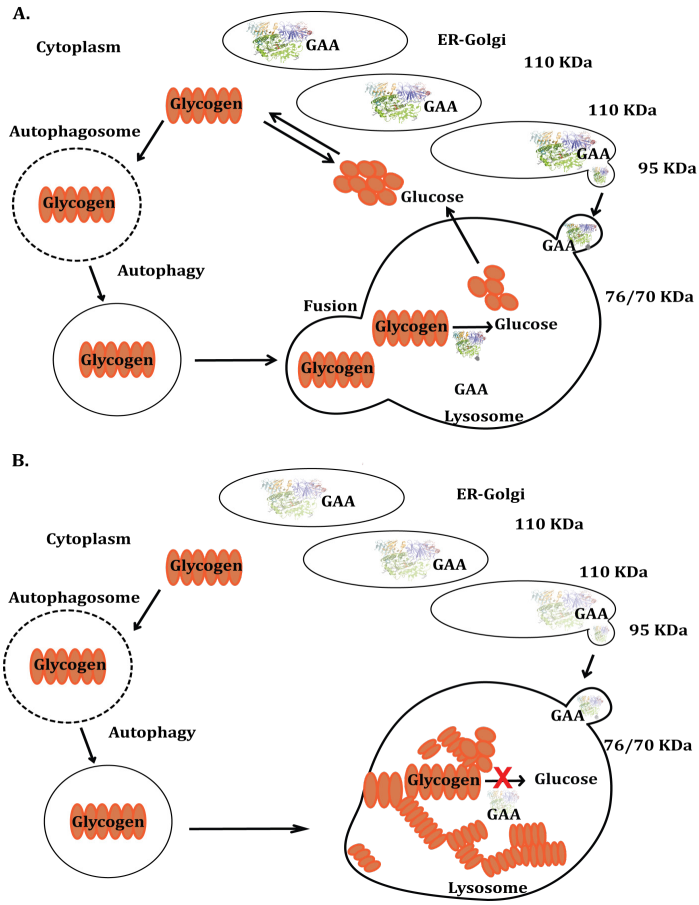


Figure 3. How glycogen enters lysosomes via autophagy, and GAA via mannose 6-phosphate receptor-mediated vesicular transport. A) Breakdown of glycogen to glucose in health. B) Accumulation of glycogen in disease. Based on [33].

The intra-cellular biosynthesis and transport of GAA is depicted in greater detail in Figure 4. The *GAA* gene is transcribed into *GAA* mRNA, which is then translated into GAA protein from ribosomes that are attached to the Rough Endoplasmic Reticulum (RER). Newly synthesized enzyme is glycosylated, phosphorylated, folded, and processed during transport through the RER and Golgi (Figure 4, stage A). The protein is synthesized as an enzymatically inactive precursor protein with a molecular mass of approximately 110 kD. The precursor enters the endoplasmic reticulum while the amino acids are being assembled, is glycosylated at the spot, and undergoes sugar chain modifications and proteolytic cleavage. One of these latter modifications involves the acquisition of a mannose 6-phosphate group, which allows for the recognition of the enzyme by the mannose 6-phosphate receptor (M6P) that collects GAA at exit from the trans-Golgi cisternae and directs it to the lysosomes (Figure 4, stage B). GAA leaves the Golgi complex in a vesicle that delivers its content to late endosomes and lysosomes through fusion of membranes. Once GAA has reached the late endosomes, the receptor-ligand complex dissociates due to the low pH, and the enzyme is delivered to the lysosomes. The mannose 6-phosphate receptor cycles back to the trans-Golgi for the next round of sorting [70] (Figure 4, stage C). Figure 4 further illustrates that the mannose 6-phosphate receptor can take a sojourn to the plasma membrane to bind, collect, and internalize mannose 6-phosphate containing GAA from the extra-cellular environment by endocytosis; enzyme replacement therapy is based on this mechanism. Endocytosis refers to a major delivery stream c.q. molecular machinery used for the internalization of cell surface and extra cellular materials (the latter materials are captured by receptors on the cell membrane). The plasma membrane invaginations form a vesicle that delivers the included material to the lysosomes [71-73]. Proteolytic processing of the GAA precursor takes place during transport from ER to lysosomes and inside the lysosomes whereby the functional, enzymatic, activity of GAA increases. The process can be followed showing the gradual appearance of molecular forms ranging from initially 110 kD via 95 kD to 76 and 70 kD [74, 75].

GAA cleaves both the 1,4 as well as the 1,6 glucose linkages in glycogen and disassembles the glucose polymer to monosaccharide units [16]. As a result of GAA deficiency, an excessive intra-lysosomal glycogen accumulation occurs (as depicted in Figure 3B) having a harmful impact on muscle cells, which become irreversibly damaged (Figure 5).

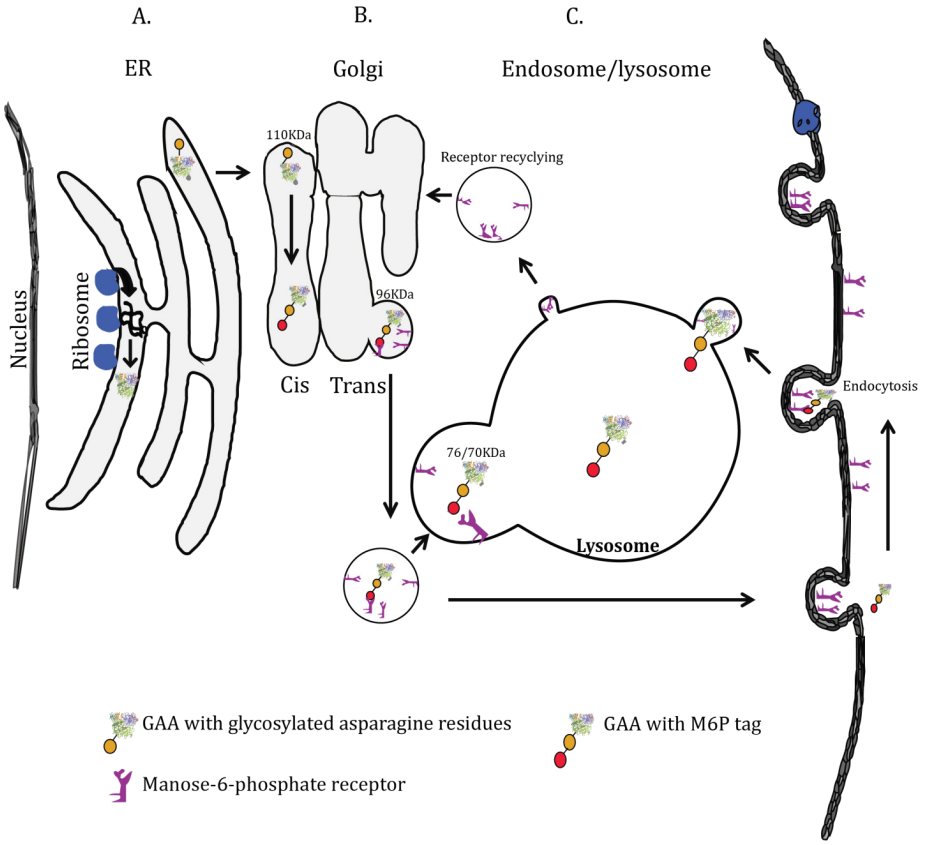


Figure 4. Synthesis and mannose 6-phosphate receptor mediated transport of GAA. A) GAA enters the endoplasmic reticulum (ER) co-translationally where it is glycosylated and folded. B) From there, GAA is transported via the Golgi complex where mannose residues are phosphorylated, and onwards to the lysosomes. The proteolytic processing occurs partially *en route* and in the lysosomes. C) The mannose 6-phosphate receptor secures selective transport of GAA from cis-Golgi to endosomes/lysosomes and recycles to the cis-Golgi or plasma membrane. Adapted from [73, 113].

Consequences of glycogen storage

Glycogen storage is the trigger to the following complex cascade of pathologic events [76]. In the simplest sense, the lysosomal system initially tries to accommodate the excess glycogen, as it seems, by 'growing' in lysosomal volume. The lysosomal system starts to dysfunction and this affects first of all the autophagic system that can no longer deliver its contents to the lysosomes, but also the endocytic system that communicates with the autophagic system and delivers its content derived from the extracellular space also to the lysosomes. When these processes are blocked, the cell ends up in chaos and secondary changes also occur in mitochondrial energy supply.

Apparently, these changes together are so damaging to the integrity of the cell that it stops functioning and dies. Initially, the lysosomal, autophagic, and lipofuscin inclusion by itself leads to mechanical loss of muscle contractile power and muscle function. These stages in the disease progression of Pompe disease are illustrated in Figure 5 [76].

The last step, from loss of muscle function and muscle mass to clinical signs, can be explained by an imbalance that arises between muscle cell repair and muscle regeneration on the one hand and continuous loss of muscle tissue at the other.

The Pompe disease variant database

The *GAA* gene spans approximately 20 kb and is located on the long arm of chromosome 17q25.2-q25.3 [12, 13, 77]. It contains 20 exons (the first exon is noncoding) [78] and comprises 2859 nucleotides (letters) of coding sequence for the lineup of 952 amino acids as GAA building blocks [78].

Pompe disease is an autosomal recessive disorder, which means that both inherited *GAA* alleles (gene copies) need to carry a disease-associated variant before the disease manifests itself. This happens when both parents are passing on one disease-associated variant to their offspring. If both parents are 'carrier' they have a 25% chance that their offspring is affected [79]. As of July 2019, 562 *GAA* variants had been identified and reported; of which 422 are listed as disease-associated and 140 are classified as unknown (www.pompevariantdatabase.nl) [80].

The way in which disease-associated *GAA* variants result in GAA deficiency, leading to lysosomal glycogen storage, inducing muscle damage, culminating in muscle weakness and related clinical signs is shown in Figure 6 [15, 81].

The classic infantile form of Pompe disease is rather homogeneous in its clinical presentation [19, 44, 50, 58], but the age of onset and the age related severity of symptoms varies substantially among childhood and adult patients (non-classic forms) [82, 83]. It seems logic to assume that a combination of two very severe 'null' (based on the fact that GAA enzyme activity is completely abolished) disease-associated *GAA* variants results in a complete GAA enzyme deficiency (0-2%) causing classic infantile Pompe disease, whereas a combination of less severe 'non-null' variants leads to less severe phenotypes, in which some residual enzyme activity (2-25%) remains and the disease can manifest at any time throughout life (Figure 7).

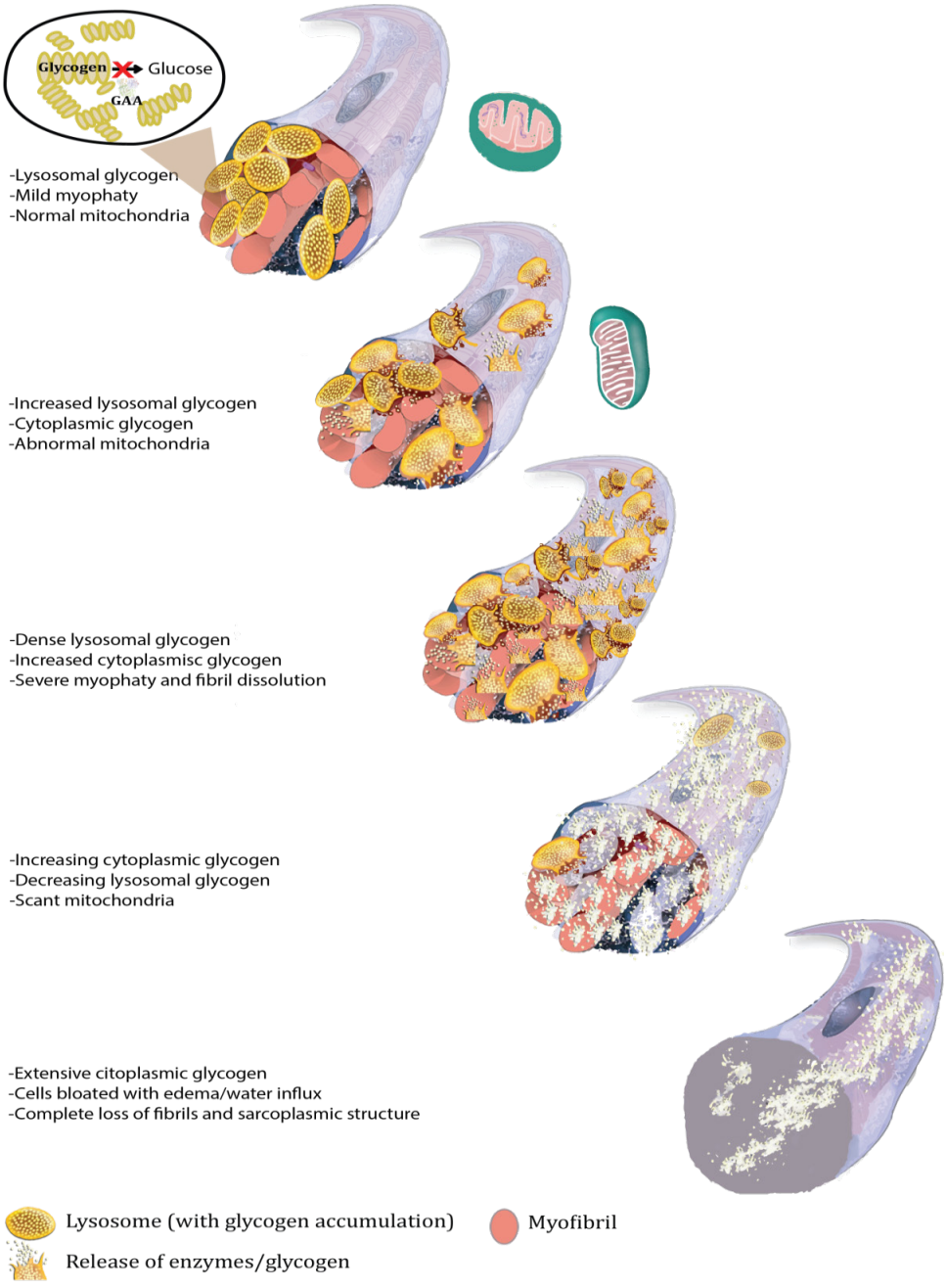


Figure 5. From lysosomal glycogen storage to muscle pathology. GAA deficiency induces a cascade of pathologic events starting with the lysosomal accumulation of glycogen. The ensuing lysosomal dysfunction affects autophagy, endocytosis and other cellular processes. In addition, the formation of vacuolar inclusions hampers muscle contraction mechanically. The muscle cells become irreversibly damaged and are locally replaced by connective tissue. Adapted from [76].

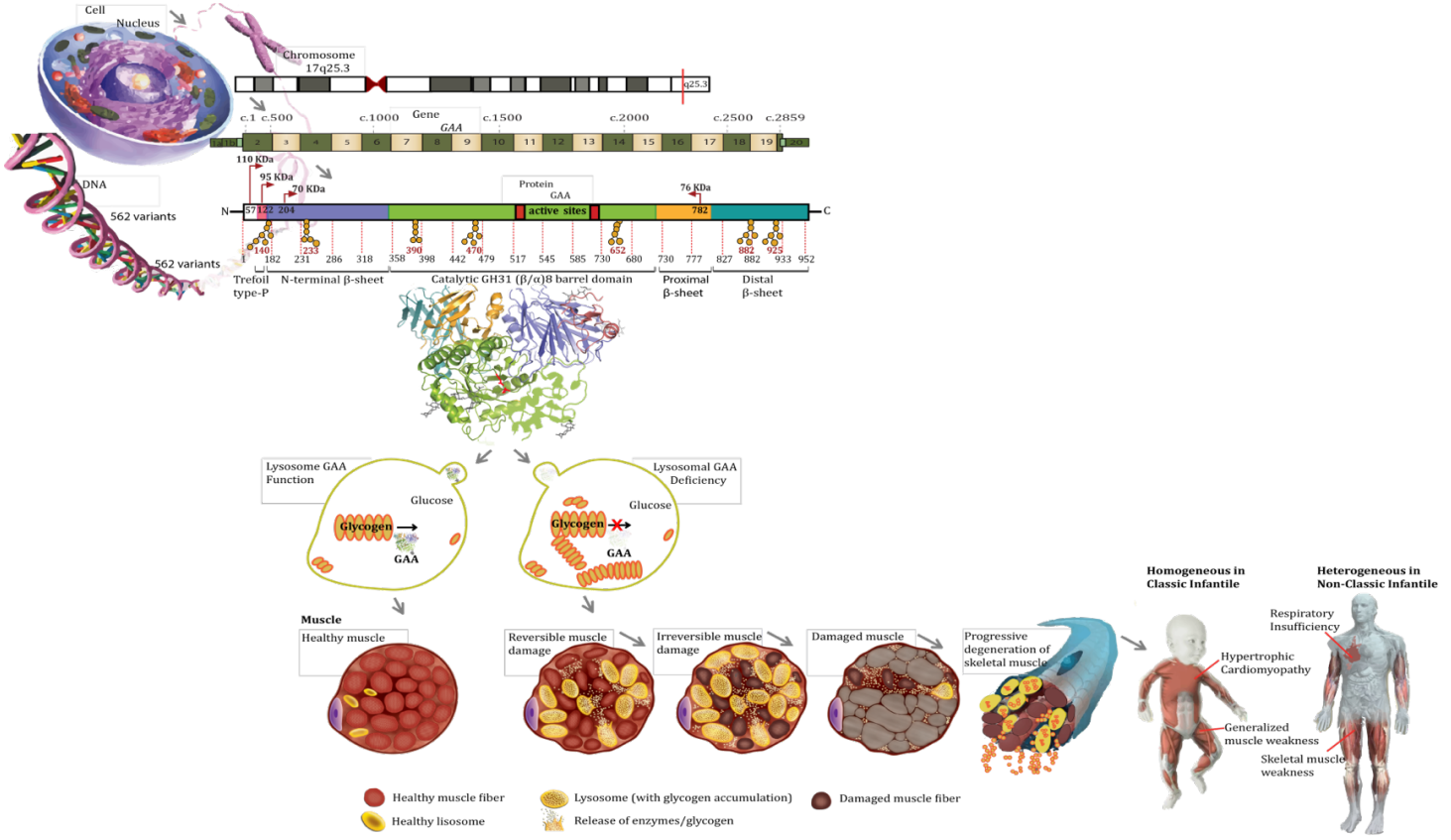


Figure 6. From DNA to Pompe disease phenotypes. Disease-associated *GAA* variants cause GAA deficiency, which results in intra-lysosomal glycogen storage that induces muscle damage, culminating in muscle weakness and clinical signs. Depending on the nature of the variants the phenotype can manifest at very early age and be rapidly progress, or late in adulthood with a usually slower progression. A composition of elements from several publications [15, 81] and [https://repub.eur.nl/pub/9453/].

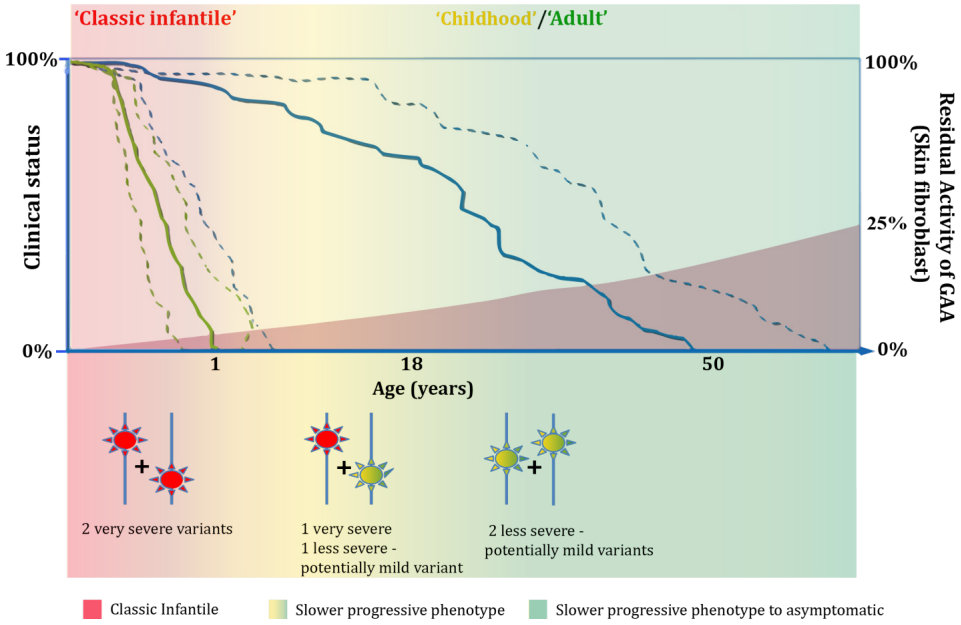


Figure 7. Genotype-phenotype correlations in Pompe disease. The left vertical axis depicts the clinical status of the patient; from 100% symptom free to deceased; the lines indicate the declining clinical condition over time. The right vertical axis indicates the approximate percentage of GAA activity compared to average normal. Patients with classic infantile Pompe disease have hardly any measurable GAA activity, affected children and adults usually more. The level of GAA activity is primarily dictated by the type and combination of disease-associated sequence variants on the two *GAA* alleles. Based on [1].

Whether such correlation exists can be investigated by systematic review of all *GAA* sequence variants that have been published so far with regard to their predicted effect or the combination in which they were identified in patients with well-described phenotypes.

The previous version of the Pompe disease mutation database tried to keep track of all published *GAA* sequence variants and, when possible, assigned to each variant a severity rating based on functional studies, *in silico* analysis and/or available clinical information [84, 85]. The accuracy of the predictions in this previous version of the database has never been challenged, but was not investigated either using unanimous criteria. Two chapters of this Thesis (**Chapters 3 and 4**) are devoted specifically to that goal: they systematically review and investigate the functional effect of each listed *GAA* variant not only on the basis of *in silico* predictions or functional assays, but also by deciphering all the different combinations in which the variants occur in pre-specified phenotypes: classic-infantile or childhood/adult Pompe disease. These two chapters describe the extension of the updated Pompe disease *GAA* variant database

at www.pompevariantdatabase.nl, which now includes a number of new features to improve its utility. Most important is to distinguish *GAA* variants leading to total loss of function from those leading to partial loss of function in order to assist in the diagnosis of newborns, predict the clinical course, and provide support to geneticists, genetic counselors, and affected families [80, 86].

From genetic modifying factors to markers of disease progression

Modifying factors, the way they are mentioned here, are circumstantial non-pathogenic variants in the genetic constitution of patients with Pompe disease that can affect (modify) the clinical outcome of disease severity and/or GAA deficiency independent of the *GAA* genotype. For instance, some people are more muscular than others from birth on, and some have a tendency to gain weight while others don't. Undoubtedly, there are numerous genetic factors and combinations thereof that co-determine the outcome of a GAA deficiency. For instance, there could be factors affecting the onset of symptomatology by delaying or accelerating lysosomal glycogen storage. They could do so by variably regulating the mitochondrial balance and function or the autophagic system [87-89]. Other factors may influence disease progression through their potential effect on the repair of cell damage and the system of muscle renewal via satellite cell activation [90, 91]. Obviously, none of these factors would have great impact on the clinical course of Pompe disease when the GAA enzyme activity is totally lacking, but when residual activity remains a modifying effect is likely. In fact, there is quite strong evidence for the contribution of modifying factors to the clinical course since patients with the most common disease-associated *GAA* genotype [c.-32-13T>G (IVS1) / 'null'] have been described with ages of onset varying from under 1 year to 67 years of age. The c.-32-13T>G (IVS1) variant is the most frequent *GAA* variant among Caucasians with Pompe disease and is observed in more than two-thirds of Caucasian patients worldwide. The c.-32-13T>G transversion is located in front of the splice acceptor site of exon 2, and results in the skipping of this exon. It also allows a low (10-15%) level of normal splicing, which explains the late onset phenotype of the patients. [92-96]. Usually, the IVS1 variant is combined with a 'null' variant on the second allele. A 'null' variant can be any variant that completely abolishes GAA enzyme activity.

Recently, a novel insertion (m.317-318insCCC) in the mitochondrial D-loop region was detected as a new mitochondrial variant that could influence disease progression in patients with Pompe disease. This study showed that D-loop variant frequency is higher in infantile patients than in adult patients. It seems that mitochondrial variants may have a secondary role in the pathogenesis of Pompe disease [89]. So far, only two genes have been identified as candidate trans-acting modifying factor by occurring in more than

one genetic variety (allelic form). These are: 1) an insertion (I) or deletion (D) of a 287-base-pair alu repeat within intron 16 of the angiotensin converting enzyme (*ACE*) gene, resulting in an II, DI or DD genotype, and 2) the R577X polymorphism in the alpha-actin-3 (*ACTN3*) gene, caused by a C to T transversion at position 1,747 in exon 16, and resulting in an XX, RX or RR genotype. A first study found association between these polymorphisms and an earlier age of onset and muscle pain, however results obtained by a second group could not find a correlation between ACE genotype and disease severity. It should however be noted that these results could have been influenced by the response to ERT. Clearly, more research is required to understand the mechanism underlying *ACE/ACTN3* polymorphism, disease severity and/or response to the ERT [97-99]. The role of the *ACE* gene is further investigated in **Chapter 5** of this Thesis, while **Chapter 6** presents the results of an extensive study searching for genes that can be used as prognostic indicator of disease progression via analysis of gene expression at the RNA level (RNAseq analysis) in two patient populations with *GAA* genotype [c.-32-13T>G / 'null', in 35/39 cases], but age of onset under 8 years versus age of onset >35 years.

Diagnosis

The clinical diagnosis is based on the presentation of clinical symptoms and the awareness of the treating physician about a rare disorder like Pompe disease. In recent years, several patients were diagnosed among groups of patients with high CK (serum creatin kinase level) pointing to a muscle disorder and or limb-girdle weakness of unknown cause [100, 101]. Muscle biopsies usually reveal glycogen storage, but by hitting a well-preserved muscle fiber bundle the diagnosis can be missed. Another pitfall of diagnosing Pompe disease on the basis of a muscle biopsy is that the glycogen washes out when the material is not properly fixed. Nowadays, an enzymatic GAA activity assay in which bloodspots are used as sample source is an often used, first tier, test for screening purposes, but not suitable for establishing the final diagnosis [102]. An assay on leukocytes is more reliable for that purpose as long as the glucoamylase activity of the granulocytes is inhibited with acarbose. The use of glycogen as natural substrate in the assay system provides a better distinction between affected and unaffected individuals than the use of the artificial 4-methylumbellifery α-D-glucopyranoside substrate (4MUG), but marks individuals with a *GAA2* pseudodeficiency as patients. Cultured fibroblasts derived from a skin biopsy can also be used as diagnostic sample source [103, 104]. They provide the most reliable results, and patients with classic-infantile Pompe disease can be distinguished from those with less severe and less progressive phenotypes on the basis of residual GAA enzyme activity. The functional assays based on measuring the GAA enzyme activity using different sample types and different procedures comparing one method with the other to determine the cut-off value

of each assay below which the outcome is taken as positive for Pompe disease, and which method applies best in cases of GAA pseudodeficiency of the gene is further investigated in **Chapter 2** of this Thesis.

DNA analysis can be applied for diagnosing Pompe disease as long it is realized that a positive diagnosis can only be made if two *GAA* variants are identified, one in each *GAA* allele, with affirmed pathologic effect.

The choice between enzyme activity assays as opposed to a DNA testing is partially dependent on the diagnostic situation. In a 'novel presenting case' scenario an enzyme activity assay has the advantage that it measures the functional effect of all disease-associated *GAA* variants in the patient's two *GAA* alleles, while some may escape detection by DNA analysis. Currently the majority (~100%) of patients within our center carry 2 disease-associated *GAA* variants, while few are known with only one *GAA* variant by standard diagnostic DNA analysis. Reuser et al, 2019 recently published that up to 10% of patients in the Pompe Registry (initiated by Sanofi-Genzyme) had only one *GAA* variant identified based on standard DNA diagnostics [105]. To discriminate between unaffected carriers and patients with Pompe disease, extended diagnostic analyses may help to identify the second DNA variant in these rare cases (In 't Groen et al., in revision). In a family scenario with an index patient and known disease-associated *GAA* variants, DNA analysis is the far better method for distinguishing between (pre-symptomatic) affected, unaffected, and carrier than a GAA enzyme activity assays that may fail to do so, especially in adult-onset forms of Pompe disease with high residual GAA activity.

Therapy

Enzyme replacement therapy (ERT) with recombinant human GAA (rhGAA) was approved in 2006 for patients across the entire clinical spectrum, after decades of pre-clinical research and successful Phase 3 trials. Without going into much detail, as this would go beyond the scope of this thesis, it is generally accepted that infants benefit from treatment as it largely prevents and reverses the hypertrophic cardiomyopathy, improves the patients' development and mobility, and can extend the life expectancy by at least more than two decades in best cases. These effects have also been demonstrated in infants that do not produce any endogenous GAA, so called cross reactive immunological material (CRIM) negative patients. In these patients the effect of ERT might be counteracted by the formation of anti-rhGAA antibodies [106-108]. Over 95% of affected infants treated with ERT develop antibodies against rhGAA [109]. CRIM negative infants who usually develop high and sustained antibody titers tend to have

a poor clinical outcome [108]. In contrast, CRIM positive patients, who synthesize a defective form of GAA, usually have lower antibody titers and a better response to ERT [110]. However, classic infantile Pompe disease is a very serious, life threatening condition, and close to 50% of the infants does not survive ventilator free for more than 5 years despite ERT. Less affected children and adults benefit from ERT by initial gain of muscle strengths and pulmonary function, but their disease progresses albeit more slowly than when untreated [111-113]. As in infants, the effect of ERT in children and adults varies on a patient-to-patient basis.

AIMS OF THE RESEARCH DESCRIBED IN THIS THESIS

Over the years great progress has been made with identifying sequence variants in the *GAA* gene and elucidating their effect on the natural course of Pompe disease, with developing ERT, and with optimizing diagnostic tests.

The Erasmus MC University Medical Center in Rotterdam, The Netherlands, has been actively involved in these activities and has, thereby, contributed substantially to the current understanding of the disease. For instance, the rapidly expanding number of *GAA* sequence variants that were discovered over the past 15 years has led to the start of the Pompe disease mutation database and to attempts at predicting the severity of each *GAA* variant.

The aims of the research described in this thesis can be summarized under the common denominator of "elucidating the genotype-phenotype correlation in Pompe disease", whereto the following research questions and goals were formulated:

1. How to diagnose Pompe disease enzymatically? A study based on diagnostic outcomes over a period of 28 years using several sample sources and different methods. (**Chapter 2**)
2. Upgrading of the Pompe disease mutation database and facilitation of the analysis of the genotype-phenotype correlation. A study in which all publicly available *GAA* sequence data is reviewed, entered into a new format of the Pompe disease mutation database, and pathogenic genotypes are linked to clinical phenotypes. (**Chapters 3 and 4**)
3. Does the common *ACE* I/D polymorphism contribute to the heterogeneity of Pompe disease? A study in which the previously claimed impact of this polymorphism on disease severity and effect of ERT is verified in a large sample collection. (**Chapter 5**)

4. Can disease markers be identified that correlate with symptom onset**?** A study wherein the genome-wide mRNA expression level of genes other than *GAA* are analyzed using cultured fibroblasts of affected children and adults with similar *GAA* genotypes (c.-32-13T>G / 'null', in 35/39 cases) **(Chapter 6)**

The results of these studies are discussed in broader context, overall conclusions are drawn, and future perspective are formulated in **Chapter 7**.

REFERENCES

1. van der Ploeg, AT, and Reuser, AJ (2008). Pompe's disease. *Lancet* **372**: 1342-1353.
2. Reuser, AJJ, Hirschhorn, R, and Kroos, MA (2019). Pompe Disease: Glycogen Storage Disease Type II, Acid α-Glucosidase (Acid Maltase) Deficiency. In: Valle, D, S Antonarakis, A Ballabio, A Beaudet and GA Mitchell (eds). *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Education: New York, NY.
3. Pompe, JC (1932). Over idiopathische hypertrofie van het hart. *Ned Tijdsch Geneesk* **76**: 304-311.
4. Melvin, JJ (2000). Pompe's disease. *Arch Neurol* **57**: 134-135.
5. Putschar, W (1932). Uber angeborne Glycogenspeicherkrankheit des herzens:Thesaurismosis glycogenica (v.Gierke). *Beitr Pathol Anat* **90**: 222-.
6. Zeidman, LA (2012). Johannes C. Pompe, MD, hero of neuroscience: the man behind the syndrome. *Muscle Nerve* **46**: 134-138.
7. Cori, GT (1954). Glycogen structure and enzyme deficiencies in glycogen storage disease. *Harvey Lectures* **8**: 145.
8. De Duve, C, Pressman, BC, Gianetto, R, Wattiaux, R, and Appelmans, F (1955). Tissue fractionation studies. 6. Intracellular distribution patterns of enzymes in rat-liver tissue. *Biochem J* **60**: 604-617.
9. Engel, AG (1969). Acid maltase deficiency of adult life. *Trans Am Neurol Assoc* **94**: 250-252.
10. Courtecuissf V, RP, Habib R, Monnier C, Demos J. (1965). Muscular glycogenosis caused by alpha-1,4-glucosidase deficiency simulating progressive muscular dystrophy. *Arch Fr Pediatr* **22**: 1153-1164.
11. Isch F, JJ, Sacrez R, Thiebaut F. (1966). Muscular glycogenosis of myopathic form caused by acid maltase deficiency. *Pediatrie* **21**: 71-86.
12. Solomon, E, Swallow, D, Burgess, S, and Evans, L (1979). Assignment of the human acid alpha-glucosidase gene (alphaGLU) to chromosome 17 using somatic cell hybrids. *Ann Hum Genet* **42**: 273-281.
13. D'Ancona, GG, Wurm, J, and Croce, CM (1979). Genetics of type II glycogenosis: assignment of the human gene for acid alpha-glucosidase to chromosome 17. *Proceedings of the National Academy of Sciences of the United States of America* **76**: 4526-4529.
14. Weil D, VCN, Gross MS, Frézal J. (1979). Localization of the gene for human acid alpha-glucosidase (alpha-GLUa) on the 17q21 to 17qter by interspecific hybridization. *Human Genet* **52**: 249-257.
15. Hirschhorn, RR, A.J. (2001). Glycogen storage disease type II; acid alpha-glucosidase (acid maltase) deficiency. In: C.R., B, A.L., Sly, W. & Valle, D, (ed). *The metabolic and molecular bases of inherited disease*, 8th ed. New York: McGraw-Hill.: New York. pp 3389-3420.
16. Hers, HG (1963). alpha-Glucosidase deficiency in generalized glycogen storage disease (Pompe's disease). *Biochem J* **86**: 11-16.
17. Platt, FM, d'Azzo, A, Davidson, BL, Neufeld, EF, and Tifft, CJ (2018). Lysosomal storage diseases. *Nat Rev Dis Primers* **4**: 27.
18. Momosaki, K, Kido, J, Yoshida, S, Sugawara, K, Miyamoto, T, Inoue, T, *et al.* (2019). Newborn screening for Pompe disease in Japan: report and literature review of mutations in the GAA gene in Japanese and Asian

patients. *J Hum Genet* **64**: 741-755.

19. Ausems, MG, Verbiest, J, Hermans, MP, Kroos, MA, Beemer, FA, Wokke, JH, *et al.* (1999). Frequency of glycogen storage disease type II in The Netherlands: implications for diagnosis and genetic counselling. *Eur J Hum Genet* **7**: 713-716.
20. Martiniuk, F, Chen, A, Mack, A, Arvanitopoulos, E, Chen, Y, Rom, WN, *et al.* (1998). Carrier frequency for glycogen storage disease type II in New York and estimates of affected individuals born with the disease. *Am J Med Genet* **79**: 69-72.
21. Chiang, SC, Hwu, WL, Lee, NC, Hsu, LW, and Chien, YH (2012). Algorithm for Pompe disease newborn screening: results from the Taiwan screening program. *Mol Genet Metab* **106**: 281-286.
22. Pinto, R, Caseiro, C, Lemos, M, Lopes, L, Fontes, A, Ribeiro, H, *et al.* (2004). Prevalence of lysosomal storage diseases in Portugal. *Eur J Hum Genet* **12**: 87-92.
23. Meikle, PJ, Hopwood, JJ, Clague, AE, and Carey, WF (1999). Prevalence of lysosomal storage disorders. *JAMA* **281**: 249-254.
24. Poorthuis, BJ, Wevers, RA, Kleijer, WJ, Groener, JE, de Jong, JG, van Weely, S, *et al.* (1999). The frequency of lysosomal storage diseases in The Netherlands. *Human Genetics* **105**: 151-156.
25. Ausems, MG, ten Berg, K, Kroos, MA, van Diggelen, OP, Wevers, RA, Poorthuis, BJ, *et al.* (1999). Glycogen storage disease type II: birth prevalence agrees with predicted genotype frequency. *Community Genet* **2**: 91-96.
26. Hopkins, PV, Campbell, C, Klug, T, Rogers, S, Raburn-Miller, J, and Kiesling, J (2015). Lysosomal storage disorder screening implementation: findings from the first six months of full population pilot testing in Missouri. *J Pediatr* **166**: 172-177.
27. Lin, CY, Hwang, B, Hsiao, KJ, and Jin, YR (1987). Pompe's disease in Chinese and prenatal diagnosis by determination of alpha-glucosidase activity. *Journal of Inherited Metabolic Disease* **10**: 11-17.
28. Chien, YH, Lee, NC, Chen, CA, Tsai, FJ, Tsai, WH, Shieh, JY, *et al.* (2014). Long-Term Prognosis of Patients with Infantile-Onset Pompe Disease Diagnosed by Newborn Screening and Treated since Birth. *J Pediatr* **166**: 985-991.
29. Mechtler, TP, Stary, S, Metz, TF, De Jesus, VR, Greber-Platzer, S, Pollak, A, *et al.* (2012). Neonatal screening for lysosomal storage disorders: feasibility and incidence from a nationwide study in Austria. *Lancet* **379**: 335-341.
30. Palmio, J, Auranen, M, Kiuru-Enari, S, Lofberg, M, Bodamer, O, and Udd, B (2014). Screening for late-onset Pompe disease in Finland. *Neuromuscul Disord* **24**: 982-5.
31. Huie, ML, Chen, AS, Brooks, SS, Grix, A, and Hirschhorn, R (1994). A de novo 13 nt deletion, a newly identified C647W missense mutation and a deletion of exon 18 in infantile onset glycogen storage disease type II (GSDII). *Human Molecular Genetics* **3**: 1081-1087.
32. Becker, JA, Vlach, J, Raben, N, Nagaraju, K, Adams, EM, Hermans, MM, *et al.* (1998). The African origin of the common mutation in African American patients with glycogen-storage disease type II. *Am J Hum Genet* **62**: 991-994.
33. Raben, N, Plotz, P, and Byrne, BJ (2002). Acid alpha-glucosidase deficiency (glycogenosis type II, Pompe

disease). *Curr Mol Med* **2**: 145-166.

34. Adams, EM, Becker, JA, Griffith, L, Segal, A, Plotz, PH, and Raben, N (1997). Glycogenosis type II: a juvenile-specific mutation with an unusual splicing pattern and a shared mutation in African Americans. *Human Mutation* **10**: 128-134.
35. Nino, MY, Mateus, HE, Fonseca, DJ, Kroos, MA, Ospina, SY, Mejia, JF, *et al.* (2013). Identification and Functional Characterization of GAA Mutations in Colombian Patients Affected by Pompe Disease. *JIMD Rep* **7**: 39-48.
36. Hermans, MM, de Graaff, E, Kroos, MA, Wisselaar, HA, Willemsen, R, Oostra, BA, *et al.* (1993). The conservative substitution Asp-645-->Glu in lysosomal alpha-glucosidase affects transport and phosphorylation of the enzyme in an adult patient with glycogen-storage disease type II. *Biochemical Journal* **289**: 687-693.
37. Amarinthnukrowh, P, Tongkobpetch, S, Kongpatanayothin, A, Suphapeetiporn, K, and Shotelersuk, V (2010). p.D645E of Acid alpha-Glucosidase Is the Most Common Mutation in Thai Patients with Infantile-Onset Pompe Disease. *Genet Test Mol Biomarkers* **14**: 835-837.
38. Tsujino, S, Huie, M, Kanazawa, N, Sugie, H, Goto, Y, Kawai, M, *et al.* (2000). Frequent mutations in Japanese patients with acid maltase deficiency. *Neuromuscul Disord* **10**: 599-603.
39. Liu, X, Wang, Z, Jin, W, Lv, H, Zhang, W, Que, C, *et al.* (2014). Clinical and GAA gene mutation analysis in mainland Chinese patients with late-onset Pompe disease: identifying c.2238G inverted question mark> inverted question markC as the most common mutation. *BMC Med Genet* **15**: 141.
40. Golden-Grant, K, Merritt, JL, 2nd, and Scott, CR (2015). Ethical considerations of population screening for late-onset genetic disease. *Clin Genet* **88**: 589-592.
41. Hermans, MM, van Leenen, D, Kroos, MA, Beesley, CE, Van Der Ploeg, AT, Sakuraba, H, *et al.* (2004). Twenty-two novel mutations in the lysosomal alpha-glucosidase gene (GAA) underscore the genotype-phenotype correlation in glycogen storage disease type II. *Hum Mutat* **23**: 47-56.
42. Oba-Shinjo, SM, da Silva, R, Andrade, FG, Palmer, RE, Pomponio, RJ, Ciociola, KM, *et al.* (2009). Pompe disease in a Brazilian series: clinical and molecular analyses with identification of nine new mutations. *J Neurol* **256**: 1881-1890.
43. Dubrovsky, A, Corderi, J, Karasarides, T, and Taratuto, AL (2012). Pompe disease, the must-not-miss diagnosis: A report of 3 patients. *Muscle Nerve* **47**: 594-600.
44. Van den Hout, HM, Hop, W, Van Diggelen, OP, Smeitink, JA, Smit, GP, Poll-The, BT, *et al.* (2003). The natural course of infantile Pompe's disease: 20 original cases compared with 133 cases from the literature. *Pediatrics* **112**: 332-340.
45. van Capelle, CI, van der Meijden, JC, van den Hout, JM, Jaeken, J, Baethmann, M, Voit, T, *et al.* (2016). Childhood Pompe disease: clinical spectrum and genotype in 31 patients. *Orphanet J Rare Dis* **11**: 65.
46. 46. Roberts, M, Kishnani, PS, van der Ploeg, AT, Muller-Felber, W, Merlini, L, Prasad, S, *et al.* (2011). The prevalence and impact of scoliosis in Pompe disease: Lessons learned from the Pompe Registry. *Mol Genet Metab* **104**: 574-582.
47. van der Beek, NA, van Capelle, CI, van der Velden-van Etten, KI, Hop, WC, van den Berg, B, Reuser, AJ, *et al.* (2011). Rate of progression and predictive factors for pulmonary outcome in children and adults with

Pompe disease. *Mol Genet Metab* **104**: 129-136.

48. Filosto, M, Todeschini, A, Cotelli, MS, Vielmi, V, Rinaldi, F, Rota, S*, et al.* (2013). Non-muscle involvement in late-onset glycogenosis II. *Acta Myol* **32**: 91-94.
49. Morales, JA, and and Anilkumar, AC (2019). Glycogen Storage Disease, Type II (Pompe Disease). In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; January, 2019: https://www.ncbi.nlm.nih.gov/books/NBK470558/.
50. Gungor, D, and Reuser, AJ (2013). How to describe the clinical spectrum in Pompe disease? *Am J Med Genet A* **161A**: 399-400.
51. Soliman, OI, van der Beek, NA, van Doorn, PA, Vletter, WB, Nemes, A, Van Dalen, BM*, et al.* (2008). Cardiac involvement in adults with Pompe disease. *J Intern Med* **264**: 333-339.
52. van der Beek, NA, Soliman, OI, van Capelle, CI, Geleijnse, ML, Vletter, WB, Kroos, MA*, et al.* (2008). Cardiac evaluation in children and adults with Pompe disease sharing the common c.-32-13T>G genotype rarely reveals abnormalities. *J Neurol Sci* **275**: 46-50.
53. van Gelder, CM, van Capelle, CI, Ebbink, BJ, Moor-van Nugteren, I, van den Hout, JM, Hakkesteegt, MM*, et al.* (2012). Facial-muscle weakness, speech disorders and dysphagia are common in patients with classic infantile Pompe disease treated with enzyme therapy. *J Inherit Metab Dis* **35**: 505-511.
54. Sandhu, D, Rizvi, A, Kim, J, and Reshi, R (2014). Diffuse cerebral microhemorrhages in a patient with adult-onset Pompe's disease: a case report. *J Vasc Interv Neurol* **7**: 82-85.
55. Sacconi, S, Bocquet, JD, Chanalet, S, Tanant, V, Salviati, L, and Desnuelle, C (2010). Abnormalities of cerebral arteries are frequent in patients with late-onset Pompe disease. *J Neurol* **257**: 1730-1733
56. Ebbink, BJ, Aarsen, FK, van Gelder, CM, van den Hout, JM, Weisglas-Kuperus, N, Jaeken, J*, et al.* (2012). Cognitive outcome of patients with classic infantile Pompe disease receiving enzyme therapy. *Neurology* **78**: 1512-1518.
57. Ebbink, BJ, Poelman, E, Plug, I, Lequin, MH, van Doorn, PA, Aarsen, FK*, et al.* (2016). Cognitive decline in classic infantile Pompe disease: An underacknowledged challenge. *Neurology* **86**: 1260-1.
58. Kishnani, PS, Hwu, WL, Mandel, H, Nicolino, M, Yong, F, and Corzo, D (2006). A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease. *J Pediatr* **148**: 671-676.
59. Chan, J, Desai, AK, Kazi, ZB, Corey, K, Austin, S, Hobson-Webb, LD*, et al.* (2017). The emerging phenotype of late-onset Pompe disease: A systematic literature review. *Mol Genet Metab* **120**: 163-172.
60. Hudgson, P, Gardner-Medwin, D, Worsfold, M, Pennington, RJ, and Walton, JN (1968). Adult myopathy from glycogen storage disease due to acid maltase deficiency. *Brain* **91**: 435-462.
61. Engel, AG, Seybold, ME, Lambert, EH, and Gomez, MR (1970). Acid maltase deficiency: comparison of infantile, childhood, and adult types. *Neurology* **20**: 382.
62. Engel, AG, Gomez, MR, Seybold, ME, and Lambert, EH (1973). The spectrum and diagnosis of acid maltase deficiency. *Neurology* **23**: 95-106.
63. Weil, D, Van Cong, N, Gross, MS, and Frezal, J (1979). [Localization of the gene for human acid alpha-glucosidase (alpha-GLUa) on the 17q21 to 17qter by interspecific hybridization (author's transl)]. Localisation du gene de l'alpha-glucosidase acide (alpha-GLUa) sur le segment q21 a qter du chromosome

17 par l'hybridation cellulaire interspecifique. *Human Genetics* **52**: 249-257.

64. Solomon, E, and Barker, DF (1989). Report of the committee on the genetic constitution of chromosome 17. *Cytogenet Cell Genet* **51**: 319-337.
65. NCBI. Reference Sequence: NT_024871.11.
66. Bergsma, AJ, In 't Groen, SL, Verheijen, FW, van der Ploeg, AT, and Pijnappel, WP (2016). From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides. *Mol Ther Nucleic Acids* **5**: e361.
67. Zhong, N, Martiniuk, F, Tzall, S, and Hirschhorn, R (1991). Identification of a missense mutation in one allele of a patient with Pompe disease, and use of endonuclease digestion of PCR-amplified RNA to demonstrate lack of mRNA expression from the second allele. *American Journal of Human Genetics* **49**: 635-645.
68. Hasilik, A, and Neufeld, EF (1980). Biosynthesis of lysosomal enzymes in fibroblasts. Synthesis as precursors of higher molecular weight. *J Biol Chem* **255**: 4937-4945.
69. Hasilik, A, and Neufeld, EF (1980). Biosynthesis of lysosomal enzymes in fibroblasts. Phosphorylation of mannose residues. *J Biol Chem* **255**: 4946-4950.
70. Kornfeld, S (1992). Structure and function of the mannose 6-phosphate/insulinlike growth factor II receptors. *Annu Rev Biochem* **61**: 307-330.
71. Wenk, J, Hille, A, and von Figura, K (1991). Quantitation of Mr 46000 and Mr 300000 mannose 6-phosphate receptors in human cells and tissues. *Biochem Int* **23**: 723-731.
72. Kohler, L, Puertollano, R, and Raben, N (2018). Pompe Disease: From Basic Science to Therapy. *Neurotherapeutics* **15**: 928-942.
73. Metha Atul, Winchester Bryan (2012). *Lysosomal Storage Disorders: A Practical Guide*, Wiley-Blackwell, Ltda.
74. Moreland, RJ, Jin, X, Zhang, XK, Decker, RW, Albee, KL, Lee, KL, *et al.* (2005). Lysosomal acid alpha-glucosidase consists of four different peptides processed from a single chain precursor. *J Biol Chem* **280**: 6780-6791.
75. Hermans, Wisselaar, HA, Kroos, MA, Oostra, BA, and Reuser, AJJ (1993). Human lysosomal a-glucosidase: functional characterization of the glycosylation sites. *Biochem J* **289**: 681-686.
76. Thurberg, BL, Lynch Maloney, C, Vaccaro, C, Afonso, K, Tsai, AC, Bossen, E, *et al.* (2006). Characterization of pre- and post-treatment pathology after enzyme replacement therapy for pompe disease. *Lab Invest* **86**: 1208-1220.
77. Kuo, WL, Hirschhorn, R, Huie, ML, and Hirschhorn, K (1996). Localization and ordering of acid a-glucosidase (GAA) and thymidine kinase (TK1) by fluorescence in situ hybridization. *Hum Genet* **97**: 404-406.
78. Hoefsloot, LH, Hoogeveen-Westerveld, M, Reuser, AJJ, and Oostra, BA (1990). Characterization of the human lysosomal alpha-glucosidase gene. *Biochem J* **272**: 493-497.
79. Taglia, A, Picillo, E, D'Ambrosio, P, Cecio, MR, Viggiano, E, and Politano, L (2011). Genetic counseling in Pompe disease. *Acta Myol* **30**: 179-181.
80. Niño, MY, In 't Groen, SLM, Bergsma, AJ, van der Beek, N, Kroos, M, Hoogeveen-Westerveld, M, *et al.*

(2019). Extension of the Pompe mutation database by linking disease-associated variants to clinical severity. *Hum Mutat* **40**: 1954-1967.

81. Group, MPW, Al Jasmi, F, Al Jumah, M, Alqarni, F, Al-Sanna'a, N, Al-Sharif, F, *et al.* (2015). Diagnosis and treatment of late-onset Pompe disease in the Middle East and North Africa region: consensus recommendations from an expert group. *BMC Neurol* **15**: 205.
82. Gungor, D, de Vries, JM, Brusse, E, Kruijshaar, ME, Hop, WC, Murawska, M, *et al.* (2013). Enzyme replacement therapy and fatigue in adults with Pompe disease. *Mol Genet Metab* **109**: 174-8.
83. Wokke, JH, Escolar, DM, Pestronk, A, Jaffe, KM, Carter, GT, van den Berg, LH, *et al.* (2008). Clinical features of late-onset Pompe disease: A prospective cohort study. *Muscle Nerve* **38**: 1236-1245.
84. Kroos, M, Pomponio, RJ, van Vliet, L, Palmer, RE, Phipps, M, Van der Helm, R, *et al.* (2008). Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. *Hum Mutat* **29**: E13-26.
85. Kroos, M, Hoogeveen-Westerveld, M, Michelakakis, H, Pomponio, R, Van der Ploeg, A, Halley, D, *et al.* (2012). Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. *Hum Mutat* **33**: 1161-1165.
86. Niño, MY, In 't Groen, SLM, Bergsma, AJ, Hoogeveen-Westerveld, M, van der Ploeg, AT, and Pijnappel, WWMP (2019). The extended Pompe mutation database reveals distinct phenotypic spectra of Pompe disease-associated GAA variants. *Hum Mutat*.
87. Nascimbeni, AC, Fanin, M, Masiero, E, Angelini, C, and Sandri, M (2012). The role of autophagy in the pathogenesis of glycogen storage disease type II (GSDII). *Cell Death Differ* **9**: 1698-708.
88. Li, HM, Feeney, E, Li, L, Zare, H, Puertollano, R, and Raben, N (2013). Clearance of lysosomal glycogen accumulation by Transcription factor EB (TFEB) in muscle cells from lysosomal alpha-glucosidase deficient mice. *Biochem Biophys Res Commun* **291**: 272-6.
89. Bahreini, F, Houshmand, M, Modaresi, MH, Tonekaboni, H, Nafissi, S, Nazari, F, *et al.* (2016). Mitochondrial Copy Number and D-Loop Variants in Pompe Patients. *Cell J* **18**: 405-415.
90. Schaaf, GJ, van Gestel, TJ, Brusse, E, Verdijk, RM, de Coo, IF, van Doorn, PA, *et al.* (2015). Lack of robust satellite cell activation and muscle regeneration during the progression of Pompe disease. *Acta Neuropathol Commun* **3**: 65.
91. Sato, Y, Kobayashi, H, Higuchi, T, Shimada, Y, Ida, H, and Ohashi, T (2016). TFEB overexpression promotes glycogen clearance of Pompe disease iPSC-derived skeletal muscle. *Mol Ther Methods Clin Dev* **3**: 16054.
92. Alejaldre, A, Diaz-Manera, J, Ravaglia, S, Tibaldi, EC, D'Amore, F, Moris, G, *et al.* (2012). Trunk muscle involvement in late-onset Pompe disease: Study of thirty patients. *Neuromuscul Disord* **22 Suppl 2**: S148-154.
93. McCready, ME, Carson, NL, Chakraborty, P, Clarke, JT, Callahan, JW, Skomorowski, MA, *et al.* (2007). Development of a clinical assay for detection of GAA mutations and characterization of the GAA mutation spectrum in a Canadian cohort of individuals with glycogen storage disease, type II. *Mol Genet Metab* **92**: 325-335.
94. Wens, SC, van Gelder, CM, Kruijshaar, ME, de Vries, JM, van der Beek, NA, Reuser, AJ, *et al.* (2013).

Phenotypical variation within 22 families with Pompe disease. *Orphanet J Rare Dis* **8**: 182.

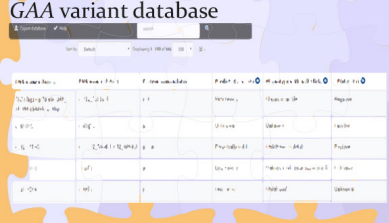
95. Huie, ML, Chen, AS, Tsujino, S, Shanske, S, DiMauro, S, Engel, AG*, et al.* (1994). Aberrant splicing in adult onset glycogen storage disease type II (GSDII): molecular identification of an IVS1 (-13T->G) mutation in a majority of patients and a novel IVS10 (+1GT->CT) mutation. *Human Molecular Genetics* **3**: 2231-2236.
96. Kroos, M, Hoogeveen-Westerveld, M, van der Ploeg, A, and Reuser, AJ (2012). The genotype-phenotype correlation in Pompe disease. *Am J Med Genet C Semin Med Genet* **160C**: 59-68.
97. De Filippi, P, Saeidi, K, Ravaglia, S, Dardis, A, Angelini, C, Mongini, T*, et al.* (2014). Genotype-phenotype correlation in Pompe disease, a step forward. *Orphanet J Rare Dis* **9**: 102.
98. de Filippi, P, Ravaglia, S, Bembi, B, Costa, A, Moglia, A, Piccolo, G*, et al.* (2010). The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. *Genet Med* **12**: 206-211.
99. Baek, RC, Palmer, R, Pomponio, RJ, Lu, Y, Ma, X, and McVie-Wylie, AJ (2016). The influence of a polymorphism in the gene encoding angiotensin converting enzyme (ACE) on treatment outcomes in late-onset Pompe patients receiving alglucosidase alfa. *Mol Genet Metab Rep* **8**: 48-50.
100. Gutierrez-Rivas, E, Bautista, J, Vilchez, JJ, Muelas, N, Diaz-Manera, J, Illa, I*, et al.* (2015). Targeted screening for the detection of Pompe disease in patients with unclassified limb-girdle muscular dystrophy or asymptomatic hyperCKemia using dried blood: A Spanish cohort. *Neuromuscul Disord* **25**: 548-53.
101. Lukacs, Z, Nieves Cobos, P, Wenninger, S, Willis, TA, Guglieri, M, Roberts, M*, et al.* (2016). Prevalence of Pompe disease in 3,076 patients with hyperCKemia and limb-girdle muscular weakness. *Neurology* **8**: 295-8.
102. Genge, A, and Campbell, N (2016). Reevaluating Muscle Biopsies in the Diagnosis of Pompe Disease: A Corroborative Report. *Can J Neurol Sci* **43**: 561-6.
103. Reuser, Verheijen, F, Kroos, M, Okumiya, T, Van Diggelen, O, Van der Ploeg, A*, et al.* (2010). Enzymatic and molecular strategies to diagnose Pompe disease. *Expert Opin Med Diagn* **4**: 79-89.
104. Kishnani, PS, Amartino, HM, Lindberg, C, Miller, TM, Wilson, A, and Keutzer, J (2014). Methods of diagnosis of patients with Pompe disease: Data from the Pompe Registry. *Mol Genet Metab* **113**: 84-91.
105. Reuser, AJJ, van der Ploeg, AT, Chien, YH, Llerena, J, Jr., Abbott, MA, Clemens, PR*, et al.* (2019). GAA Variants and Phenotypes Among 1079 Patients with Pompe Disease: Data from the Pompe Registry. *Hum Mutat* **40**: 2146-2164.
106. Amalfitano, A, Bengur, AR, Morse, RP, Majure, JM, Case, LE, Veerling, DL*, et al.* (2001). Recombinant human acid alpha-glucosidase enzyme therapy for infantile glycogen storage disease type II: results of a phase I/II clinical trial. *Genet Med* **3**: 132-138.
107. Banugaria, SG, Patel, TT, and Kishnani, PS (2012). Immune modulation in Pompe disease treated with enzyme replacement therapy. *Expert Rev Clin Immunol* **8**: 497-499.
108. van Gelder, CM, Hoogeveen-Westerveld, M, Kroos, MA, Plug, I, van der Ploeg, AT, and Reuser, AJ (2014). Enzyme therapy and immune response in relation to CRIM status: the Dutch experience in classic infantile Pompe disease. *J Inherit Metab Dis **38**:305-314*.
109. Kishnani, PS, Goldenberg, PC, DeArmey, SL, Heller, J, Benjamin, D, Young, S*, et al.* (2010). Cross-reactive

immunologic material status affects treatment outcomes in Pompe disease infants. *Mol Genet Metab* **99**: 26-33.

110. Bali, DS, Goldstein, JL, Rehder, C, Kazi, ZB, Berrier, KL, Dai, J, *et al.* (2015). Clinical Laboratory Experience of Blood CRIM Testing in Infantile Pompe Disease. *Mol Genet Metab Rep* **5**: 76-79.
111. van der Ploeg, AT, Clemens, PR, Corzo, D, Escolar, DM, Florence, J, Groeneveld, GJ, *et al.* (2010). A randomized study of alglucosidase alfa in late-onset Pompe's disease. *N Engl J Med* **362**: 1396-1406.
112. de Vries, JM, van der Beek, NA, Hop, WC, Karstens, FP, Wokke, JH, de Visser, M, *et al.* (2012). Effect of enzyme therapy and prognostic factors in 69 adults with Pompe disease: an open-label single-center study. *Orphanet J Rare Dis* **7**: 73.
113. Toscano, A, and Schoser, B (2012). Enzyme replacement therapy in late-onset Pompe disease: a systematic literature review. *J Neurol* **260**: 951-9.
114. Barranger JA, C-SM (2007). *Lysosomal Storage Disorders*, Springer: New York.

Pompe disease

GAA variant database



CHAPTER 2

ENZYMATIC DIAGNOSIS OF POMPE DISEASE: LESSONS FROM 28 YEARS OF EXPERIENCE

Monica Y. Niño Martinez[1,2,3], Mark Wijgerde[2*], Douglas Oliveira Soares de Faria[1,2,3*] Atze J. Bergsma[1,2,3], Mike Broeders[1,2,3], Nadine A.M.E. van der Beek[4], Hannerieke J.M. van den Hout[1,3], Ans T. van der Ploeg[1,3], Frans W. Verheijen[1#] and W.W.M. Pim Pijnappel[1,2,3#]

[1]Department of Pediatrics, Erasmus University Medical Center, Rotterdam, Netherlands [2]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, Netherlands

[3]Center for Lysosomal and Metabolic Diseases, Erasmus University Medical Center, Rotterdam, Netherlands

[4]Department of Neurology, Erasmus University Medical Center, Rotterdam, The Netherlands

*These authors contributed equally to this work

[#]Correspondence should be addressed to W.W.M.P. Pijnappel, Tel: +31 1070437864 and F.W.V., Tel: +31 1070443212.

e-mail:
w.pijnappel@erasmusmc.nl
f.verheijen@erasmusmc.nl

Under review.

2

Enzymatic Diagnosis of Pompe Disease: lessons from 28 years of experience

Monica Y. Niño Martinez, Mark Wijgerde, Douglas Oliveira Soares de Faria, Atze J. Bergsma, Mike Broeders, Nadine A.M.E. van der Beek, Hannerieke J.M. van den Hout, Ans T. van der Ploeg, Frans W. Verheijen and W.W.M. Pim Pijnappel

ABSTRACT

Pompe disease is a lysosomal and neuromuscular disorder caused by deficiency of acid alpha-glucosidase (GAA), and causes classic infantile, childhood onset, or adulthood onset phenotypes. The biochemical diagnosis is based on GAA activity assays in dried blood spots, leukocytes, or fibroblasts. Diagnosis can be complicated by additional factors such as pseudodeficiencies, i.e. *GAA* variants that lower GAA activity but do not cause Pompe disease. A large-scale comparison between these assays for patient samples, including exceptions and borderline cases, along with clinical diagnoses has not been reported so far. Here we analyzed GAA activity in a total of 1709 diagnostic cases over the past 28 years using a total of 2591 analyses and we confirmed the clinical diagnoses. We used the following assays: leukocytes using glycogen or 4MUG as substrate, fibroblasts using 4MUG as substrate, and dried blood spots using 4MUG as substrate. In 794 individuals, 2 or more assays were performed. We found that phenotypes could only be distinguished using fibroblasts with 4MUG as substrate. This assay also had the lowest false positive and false negative rates. Pseudodeficiencies caused by the *GAA2* allele could be ruled out using 4MUG rather than glycogen as substrate in leukocytes or fibroblasts. The Asian pseudodeficiency could only be ruled out in fibroblasts using 4MUG as substrate. We conclude that fibroblasts using 4MUG as substrate provides the most reliable assay for biochemical diagnosis and can serve to confirm or reject results from leukocytes or dried blood spots.

Keywords: Pompe disease, acid α-glucosidase, enzymatic diagnosis, newborn screening.

INTRODUCTION

Pompe disease is an autosomal recessive metabolic disorder caused by acid α-glucosidase (GAA) deficiency, which leads to intralysosomal accumulation of glycogen causing progressive damage especially to cardiac and skeletal muscles [1-3]. Pompe disease presents as a spectrum of phenotypes. The classic infantile form is associated with rapidly progressive general muscle weakness and hypertrophic cardiomyopathy, culminating in death within the first year of life if left untreated. Less severe and less progressive forms can manifest at any age from infancy to late adulthood and present with proximal muscle weakness and/or respiratory problems with minimal or no cardiac involvement [4-7]. These patients eventually lose ambulation and become ventilator dependent. Enzyme replacement therapy with recombinant human GAA is available and is often applied in combination with immunomodulation in classic infantile cases to reduce the chance of antibody formation [8-11]. Early diagnosis including phenotype prediction is required to optimize counseling and to ensure a timely start of treatment.

Current diagnostic guidelines recommend the establishment of GAA enzyme deficiency with additional confirmation of two disease-associated *GAA* variants [12, 13]. Disease-associated variants, recently linked to clinical phenotypes, are listed in the open access Pompe disease *GAA* variant database [www.pompevariantdatabase.nl] [14, 15]. Different enzymatic diagnostic assays are available in which the biological material and the choice of substrate are variables.

Biological materials can be leukocytes, dried blood spots (DBSs), fibroblasts derived from skin biopsies, and muscle tissue. Leukocytes and DBSs are easily obtained, which is important for timely diagnosis and treatment of classic infantile patients in connection with newborn screening programs. However, DBSs assays are known to have both false positive and false negative outcomes [16] and therefore need to be confirmed with an independent assay. Substrates include the natural substrate glycogen, and the artificial substrates 4-methylumbelliferyl-α-D-glucopyranoside (4MUG) [17] and the solution (7-Benzoylamino-heptyl)-{2-[4-(3,4,5-trihydroxy-6-hydroxymethyl-tetrahydro-pyran-2-yloxy)-phenylcarbamoyl]-ethyl}-carbamic acid tert-butyl ester (GAA-S), a substrate used in tandem mass spectrometry for newborn screening [18]. In certain cell types such as leukocytes, the neutral hydrolase glucoamylase activity needs to be inhibited using acarbose as it interferes with the measurement of GAA activity [19-21].

For the interpretation of diagnostic outcome, it is important to realize that at least two (combinations of) *GAA* variants can lead to pseudodeficiency. In the Caucasian population, the *GAA2* (c.271G>A) pseudodeficiency variant lowers the activity of GAA for the natural substrate glycogen, but not for the artificial substrate 4MUG [22, 23]. In the Asian population, the common *GAA* pseudodeficiency variants c.[1726G>A; 2065G>A] can lower GAA activity to levels that come close to the disease threshold, and their presence can lead to a false positive diagnosis in certain assays [21, 24, 25].

Over the past 28 years, our laboratory has processed 1709 diagnostic cases and has diagnosed over 250 patients with Pompe disease using various assays separately and in parallel. In this article we are presenting and reviewing all our test results of blood-based and fibroblast-based assays from 1990 till 2018 allowing to compare the various methods for enzymatic diagnosis and their pitfalls.

MATERIALS AND METHODS

Diagnostic materials and assay procedures

Results were obtained over the period 1990-2018. Reference ranges were developed for each assay and are given for the different phenotypes (classic infantile, childhood, adulthood, healthy controls). Values in between these ranges are specified as 'gray zone'. For fibroblasts, the reference ranges were derived from data collected over a 40 years period. For the other assays, including GAA activity in leukocytes and bloodspots, and tetrasaccharide 6-α-D-glucopyranosyl-maltotriose (Glc4) concentration in urine by mass spectrometry [26], a period of at least 8-years was used. The study was conducted according to the Declaration of Helsinki. The Medical Ethical Committee at Erasmus University MC approved the study protocol, and all patients, or their parents or legal guardians, provided written informed consent.

Activity assays were performed as previously described [20, 26-29]. A final concentration of 3 µmol/L acarbose (in the reaction mixture) was used for the glycogen assay and 8 µmol/L acarbose for the 4MUG assay to inhibit glucoamylase activity present in mixed leukocytes and which interferes with GAA activity measurements. The GAA activity in DBS was determined in the presence of 8 µmol/L acarbose for glucoamylase inhibition and after precipitation of hemoglobin as previously described [27, 28].

Clinical diagnosis

The clinical information of all patients with Pompe disease and the diagnostic cases

with GAA enzyme activities above the patient range but below the control range (the 'gray zone') was verified to support or reject the clinical diagnosis.

RESULTS

Over the past 28 years, we analyzed a total of 1709 individuals using a total of 2591 assays (Figure 1). In the majority of individuals, leukocytes were used with both glycogen and 4MUG as substrate (n=637). In 539 individuals, fibroblasts using 4MUG as substrate were used as single assay, while in 375 individuals, leukocytes with glycogen were used as single assay. Only 1 individual was analyzed with leukocytes using 4MUG as substrate as single assay. 88 individuals were analyzed with all three assays, while 69 individuals were analyzed with both leukocytes/glycogen and fibroblasts/4MUG. As part of this study we traced back the clinical data of all patients with a confirmed enzymatic diagnosis. The following categories for the diagnoses were defined: 'Classic Infantile', 'Childhood onset' (symptom onset before the age of 18), 'Adult onset' (symptom onset at 18 years and older), 'no Pompe disease', and 'Unknown'. For enzymatic assays, the following additional categories were defined: 'Unknown/Deficient' (i.e. the value of the assay was deficient but it was unknown whether the individual had Pompe disease), 'No Pompe/Deficient' (i.e. the value of the assay was deficient but Pompe disease was excluded based on other assays), 'Asymptomatic/Deficient' (i.e. the value of the assay was deficient but the individual was asymptomatic), and 'Gray Zone' (the values above patient ranges but below normal ranges).

GAA activity assay in leukocytes using glycogen as substrate

Figure 2A shows GAA activities using glycogen as substrate that were measured in 1169 individuals. Ranges (in nmol/mg/hr), taken as guidelines by our diagnostic department, were as follows: 0-3.5, classic infantile; 0-10, childhood/adult onset; 10-40, gray zone; >40-250, normal, >250, above normal. Within the range of 0-3.5 nmol/mg/hr, there were 20 classic infantile patients, 15 with childhood onset, and 111 with adult onset Pompe disease. Two individuals were 'unknown deficient'. Within the range of 3.5-10 nmol/mg/hr, there were 1 classic infantile patient, 6 with childhood onset, 20 with adult onset, 7 'no Pompe/deficient', and 3 'unknown/deficient'. This indicated that classic infantile patients were strongly enriched in the range of 0-3.5 nmol/mg/hr, while childhood and adult onset patients were present in both 0-3.5 and 3.5-10 nmol/mg/hr ranges. In both ranges, putative false positives were found to a total of 12 individuals. A subset was followed up using leukocytes/4MUG and fibroblasts/4MUG assays, and these showed mainly gray zone (7assays) and normal

(9 assays) values, while one assay in leukocytes/4MUG showed an unknown/deficient value. Within the range of 10-40 nmol/mg/hr, 52 individuals were classified as gray zone, while 1 classic infantile patient was present. This patient showed values in the patient range in leukocytes/4MUG (3.46 nmol/mg/hr) and in the classic infantile range in fibroblasts/4MUG (0.2 nmol/mg/hr). In the normal ranges above 40 nmol/mg/hr, a total of 921 healthy individuals were present. In conclusion, the leukocytes/glycogen assay can partially distinguish between classic infantile and late onset phenotypes in the range between 3.5-10 nmol/mg/hr, has a considerable risk of false positives, and a low risk of false negatives.

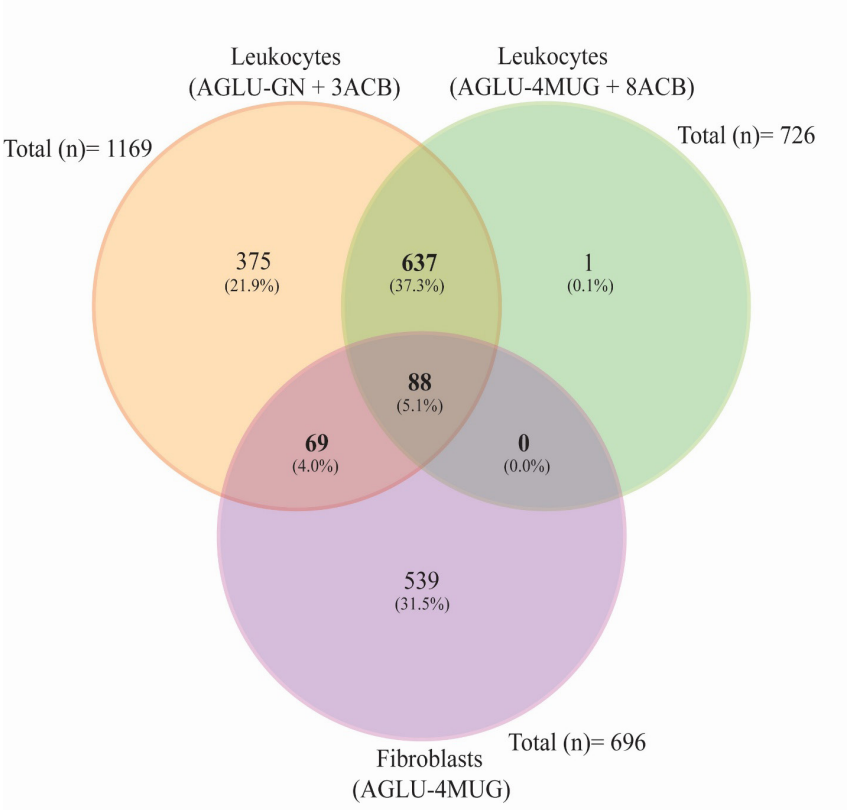


Figure 1. **Venn diagram showing how many diagnostic cases were processed with each of the 3 different enzymatic methods.** The figure shows the similarities and differences of individuals whereby the GAA activity was measured in leukocytes with glycogen as well as with 4MUG as substrate, and in fibroblast with 4MUG. Of note 11 assay results (5 from leukocytes with glycogen 4 from leukocytes with 4MUG, and 2 from fibroblast with 4MUG) were not included in this comparison since they are biological replicates. DBS assays (20 assays from 18 individuals) are not shown.

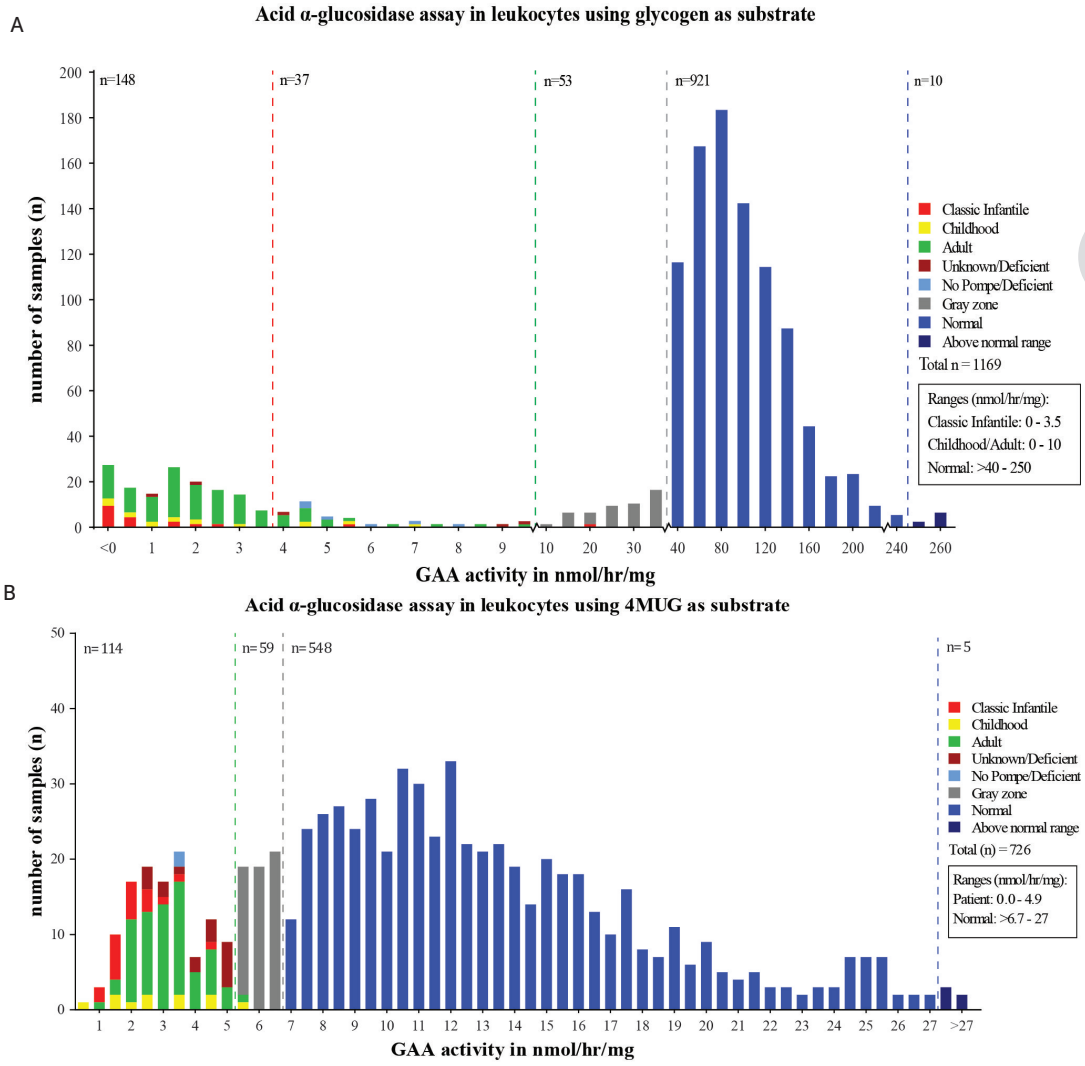


Figure 2. Distribution of GAA activity in leukocytes using: A) Glycogen as substrate in the presence of 3 µmol/L acarbose (AGLU-GN+3ACB assay). A total of 1169 individuals suspected of Pompe disease were included. B) 4MUG as substrate in the presence of 8 µmol/L acarbose (AGLU-4MUG+8ACB assay). A total of 726 individuals suspected of Pompe disease were included in this series.

GAA activity assay in leukocytes using 4MUG as substrate

In Figure 2B, the activities in leukocyte assays using 4MUG as substrate, measured in 726 individuals, are shown. The ranges (in nmol/hr/mg) in leukocytes/4MUG, taken as guidelines by our diagnostic department, were as follows: 0.0-4.9, patients; 4.9-6.7, gray zone, >6.7-27, normal, >27, above normal. In the range 0.0-4.9, there were 19 classic infantile patients, 10 with childhood and 66 with adulthood onset, 2 were

no Pompe/deficient, and 17 were unknown/deficient. The 2 individuals that were no Pompe/deficient and the 17 individuals that were unknown/deficient were either gray zone and/or normal when tested in leukocytes/4MUG and/or fibroblasts/4MUG. This showed that the leukocyte/4MUG assay did not distinguish classic infantile from late onset phenotypes and that it showed relatively frequent false positive outcomes. The range 4.9-6.7 showed 1 childhood onset patient, 1 adult onset patient, and 57 gray zone individuals. The childhood onset patient had borderline activity just above the threshold (5.14 nmol/mg/hr) and tested in the patient range (0.365 nmol/mg/hr) in the leukocyte/glycogen assay. The range > 6.7 nmol/mg/hr showed 548 normal and 5 above normal individuals. This indicated that the leukocyte/4MUG had a low false negative rate, especially when the healthy range (>6.7 nmol/mg/hr) was applied.

GAA activity assay in fibroblasts with 4MUG as substrate

In Figure 3, the activities in fibroblasts using 4MUG as substrate are shown. Ranges (in nmol/mg/hr), taken as guidelines by our diagnostic department, were as follows: 0-3, classic infantile; 4.2-20, childhood/adulthood onset; 20-45, gray zone; >45-180, normal; >180: above normal. In the range 0-3, there were 77 classic infantile patients, 8 patients with childhood and 3 with adulthood onset. One patient with childhood onset had an activity in between 3-4.2. In the range 4.2-20, there were 1 classic infantile patient, 18 with childhood onset and 108 with adulthood onset. There were 6 asymptomatic/deficient individuals. These individuals may develop symptoms later in life, as symptom onset in Pompe disease can be highly variable. Two of these individuals were also tested using leukocytes/glycogen and 4MUG, and in both individuals, these two assays yielded gray zone values. This indicated that the fibroblasts/4MUG assay could distinguish classic infantile from childhood/adulthood onset patients to a large extent, and between childhood and adulthood patients to a lesser extent. In the range >20-45 nmol/hr/mg, there were 61 gray zone individuals, while all individuals with values >45 nmol/hr/mg were normal indicating low false positive and false negative outcomes of this assay.

GAA activity assay in DBS

Table 1 shows the results of dried blood spot (DBS) assays in 18 individuals with Pompe disease. Ranges (in pmol/17h/ml) were 11-56, patients; 56-94, gray zone; 94-448, normal. In 15 cases (12 classic infantile, 2 childhood, 1 with an unknown phenotype), GAA activity was in the patient range, while three individuals (patients 30-32, who had classic infantile Pompe disease) were in the gray zone. However, all individuals including individuals 30-32 tested in the patient range in leukocytes with glycogen as substrate. One sample, from patient 32 was measured three times, of which one was in the patient range and 2 in the gray zone. Fibroblasts from this patient showed activity in the patient

range and DNA analysis of this patient and patients 30 and 31 confirmed the presence of two severe *GAA* disease-associated variants (Table 2 and Suppl. Fig S1). These results indicated that the DBS had a relatively high false negative rate (16.7 % in this analysis).

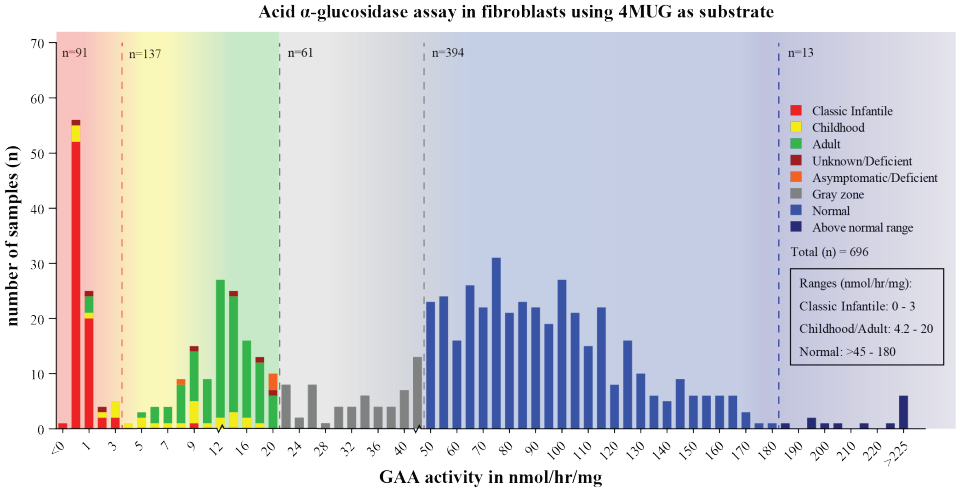


Figure 3. Distribution of GAA activity in fibroblast using the artificial substrate 4-methylumbellife-ryl-α-D-glucoside (4MUG) (AGLU-4MUGassay). A total of 695 individuals suspected of Pompe disease were included in this series.

Table 1. GAA activity in Pompe patients measured in DBS and leukocytes

Patient	DRIED BLOOD SPOT (DBS)		LEUCOCYTES		Clinical diagnosis
	AGLU-4MUG+8ACB (pmol/17h/ml)	BGAL (pmol/17h/ml)	AGLU-GN+3ACB (nmol/hr/mg)	BGAL (nmol/hr/mg)	
15	30.8	4730	1.6	152	Classic infantile Pompe disease
16	48.8	2480	1.0	131	Childhood onset Pompe disease
17	13.8	2860	-0.1	162	Classic infantile Pompe disease
18	25.2	2520	0.8	129	Classic infantile Pompe disease
19	17.4	1660	5.0	147	Classic infantile Pompe disease
20	11.4	1380	1.2	192	Classic infantile Pompe disease
21	10.6	1780	-1.1	174	Classic infantile Pompe disease
22	27.8	3100	-0.2	192	Classic infantile Pompe disease
23	14.6	2040	-1.6	167	Classic infantile Pompe disease
24	12.9	2950	0.3	207	Classic infantile Pompe disease
25	25.4	3920	0.1	172	Classic infantile Pompe disease
26	28.6	3710	2.8	104	Childhood onset Pompe disease
27	25	4990	0.9	222	Unknown
28	8.51	3060	1.2	161	Classic infantile Pompe disease
29	41.5	4700	-0.4	181	Classic infantile Pompe disease
30	59.4	3620	-2.0	225	Classic infantile Pompe disease

Table 1 continued.

Patient	DRIED BLOOD SPOT (DBS)		LEUCOCYTES		Clinical diagnosis
	AGLU-4MUG+8ACB	BGAL	AGLU-GN+3ACB	BGAL	
	(pmol/17h/ml)	(pmol/17h/ml)	(nmol/hr/mg)	(nmol/hr/mg)	
31	58.6	6050	-1.1	169	Classic infantile Pompe disease
	85.3	4710			Classic infantile Pompe disease
32*	68	5230	-0.7	139	Classic infantile Pompe disease
	48.6	3250			Classic infantile Pompe disease
Normal range	**94 - 448**	**476 - 4680**	**>40 - 250**	**50 - 326**	
Patient range	**11 - 56**	**88**	**Classic Infantile: 0 - 3.5 Childhood/Adult: 0 - 10**	**0.6 - 6.3**	

Cases 30-32 were borderline inconclusive using the DBS, but were positively confirmed with leukocytes as sample source. Patient 32 showed variable measurements of GAA activity and BGAL in DBS (technical replicates were performed). Patient 16 and 26 were diagnosed with childhood Pompe disease, and individual 27 was classified as Unknown/Deficient due lack of clinical information. Red: values within the patient range.

Table 2. Follow up of three cases with inconclusive DBS data.

Patient	Genotype DNA (protein)	Predicted severity (Pompe disease *GAA* variant database)	ACMG variant classification	Fibroblast (GAA activity) AGLU-4MUG (nmol/hr/mg)	Clinical diagnosis
30	c.1460T>C p.(Phe487Ser);	Potentially less severe	Likely pathogenic	0.432	Classic infantile Pompe disease
	c.1460T>C p.(Phe487Ser)	Potentially less severe	Likely pathogenic		
31	c.379_380del p.(Cys127Leufs*18);	Very severe	Pathogenic	0.35	Classic infantile Pompe disease
	c.525del p.(Glu176Argfs*45)	Very severe	Pathogenic		
32	c.2481+102_2646+31del p.(Gly828_Asn882del);	Very severe	Pathogenic	0.415	Classic infantile Pompe disease
	c.525del p.(Glu176Argfs*45)	Very severe	Pathogenic		
Ranges (GAA activity)			Normal Range	>45 - 180	
			Classic Infantile	0 - 3	
			Childhood/Adult	4.2 - 20	

DNA analysis and GAA activity in cultured fibroblasts left no doubt that all three patients had classic infantile Pompe disease.

GAA2 pseudodeficiency

Individuals with pseudodeficiencies can give false positive outcomes in GAA enzyme assays. To illustrate this, we compared individuals with confirmed presence of the *GAA2* (c.271G>A) and Asian c.[1726G>A; 2065G>A] pseudodeficiencies using different GAA activity assays.

In 4 individuals with the *GAA2* variant at heterozygous state without other known GAA

disease-associated variants, activities were in the gray zone (2 cases) or in the normal range (2 cases) in leukocytes/glycogen (Table 3). In leukocytes/4MUG, 1 case had activity in the gray zone, and 3 in the normal range. In fibroblasts/4MUG, 3 patients were tested, 2 of whom were in the normal range and one (individual 4) in the patient range. This indicated that the *GAA2* allele at heterozygous state can already lower GAA activity in leukocyte-based assays to values in the gray zone. Individual 4 had a medical record of hypertrophic cardiomyopathy. Follow up using genetic testing led to the identification of a disease-associated variant in exon 19 (c.1831G>A, p.E611K) of the MYBPC1 gene (Myosin Binding Protein C, Slow Type), which plays an important role in muscle contraction and cardiac conduction. We suspect that individual 4 may contain a yet unidentified 2nd *GAA* disease-associated variant.

Table 3. GAA enzyme activity in leukocytes and fibroblast from individuals carrying the *GAA2* pseudodeficiency allele (c.271G>A p.(Asp91Asn))

Individual/ Patient (Gender)	Genotype DNA (protein)	Enzymatic Assay			Clinical diagnosis
		LEUKOCYTES		FIBROBLAST	
		AGLU-GN+3ACB (nmol/hr/mg)	AGLU-4MUG+8ACB (nmol/hr/mg)	AGLU-4MUG (nmol/hr/mg)	
1 (M)	c.271G>A p.(Asp91Asn)	36.2	7.9	61.7	No Pompe disease
2 (M)	c.271G>A p.(Asp91Asn)	46.6	8.5	50.4	No Pompe disease
3 (M)	c.271G>A p.(Asp91Asn)	77.6	12.2	N.D.	No Pompe disease
4 (M)[1]	c.271G>A p.(Asp91Asn)	24.0	5.7	19.5	No Pompe disease
5 (M)	c.271G>A p.(Asp91Asn); c.271G>A p.(Asp91Asn)	7.6	10.9	49.8	No Pompe disease
6 (M)*	c.271G>A p.(Asp91Asn); c.271G>A p.(Asp91Asn)	4.4/7.7	11.1	86.9	No Pompe disease
7 (F)	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0)	2.2	6.1	25.9	No Pompe disease
8 (M)	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0)	4.1	N.D.	51.0	No Pompe disease
9 (F)	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0)	4.7	6.5	53.5	No Pompe disease
10 (F)	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0)	6.6	N.D.	53.1	No Pompe disease
11 (F)[2]	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0); c.1076-22T>G p.?	5.7	N.D.	33.3	No Pompe disease
12 (F)	c.271G>A p.(Asp91Asn); c.-32-13T>G p.(=), p.(0); c.1447G>A p.(Gly483Arg)	1.5	1.9	7.6	Adult onset Pompe disease
Ranges	**Normal range**	>40 - 250	>6.7 - 27	>45 - 180	
	Patient range	Classic Infantile: 0 - 3.5 Childhood/Adult: 0 - 10	1.1 - 4.9	Classic Infantile: 0 - 3 Childhood/ Adult: 4.2 - 20	

[1]Unaffected by Pompe disease, but a disease-associated variant was found in the MYBPC1 gene, which is associated with cardiac hypertrophy.
[2]Unaffected by Pompe disease. * The biological replicate was 7.7 nmol/hr/mg in leukocytes using glycogen as substrate. Red: values within the patient range.

In individuals that were homozygous for *GAA2* (number 5 and 6), GAA activities were in the patient range in leukocytes/glycogen, but in the normal range in leukocytes/4MUG and in fibroblasts/4MUG (Table 3). In individuals that were compound heterozygous for *GAA2* in combination with a second disease-associated *GAA* variant (numbers 7-11), GAA activities were in the patient range in leukocytes/glycogen in all cases (individual 11 contained an additional *GAA* variant classified as childhood when combined with a null allele according to www.pompevariantdatabase.nl) (Table 3). In leukocytes/4MUG and fibroblasts/4MUG, values were either in the gray zone or in the normal range. This highlights that the *GAA2* pseudodeficiency at homozygous or compound heterozygous state can give false positive outcomes when using the leukocyte/glycogen assay, while correct outcomes can be obtained by using either the leukocytes/4MUG or fibroblast/4MUG assays. Patient 12 contained 2 *GAA* disease-associated variants in addition to the *GAA2* allele at heterozygous state, and showed GAA values in the patient range in all three assays consistent with the diagnosis of Pompe disease (adulthood onset).

Asian c.[1726G>A; 2065G>A] pseudodeficiency

Individual 13 was compound heterozygous and individual 14 homozygous for the Asian c.[1726G>A; 2065G>A] pseudodeficiency (Table 4). In both cases, assays using leukocytes, either with glycogen or with 4MUG as substrate, resulted in activities in the patient range, while activities in fibroblasts/4MUG were in the gray zone, just above the patient range. Additional diagnostic analyses showed that individual 13 had normal CK, ASAT, and ALAT levels, but slightly elevated TGLC levels. Individual 14 did have elevated CK, ASAT, and ALT levels but normal TGLC levels. Both individuals were diagnosed not to have Pompe disease based on the diagnostic information and the lack of clinical signs associated with Pompe disease. This indicated that the presence of the Asian pseudodeficiency can seriously affect the diagnostic outcome of enzymatic assays that are based on leukocytes, independent of the substrate used.

Table 4. Lack of diagnosis in cases with the Asian pseudodeficiency (*c.1726G>A and c.2065G>A)*

Individual (Gender, age in years)	Genotype DNA (protein)	Enzymatic assay			Ancillary Studies				Clinical diagnosis
		LEUKOCYTES		FIBROBLAST	Plasma			Urine	
		AGLU-GN+3ACB	AGLU-4MUG+8ACB	AGLU-4MUG	CK	ASAT	ALAT	TGLC	
		(nmol/hr/mg)	(nmol/hr/mg)	(nmol/hr/mg)	(U/l)	(U/l)	(U/l)	(mmol/mol creatinine)	
13 (F, 74 y)	c.1726G>A p.(Gly576Ser); c.2065G>A p.(Glu689Lys); c.-32-13T>G p.(=), p.(0)	9.6	3	22	58-104	27-34	19-32	2.3*	No Pompe disease
14 (M, 53 y)	c.1726G>A p.(Gly576Ser); c.2065G>A p.(Glu689Lys); c.1726G>A p.(Gly576Ser); c.2065G>A p.(Glu689Lys)	18.7	3.9	20.75	992	50	57	0.8	No Pompe disease
Ranges	**Normal range**	>40- 250	>6.7- 27	>45- 180	W>17 years: <170 M>17 years: <200	W>17 years: <31 M>17 years: <37	W>17 years: <31 M>17 years: <41	> 20 years: 0-2.2	
	Patient range	Classic Infantile: 0 - 3.5 Childhood/Adult: 0 - 10	1.1- 4.9	Classic Infantile: 0 - 3 Childhood/ Adult: 4.2- 20				> 20 years: 2.3-130	

GAA activities were measured both in leukocytes using the natural and artificial substrates as well as in cultured fibroblasts. *2th measurement of TGLC was 5.1 and a 3rd measurement was 11.6 mmol/mol creatinine. Red: values within the patient range.

DISCUSSION

We compared the outcome of different assay procedures for measuring the GAA activity in various sample types using different substrates measured at the Erasmus MC in the last 28 years. The pro's and con's of the different diagnostic procedures are discussed below.

Blood based assays

Leukocytes

We used leukocyte pellets for diagnostic purposes and applied both glycogen as well as 4MUG as substrate [27]. The results presented in Figure 2A and 2B illustrate that a diagnosis can be established in most cases and with both substrates if proper cutoff values are chosen. As previously suggested by van Diggelen et al [27], our long-term data also demonstrate that the dynamic range is broader with glycogen than with 4MUG. None of the two assays fully discriminate between classic and milder (childhood, adult) phenotypes, but by using both assays in parallel, pseudodeficiencies caused by the Caucasian *GAA2* allele can be excluded (Figure 4B). Another pseudodeficiency that complicates the assay is the Asian c.[1726G>A; 2065G>A] pseudodeficiency. This problem applies to all enzymatic procedures used.

Bloodspots

In our patient cohort, 3 (16.7%) of 18 patients (assays 30-32 in Suppl. Figure S1) came out as false negative in the DBS assay but were correctly diagnosed with other diagnostic methods. Analysis of the activity of a second lysosomal enzyme (for instance BGAL) as reference enzyme may help to judge the outcome of the assay. For example, a normal GAA activity should not be trusted if the reference enzyme shows activity outside of the normal ranges. In such an event additional assays are required.

Despite the ten times higher activity of GAA for glycogen compared to 4MUG, the small sample size of the bloodspot assay precludes the use of glycogen for DBS testing due to the lower sensitivity of the colorimetric assay used with glycogen compared to the fluorimetric detection using 4MUG. The tandem mass spectrometry methodology, employing GAA-S has proven to be a valuable alternative [18]. A recent refinement of that method was reported to separate 96% of the Taiwanese newborns with GAA pseudodeficiency and all Pompe disease carriers from patients with Pompe disease [30]. A recent study from Japan supports this claim [31].

Skin fibroblast assays

Figures 2 (A, B) and 3 clearly demonstrate that the combination of fibroblasts as diagnostic material and the fluorimetric assay using 4MUG as artificial substrate provide the most robust and reliable enzymatic assay for diagnosing Pompe disease. In the majority of cases it distinguishes classic infantile Pompe disease from non-classic phenotypes on the basis of residual activity, while none of the other methods do. Assays with intermediate GAA activity were likely derived from Pompe disease carriers. DNA analysis usually provides a clear answer in these cases, but some remained unsolved. For instance, in individual 13, that carried the Asian pseudodeficiency c.[1726G>A; 2065G>A] in combination with the IVS1 variant, biochemical results suggested Pompe disease while clinical signs did not, but it leaves doubt whether adult/late onset Pompe disease can be fully ruled out in this patient.

Conclusions and recommendations

We conclude that cultured skin fibroblasts provide the best sample source. Leukocytes isolated from peripheral blood are the second-best diagnostic sample source, but the reaction mixture needs to contain acarbose to inhibit glucoamylase. The use of glycogen as natural substrate enhances the resolution between affected and unaffected, but *GAA2* pseudodeficiency occurs in the Caucasian population and has to be excluded by also using 4MUG as substrate in case of exceptionally low GAA activity (Figure 4B). The implementation of DBSs has enabled (newborn) screening, and several analytical methods have proven their value as first tier test. There is general agreement that additional tests remain necessary to finally establish the diagnosis. For reasons mentioned in the text, we advise to conduct genetic testing in addition to any type of enzymatic testing, as recommended recently by the EPOC consortium [32].

The recommended diagnostic flow is as follows: 1) start with leukocytes using 4MUG and glycogen as substrate; 2) confirm with DNA analysis, in which two disease-associated variants should be present, while a pseudodeficiency does not qualify as disease-associated variant; 3) if there is doubt, for example when values are close to the gray zone or when a DNA variant is unknown, use fibroblasts as confirmation (Figure 4A, B).

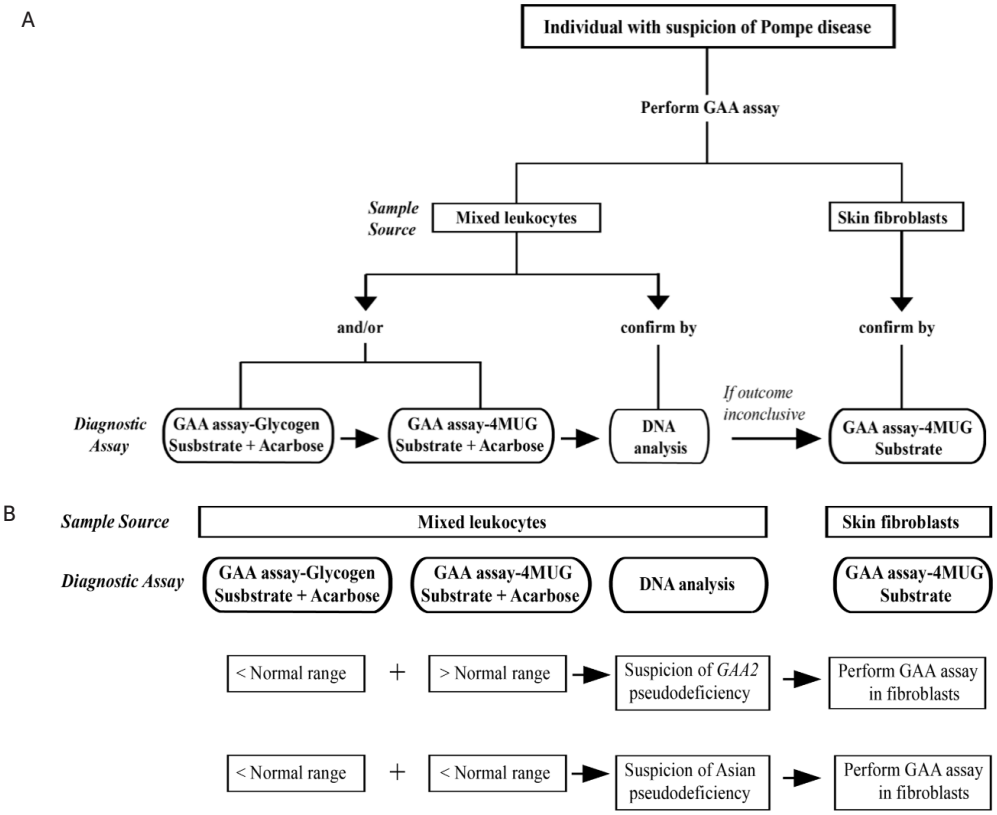


Figure 4. Flow chart for diagnosis of Pompe disease. A) General flow chart. B) Flow chart for cases with pseudodeficiency.

Acknowledgements

We thank Dr. Arnold Reuser for invaluable discussions.

Conflict of Interest

Ans T. van der Ploeg has provided consulting services for various industries in the field of Pompe disease under an agreement between these industries and Erasmus MC, Rotterdam, The Netherlands. All other authors do not declare any conflict of interest.

Funding Source

This work was supported by Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias), Colombia (Administrative Department of Science, Technology and Innovation, Colciencias, Colombia), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil), Zeldzame Ziekten Fonds/

WE Foundation, the Sophia Foundation for Medical Research (SSWO; project number s17-32), Metakids (project number 2016 - 063 and 2018-082), and Stofwisselkracht.

REFERENCES

1. Hers, HG (1963). alpha-Glucosidase deficiency in generalized glycogen storage disease (Pompe's disease). *Biochem J* **86**: 11-16.
2. Dasouki, M, Jawdat, O, Almadhoun, O, Pasnoor, M, McVey, AL, Abuzinadah, A, *et al.* (2014). Pompe Disease: Literature Review and Case Series. *Neurol Clin* **32**: 751-776.
3. Reuser, AJJ, Hirschhorn, R and Kroos, MA (2018). Pompe Disease: Glycogen Storage Disease Type II, Acid α-Glucosidase (Acid Maltase) Deficiency. In: Scriver, CR, AL Beaudet, D Valle and WS Sly (eds). *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Medical: NY.
4. van der Ploeg, AT, and Reuser, AJ (2008). Pompe's disease. *Lancet* **372**: 1342-1353.
5. Herzog, A, Hartung, R, Reuser, AJ, Hermanns, P, Runz, H, Karabul, N, *et al.* (2012). A cross-sectional single-centre study on the spectrum of Pompe disease, German patients: molecular analysis of the GAA gene, manifestation and genotype-phenotype correlations. *Orphanet J Rare Dis* **7**: 35.
6. Laforet, P, Laloui, K, Granger, B, Hamroun, D, Taouagh, N, Hogrel, JY, *et al.* (2013). The French Pompe registry. Baseline characteristics of a cohort of 126 patients with adult Pompe disease. *Rev Neurol (Paris)* ***169****: 595-602*.
7. van Capelle, CI, van der Meijden, JC, van den Hout, JM, Jaeken, J, Baethmann, M, Voit, T, *et al.* (2016). Childhood Pompe disease: clinical spectrum and genotype in 31 patients. *Orphanet J Rare Dis* **11**: 65.
8. Poelman, E, Hoogeveen-Westerveld, M, Kroos-de Haan, MA, van den Hout, JMP, Bronsema, KJ, van de Merbel, NC, *et al.* (2018). High Sustained Antibody Titers in Patients with Classic Infantile Pompe Disease Following Immunomodulation at Start of Enzyme Replacement Therapy. *J Pediatr* **195**: 236-243 e233.
9. Stenger, EO, Kazi, Z, Lisi, E, Gambello, MJ, and Kishnani, P (2015). Immune Tolerance Strategies in Siblings with Infantile Pompe Disease-Advantages for a Preemptive Approach to High-Sustained Antibody Titers. *Mol Genet Metab Rep* **4**: 30-34.
10. Elder, ME, Nayak, S, Collins, SW, Lawson, LA, Kelley, JS, Herzog, RW, *et al.* (2013). B-Cell depletion and immunomodulation before initiation of enzyme replacement therapy blocks the immune response to acid alpha-glucosidase in infantile-onset Pompe disease. *J Pediatr* **163**: 847-854 e841.
11. Poelman, E, Hoogeveen-Westerveld, M, van den Hout, JMP, Bredius, RGM, Lankester, AC, Driessen, GJA, *et al.* (2019). Effects of immunomodulation in classic infantile Pompe patients with high antibody titers. *Orphanet J Rare Dis* **14**: 71.
12. Musumeci, O, and Toscano, A (2019). Diagnostic tools in late onset Pompe disease (LOPD). *Ann Transl Med* **7**: 286.
13. van der Ploeg, AT, Kruijshaar, ME, Toscano, A, Laforet, P, Angelini, C, Lachmann, RH, *et al.* (2017). European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. *Eur J Neurol* **24**: 768-e731.
14. Kroos, M, Hoogeveen-Westerveld, M, Michelakakis, H, Pomponio, R, Van der Ploeg, A, Halley, D, *et al.* (2012). Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. *Hum Mutat* **33**: 1161-1165.

15. Niño, MY, In 't Groen, SLM, Bergsma, AJ, van der Beek, N, Kroos, M, Hoogeveen-Westerveld, M, *et al.* (2019). Extension of the Pompe mutation database by linking disease-associated variants to clinical severity. *Hum Mutat* ***40****: 1954-1967*.
16. Bodamer, OA, Scott, CR, Giugliani, R, and Pompe Disease Newborn Screening Working, G (2017). Newborn Screening for Pompe Disease. *Pediatrics* **140**: S4-S13.
17. Robinson, D, Price, RG, and Dance, N (1967). Rat-urine glycosidases and kidney damage. *Biochem J* **102**: 533-538.
18. Li, Y, Scott, CR, Chamoles, NA, Ghavami, A, Pinto, BM, Turecek, F, *et al.* (2004). Direct Multiplex Assay of Lysosomal Enzymes in Dried Blood Spots for Newborn Screening. *Clin Chem* **50**: 1785-1796.
19. Gelb, MH, Turecek, F, Scott, CR, and Chamoles, NA (2006). Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for the newborn screening of lysosomal storage disorders. *J Inherit Metab Dis* **29**: 397-404.
20. Okumiya, T, Keulemans, JL, Kroos, MA, Van der Beek, NM, Boer, MA, Takeuchi, H, *et al.* (2006). A new diagnostic assay for glycogen storage disease type II in mixed leukocytes. *Mol Genet Metab* **88**: 22-28.
21. Shigeto, S, Katafuchi, T, Okada, Y, Nakamura, K, Endo, F, Okuyama, T, *et al.* (2011). Improved assay for differential diagnosis between Pompe disease and acid alpha-glucosidase pseudodeficiency on dried blood spots. *Mol Genet Metab* **103**: 12-17.
22. Martiniuk, F, Bodkin, M, Tzall, S, and Hirschhorn, R (1990). Identification of the base-pair substitution responsible for a human acid alpha glucosidase allele with lower "affinity" for glycogen (GAA 2) and transient gene expression in deficient cells. *Am J Hum Genet* **47**: 440-445.
23. Swallow, DM, Kroos, M, Van der Ploeg, AT, Griffiths, B, Islam, I, Marenah, CB, *et al.* (1989). An investigation of the properties and possible clinical significance of the lysosomal alpha-glucosidase GAA*2 allele. *Ann Hum Genet* **53**: 177-184.
24. Kumamoto, S, Katafuchi, T, Nakamura, K, Endo, F, Oda, E, Okuyama, T, *et al.* (2009). High frequency of acid alpha-glucosidase pseudodeficiency complicates newborn screening for glycogen storage disease type II in the Japanese population. *Mol Genet Metab* **97**: 190-195.
25. Tajima, Y, Matsuzawa, F, Aikawa, S, Okumiya, T, Yoshimizu, M, Tsukimura, T, *et al.* (2007). Structural and biochemical studies on Pompe disease and a "pseudodeficiency of acid alpha-glucosidase". *J Hum Genet* **52**: 898-906.
26. Sluiter, W, van den Bosch, JC, Goudriaan, DA, van Gelder, CM, de Vries, JM, Huijmans, JG, *et al.* (2012). Rapid Ultraperformance Liquid Chromatography-Tandem Mass Spectrometry Assay for a Characteristic Glycogen-Derived Tetrasaccharide in Pompe Disease and Other Glycogen Storage Diseases. *Clin Chem* ***58****: 1139-114*.
27. van Diggelen, OP, Oemardien, LF, van der Beek, NA, Kroos, MA, Wind, HK, Voznyi, YV, *et al.* (2009). Enzyme analysis for Pompe disease in leukocytes; superior results with natural substrate compared with artificial substrates. *J Inherit Metab Dis* **32**: 416-423.
28. Oemardien, LF, Boer, AM, Ruijter, GJ, van der Ploeg, AT, de Klerk, JB, Reuser, AJ, *et al.* (2011). Hemoglobin precipitation greatly improves 4-methylumbelliferone-based diagnostic assays for lysosomal storage

diseases in dried blood spots. *Mol Genet Metab* **102**: 44-48.

29. Smith, PK, Krohn, RI, Hermanson, GT, Mallia, AK, Gartner, FH, Provenzano, MD, *et al.* (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**: 76-85.
30. Liao, HC, Chan, MJ, Yang, CF, Chiang, CC, Niu, DM, Huang, CK, *et al.* (2017). Mass Spectrometry but Not Fluorimetry Distinguishes Affected and Pseudodeficiency Patients in Newborn Screening for Pompe Disease. *Clin Chem* **63**: 1271-1277.
31. Mashima, R, and Okuyama, T (2017). Enzyme activities of alpha-glucosidase in Japanese neonates with pseudodeficiency alleles. *Mol Genet Metab Rep* **12**: 110-114.
32. Schoser, B, Laforet, P, Kruijshaar, ME, Toscano, A, van Doorn, PA, van der Ploeg, AT, *et al.* (2015). 208th ENMC International Workshop: Formation of a European Network to develop a European data sharing model and treatment guidelines for Pompe disease Naarden, The Netherlands, 26-28 September 2014. *Neuromuscul Disord* **25**: 674-678.

SUPPLEMENTARY TEXT

Pseudodeficiencies and borderline cases

c.271G>A (GAA2): common Caucasian background

Table 3 lists individuals that contained the *GAA2* pseudodeficiency allele. In individuals 1-4, only the *GAA2* variant was detected in heterozygous state but no other disease-associated *GAA* variant were detected. GAA enzyme activities were generally low and in some cases in the gray zone, but above the patient range in all three assays (leukocytes using 4MUG or glycogen as substrate and fibroblasts using 4MUG as substrate), except for individual 4. This person showed an activity that was within the patient range in fibroblasts using 4MUG a substrate (19.5 nmol/mg/hr, patient range 4.2-20), while the activity was slightly above the disease threshold in leukocytes using 4MUG as substrate. This individual was diagnosed not to have Pompe disease. The genetic testing led to the identification of a disease-associated variant in exon 19 (c.1831G>A, p.E611K) of the MYBPC1 gene (Myosin Binding Protein C, Slow Type). The GAA enzymatic diagnostic outcome remained enigmatic.

Individual 5 and 6 were homozygous for *GAA2*. In both cases, the activity for glycogen was within the patients' range, but for 4MUG in the normal range. Also based on the normal activity in cultured fibroblasts, both cases were diagnosed not to have Pompe disease.

Individual 7-10 were compound heterozygous for *GAA2* and c.-32-13T>G (IVS1). Also in these cases, the activities using glycogen as substrate were in the patient range, but activities using 4MUG as substrate were in the normal range or gray zone.

Individuals 11 and 12 contained the *GAA2* allele in heterozygous state in addition to two disease-associated variants. Case 11 concerned an unaffected of six children, two females of whom have the adult Pompe disease phenotype. DNA analysis of the mother revealed the presence of two disease-associated variants on one *GAA* allele c.[-32-13T>G; 1076-22T>G] and the *GAA2* variant on the other. The activity in leukocytes was in the patient range when measured with glycogen as substrate, but in the gray zone (33.3 nmol MU/hr/mg) when measured in fibroblasts with 4MUG as substrate. She was diagnosed not to have Pompe disease. Case 12 demonstrated GAA deficiency in all three assays and was diagnosed with adult onset of Pompe disease. She contained [c.271G>A; c.-32-13T>G on one allele, and the disease-associated variant c.1447G>A p.(Gly483Arg) on the other. This was supported by her clinical history, mentioning unbalanced walking at the age of 26 years, progressing to difficulty with running,

respiratory involvement at the age of 43, and weak back and walking difficulties one decade later. She developed walking disability and started ERT at the age of 64 years.

In conclusion, the pseudodeficiency *GAA2* allele can lower GAA enzyme activity measurements when using glycogen as substrate, but not when using 4MUG as substrate. This can result in values within the patient range when using leukocytes with glycogen as substrate in patients that carry a disease-associated variant on one allele and *GAA2* on the second allele.

c.1726G>A and c.2065G>A: common Asian background

Individual 13 was referred at the age of 73 and was compound heterozygous for c.[1726G>A, c.2065G>A] and c.-32-13T>G (IVS1) (Table 4). Activities in leukocytes were in the patient range, both when using glycogen and 4MUG as substrate. The activity in fibroblasts was just above the patient range in the gray zone. This individual also had elevated urinary glucose tetrasaccharide levels, but no pathology in a muscle biopsy. Individual 14 was referred at the age of 60 and was homozygous for c.[1726G>A, c.2065G>A]. GAA enzyme activity was in the patient range in leukocytes with 4MUG as substrate but not with glycogen as substrate; in fibroblasts the activity was just above the patient range. This case also had elevated CK, ASAT, and ALAT levels. Both cases did not display clinical signs of Pompe disease and were diagnosed not to have Pompe disease. They illustrate the difficulty of GAA diagnosis solely based on biochemical assays.

Genetic analysis

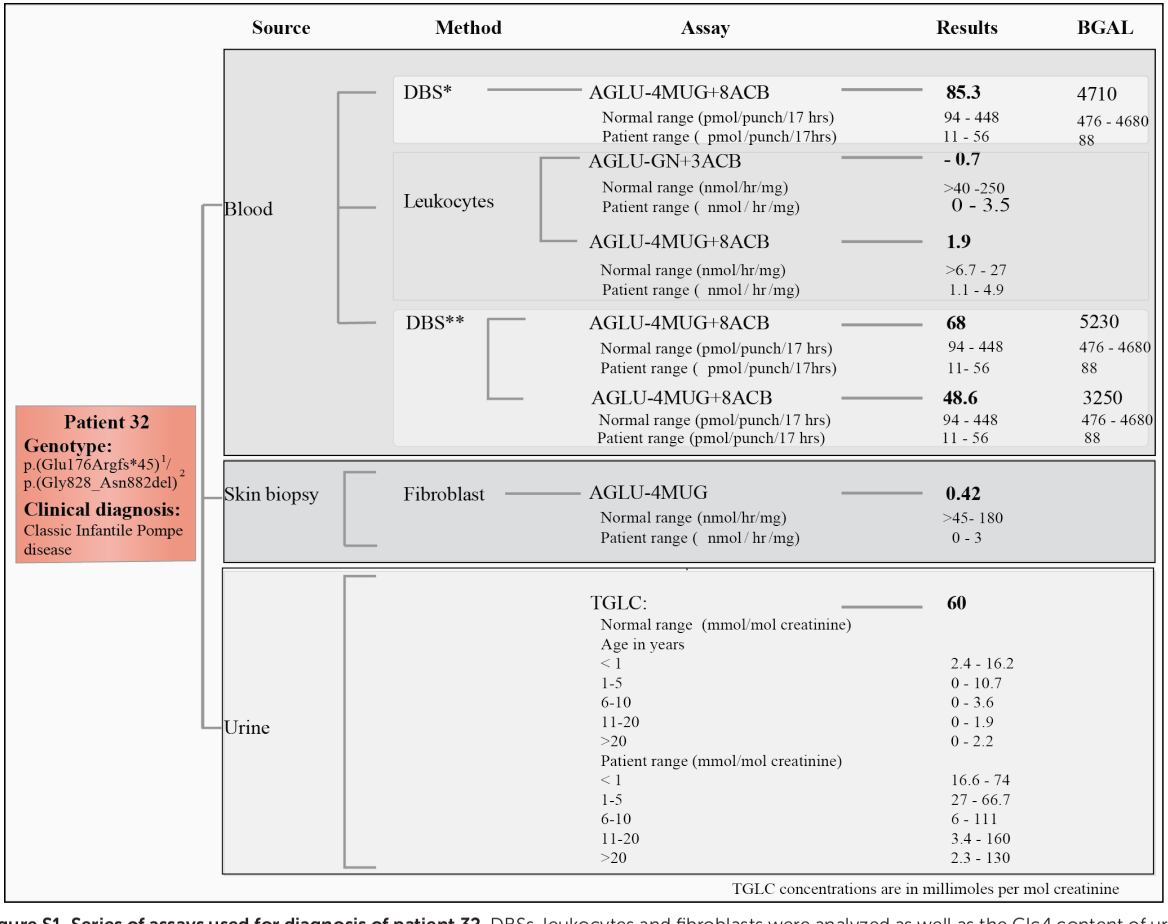
DNA sequence analysis has become an alternative or supportive diagnostic tool in Pompe disease. We recently extended the Pompe mutation database, now termed the Pompe disease *GAA* variant database, with clinical information [www.pompevariantdatabase.nl] [1]. There are at least two reasons not to put DNA analysis first in line. First, two disease-associated *GAA* variants cannot always be identified by standard diagnostics, for example when the variant is located in an intron [2]. Second, it is not always known whether the variant is disease-associated, despite the large number of variants listed at www.pompevariantdatabase.nl and at the LOVD.

Our present studies focused on the performance of various enzymatic procedures for diagnosing Pompe disease. The pro's and con's of various genetic procedures are largely known and we mention them only briefly. Sequencing of only the *GAA* exons and adjacent regions bears the risk that disease-associated variants in the promoter sequences or deeper intronic sequences are missed. This may be approached using

analysis at the mRNA level as shown by us previously [2, 3]. Another limitation is the identification of variants of uncertain significance (VUS) necessitating the performance of an expression assay [4, 5]. NGS procedures, developed for disease specific gene panels, are gaining field, but their application is expected to increase the number of VUSs. The great advantage of an enzymatic assay over DNA analysis is that the enzymatic assay captures in principle all disease-associated *GAA* variants in one. However, when it comes to carrier detection and genetic counseling in families with an index Pompe case, DNA analysis is essential since the activity ranges of carriers and healthy individuals show overlap [6].

References supplementary data

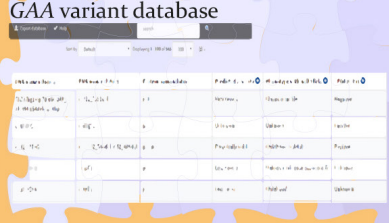
1. Niño, MY, In 't Groen, SLM, Bergsma, AJ, van der Beek, N, Kroos, M, Hoogeveen-Westerveld, M, *et al.* (2019). Extension of the Pompe mutation database by linking disease-associated variants to clinical severity. *Hum Mutat* **40**:1954-1967
2. Bergsma, AJ, In 't Groen, SL, Verheijen, FW, van der Ploeg, AT, and Pijnappel, W (2016). From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides. *Mol Ther Nucleic Acids* **5**: e361.
3. Bergsma, AJ, Kroos, M, Hoogeveen-Westerveld, M, Halley, D, van der Ploeg, AT, and Pijnappel, WW (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. *Hum Mutat* **36**: 57-68.
4. Kroos, M, Pomponio, RJ, van Vliet, L, Palmer, RE, Phipps, M, Van der Helm, R, *et al.* (2008). Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. *Hum Mutat* **29**: E13-26.
5. Kroos, M, Hoogeveen-Westerveld, M, Michelakakis, H, Pomponio, R, Van der Ploeg, A, Halley, D, *et al.* (2012). Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. *Hum Mutat* **33**: 1161-1165.
6. Reuser, Verheijen, F, Kroos, M, Okumiya, T, Van Diggelen, O, Van der Ploeg, A, *et al.* (2010). Enzymatic and molecular strategies to diagnose Pompe disease. *Expert Opin Med Diagn* **4**: 79-89.



Supplementary Figure S1. Series of assays used for diagnosis of patient 32. DBSs, leukocytes and fibroblasts were analyzed as well as the Glc4 content of urine. [1] = pathogenic variant from the mother, [2]= pathogenic variant from the father, * a pilot study by DBS showed GAA activity in the grey zone, and BGAL activity higher than normal ranges. Technical replicates were therefore performed (**), as well as assays in leukocytes using glycogen or 4MUG (plus acarbose) as substrate and in fibroblast with 4MUG as substrate.

Pompe disease

GAA variant database



CHAPTER 3

EXTENSION OF THE POMPE MUTATION DATABASE BY LINKING DISEASE-ASSOCIATED VARIANTS TO CLINICAL SEVERITY

Monica Y. Niño[1,2,3,#], Stijn L.M. in 't Groen[1,2,3,#], Atze J. Bergsma[1,2,3], Nadine A.M.E. van der Beek[1,3], Marian Kroos[2], Marianne Hoogeveen-Westerveld[2], Ans T. van der Ploeg[1,3] and W.W.M. Pim Pijnappel[1,2,3,*]

[1]Department of Pediatrics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

[2]Department of Clinical Genetics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

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

[#]equal contribution

*Correspondence: w.pijnappel@erasmusmc.nl

Human Mutation. 2019 Nov; 40(11):1954-1967. doi.org/10.1002/humu.23854
Front Cover, *Human Mutation*, Volume 40, Issue 11, Nov 2019. doi:10.1002/humu.23934

3

Extension of the Pompe mutation database by linking disease-associated variants to clinical severity

Monica Y. Niño, Stijn L.M. in 't Groen, Atze J. Bergsma, Nadine A.M.E. van der Beek, Marian Kroos, Marianne Hoogeveen-Westerveld, Ans T. van der Ploeg and W.W.M. Pim Pijnappel

ABSTRACT

Pompe disease is an autosomal recessive lysosomal storage disorder caused by disease-associated variants in the acid alpha-glucosidase (*GAA*) gene. The current Pompe mutation database provides a severity rating of *GAA* variants based on *in silico* predictions and expression studies. Here, we extended the database with clinical information of reported phenotypes. We added additional in silico predictions for effects on splicing and protein function and for cross reactive immunologic material (CRIM) status, minor allele frequencies, and molecular analyses. We analyzed 867 patients and 562 *GAA* variants. Based on their combination with a *GAA* null allele (i.e. complete deficiency of GAA enzyme activity), 49% of the 422 disease-associated variants could be linked to classic infantile, childhood, or adult phenotypes. Predictions and immunoblot analyses identified 131 CRIM negative and 216 CRIM positive variants. While disease-associated missense variants were found throughout the GAA protein, they were enriched up to seven-fold in the catalytic site. Fifteen percent of disease-associated missense variants were predicted to affect splicing. This should be confirmed using splicing assays. Inclusion of clinical severity rating in the Pompe mutation database provides an invaluable tool for diagnosis, prognosis of disease progression, treatment regimens, and the future development of personalized medicine for Pompe disease.

KEYWORDS cardiac and skeletal muscle disorder, genotype-phenotype relationship, glycogen storage disease type II, lysosomal storage disease, www.pompecenter.nl

INTRODUCTION

Pompe disease or glycogen storage type II (MIM# 232300) is an autosomal recessive disorder caused by deficiency of acid α-glucosidase (GAA; NP_000143.2). Partial or complete GAA deficiency is caused by disease-associated sequence variants in the *GAA* gene, resulting in lysosomal glycogen accumulation in many cell types, with profound pathology in cardiac and skeletal muscle.

Pompe disease presents as a spectrum of phenotypes. Classic infantile patients have a rapidly progressing phenotype, with hypertrophic cardiomyopathy and general muscle weakness. Without therapy, these patients die within the first year of life. Pompe patients with symptom onset at childhood or adulthood have a more slowly progressive disease, leading to mobility problems and respiratory difficulties, but generally without hypertrophic cardiomyopathy. Most of these patients will become ventilator and wheelchair dependent at some point in time (Gungor & Reuser, 2013; Hirschhorn & Reuser, 2001; Schoser et al., 2015).

The standard treatment for Pompe disease is enzyme replacement therapy (ERT) with recombinant human GAA (rhGAA). In classic infantile Pompe disease, ERT reduces cardiac hypertrophy, improves survival, and helps to achieve important developmental milestones that were previously unmet (Chakrapani, Vellodi, Robinson, Jones, & Wraith, 2010; Kishnani et al., 2009; Van den Hout et al., 2000). Numerous studies in children and adults have shown improvement of muscle strength, stabilization of pulmonary function, decreased fatigue and extended survival. However, there is variation in response and not all patients respond equally well (Broomfield et al., 2016; de Vries et al., 2012; Gungor et al., 2013; Gungor et al., 2016; Hahn et al., 2018; Kuperus et al., 2017; Papadopoulos et al., 2017; Regnery et al., 2012; Stepien, Whitby, Roberts, & Sharma, 2015; Strothotte et al., 2010; van Capelle et al., 2010; van der Ploeg et al., 2010).

Patients who do not produce detectable GAA protein are termed cross reactive immunologic material (CRIM) negative. In these patients, both alleles carry a CRIM negative variant, and the phenotype is per definition classic infantile. Whereas CRIM positive patients have at least one disease-associated GAA variant that allows expression of GAA protein, the protein produced is enzymatically inactive due to misfolding, defects in glycosylation, defects in intracellular routing to the lysosome, and so forth. The second allele in CRIM positive patients can contain either a CRIM positive or a CRIM negative GAA variant. CRIM positive patients who have no detectable

GAA enzyme activity develop a classic infantile phenotype, whereas patients with a childhood or adult phenotype still retain some enzymatic activity and are therefore always CRIM positive. CRIM negative infants have a greater tendency to develop anti-GAA neutralizing antibodies compared to CRIM positive infants (Kishnani et al., 2010; van Gelder et al., 2016). Therefore, it is important to establish the CRIM status of GAA variants.

The rapidly expanding number of GAA sequence variants that have been discovered over the past 15 years has provided the basis for the establishment of the Pompe disease mutation database at www.pompevariantdatabase.nl. The present version of the Pompe mutation database lists all known GAA variants and provides a prediction of pathogenicity based on results obtained by in silico prediction algorithms, from the measured effect on GAA synthesis and enzymatic function when expressed in vitro, and from the type of variant. This approach resulted in a rating system with six different categories ranging from "very severe" to "nonpathogenic" (M. Kroos et al., 2008; M. A. Kroos et al., 2012). Here, we have extended the in silico predictions by analyzing all GAA variants using Alamut software for effects on RNA processing and protein function, and used these in silico predictions as well as published data to report or predict CRIM status. We scrutinized the literature and report clinical phenotypes that are associated with GAA variants. The results of our study described here have been included in the update of the Pompe mutation database, that will from now on be referred to as the "Pompe disease GAA variant database" under the link (www.pompevariantdatabase.nl). The improved database should help doctors, genetic counselors and scientists to better predict disease outcome in patients diagnosed with Pompe disease. As well as providing new insight into variant severity, it will also improve prediction of prognosis in newborn screening programs, and support decision-making on therapeutic intervention.

The results of our study described here have been included in the update of the Pompe mutation database, that will from now on will be referred to as the "Pompe disease *GAA* variant database" under the same link (www.pompecenter.nl). The improved database should help doctors, genetic counselors and scientists to better predict disease outcome in patients diagnosed with Pompe disease. As well as providing new insight into variant severity, it will also improve prediction of prognosis in newborn screening programs, and support decision-making on therapeutic intervention.

METHODS

Annotation of variants

Variant annotations and classification conform to recommendations of the human genome variation society (HGVS) according to update 2016 (den Dunnen et al., 2016). NM_000152.3 was used as reference sequence for GAA mRNA. Annotation of variants present in intronic regions were based on the LRG_673 reference sequence. Position c.1 represents the first nucleotide of the translation start codon ATG located in exon 2. NP_000143.2 was used for annotation of GAA protein variants.

Minor allele frequency and reference SNP number

Minor allele frequency (MAF) was examined using dbSNP, Exac, Genome of the Netherlands (GoNL), ESP, and HGVD, with 1 January 2017 as last entry. The MAF reported in the database was based on the highest MAF score seen in any of the aforementioned databases. The link for the most recent information on MAF has been provided in the database. Reference SNP (RS) numbers, when available, are also provided.

In silico prediction of pathogenicity

These were performed using Alamut software. Missense predictions were performed using Mutation Taster, SIFT, and Align GVGD.

Splicing prediction

Splice variants were defined as intronic positions –1 and –2 relative to the canonical splice acceptor site, and intronic positions +1 and +2 relative to the canonical splice donor site. So far, all splice variants defined in this way had been listed as "very severe". To better evaluate variants' possible effect on pre-mRNA splicing, we reevaluated all variants using Alamut® software and provided this information in the database (Algorithms used are listed in Table S1). Splice predictions were deemed to potentially affect splicing when at least two of the five algorithms predicted more than a 10% difference in splice site strength as described previously (Bergsma et al., 2015). In some cases it was possible to use the literature to retrieve functional evidence of variants that affect splicing. These findings were added to the column "biochemical evidence of pathogenicity".

Prediction of CRIM status

If CRIM status had been determined experimentally by endogenous protein (i.e., immunoblot) or RNA expression (i.e., RT-qPCR) analysis, this was indicated at the "variant info" page under "biochemical evidence of CRIM status" or "biochemical

evidence of pathogenicity", respectively. We also predicted CRIM status based on the type of variant for all variants, and reported this in the database. Variants that caused a frame shift were predicted to be CRIM negative. Missense variants were predicted to be CRIM positive, provided that no effect on splicing was predicted. Known and predicted splice variants were classified with an "unknown" CRIM status, unless experimental evidence was reported on expression of in-frame mRNA transcripts. Please note that a) all prediction need to be confirmed experimentally. For example, it is known that certain variants that cause a premature stop codon can escape nonsensemediated decay. In those cases, the actual CRIM status could be positive rather than negative. It is also known that splicing outcome is difficult to predict (e.g., see Bergsma et al., 2015). b) the CRIM status of a patient is determined by two disease-associated variants, and all patients with symptom onset at childhood or adulthood have some residual GAA enzyme activity. For this reason, these patients always have a CRIM positive status, caused by the presence of at least one CRIM positive variant.

Patient classification and phenotypic information

We classified published patient information based on the criteria stated in Gungor and Reuser (2013). Patients were classified with classic infantile Pompe disease if they presented symptoms at or under 12 months of age, and had evident signs of a hypertrophic cardiomyopathy (cardiac enlargement visible on chest X-ray, evidence of hypertrophy by echography). Patients were classified with childhood Pompe disease if the age of symptom onset was before 18 years of age and evident hypertrophic cardiomyopathy was absent. Patients were classified with adult Pompe disease if the first symptoms presented at the age of 18 years or later. If specific clinical information was reported in the literature, it was included in the database as indicated. Since familial connections were not always clearly stated in publications, no familial connections were taken into account and all cases were reported independently. We did all we could to avoid duplicating patients who had been described in more than one publication. It nonetheless remains possible that some duplicate cases were not detected.

Availability of the data

The open access Pompe disease *GAA* variant database is available at www.pompecenter.nl. In addition, all variants and patient phenotypes associated with the variants are included in the Leiden Open Variation Database (LOVD) at www.lovd.nl/gaa.

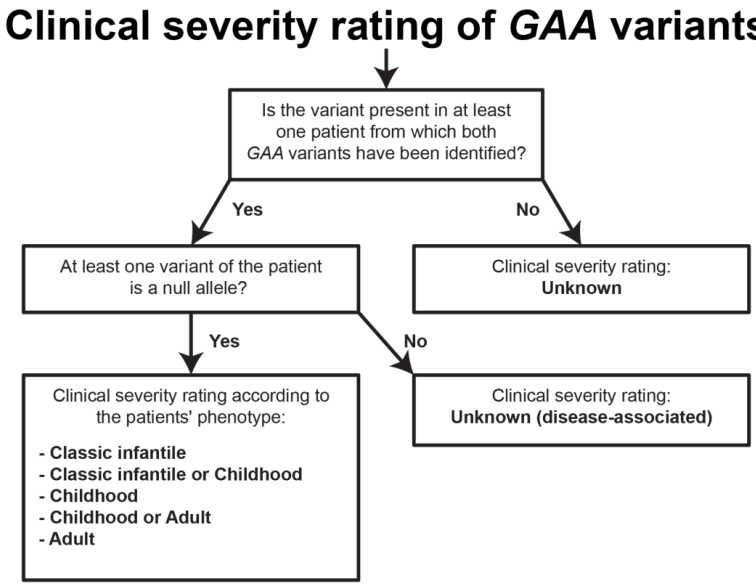
RESULTS

Overview of the updated Pompe database

A flow chart for the criteria used for clinical severity rating is given in Figure 1a. Clinical severity rating of variants was performed based on patients with two identified GAA variants, at least one of which was a null allele, which is defined as a GAA variant who's protein product lacks any detectable GAA enzyme activity. The classification of null alleles (listed in Table S2) was made based on their association with the classic infantile phenotype. In some cases data have been reported that are derived from site directed mutagenesis of GAA expression constructs, but no patients have been reported that have such variant. In those cases, the clinical phenotype was classified as "unknown". When patients were reported but the second GAA allele was not a null allele, was not reported or had an unknown severity, the clinical phenotype was classified as unknown (disease-associated). In those cases, the variantdoes causePompe disease because patientshave beenreported, and thus this variant is disease -associated. However the variant cannot be associated to a certain phenotype because the contribution of the second allele to the phenotype is unknown. The updated open access Pompe disease GAA variant database is available at www.pompevariantdatabase.nl. The main access page contains the most essential information, including a) the location of the variant based on exon or intron number, b) the DNA, RNA and protein nomenclature according to the recommendations of the HGVS (version 2016) (den Dunnen et al., 2016), c) the predicted severity, d) the phenotype when combined with a null alelle, e) the CRIM status according to literature reports or based on prediction according to type of variant, and f) links to the variants and to the patients described in the literature. (Figure 1b).

The link to variants includes the type of variants, their predicted severity, the phenotype of variants when combined with a null allele, biochemical evidence and in silico prediction of (non-) pathogenicity, MAF, RS-number, biochemical evidence and in silico prediction of CRIM status, splicing prediction for all variants using different prediction algorithms, severity predictions for all missense variants, and the link to publications on the variant. The link to patients contains the clinical phenotype of patients found in the literature with that variant in combination with a second disease-associated variant, additional clinical symptoms, the patients' geographical location, and the link to publications in which the patient was reported. The extended database includes an analysis of all data contained in the Pompe disease mutation database. Eight new variants have been added since the last update in May 2016 (Table S3).

A



B

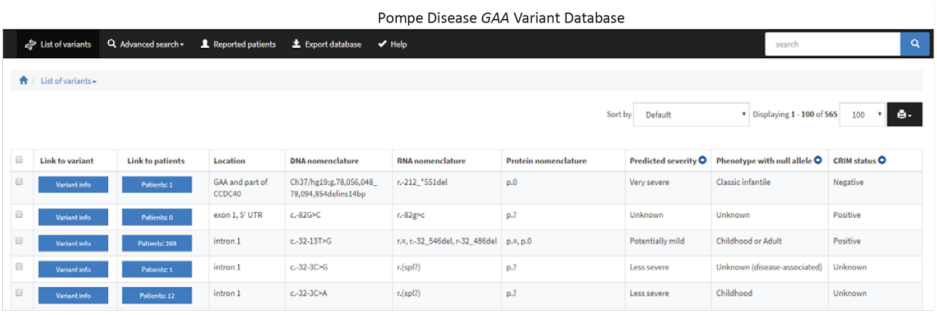


Figure 1: The Pompe Disease *GAA* Variant Database. **(A)** Flowchart for determining the clinical severity rating in the updated database. **(B)** Screenshot of the updated Pompe Disease *GAA* Variant database home page at **www.pompevariantdatabase.nl**

Analysis of GAA variants based on clinical phenotype

Of the total number of 562 described variants, 422 are listed as disease-associated. We analyzed types of GAA variants according to their RNA and protein nomenclature (HGVS version 2016), and stratified these based on the clinical phenotype, being classic infantile, childhood, or adult (Figure 2). The variants present in a total of 867 patients were analyzed (Figure S2). In 49% of patients with Pompe disease reported in the mutation database (data not shown), the second GAA allele contained a null allele, and these patients were used for clinical severity rating of GAA variants. The reason for this is as follows. Usually, the second allele is a null allele, from which no residual GAA enzyme

activity is produced. Any residual GAA enzyme activity is caused by the first allele. In this way, the disease-associated variant on the first allele can be rated. When the second allele would also have residual GAA enzyme activity, both alleles will contribute to the total residual GAA enzyme activity, and it would not be possible anymore to attribute the activity and the clinical severity rating to one of the two alleles. The remaining 140 of 562 GAA variants are classified as unknown (Table S4). The nondisease-associated variants include alleles such as the Caucasian variant c.271G>A (GAA2; MAF 0.02), and the Asian GAA variants c.1726G>A (MAF 0.0179) plus c.2065G>A (MAF 0.08797; Table S5). These variants have previously been shown to lower GAA enzymatic activity to some extent, however, they cannot cause Pompe disease (M.A. Kroos et al., 2008; Shigeto et al., 2011; Swallow et al., 1989; van Diggelen et al., 2009).

First, we performed stratification of GAA variants. At the RNA level, substitutions were the predominant type of variants in all three phenotype groups. Other types of variants including splicing variants and deletions were present at a lower percentage without a clear trend relative to the phenotype. At the protein level, analysis of substitutions showed that the largest fraction was represented by missense variants, followed by nonsense variants, again without a clear trend relative to the phenotype. In 15% of cases, splicing prediction of disease-associated missense variants revealed possible effects on splicing (Figure S1). The total number of unique variants was similar between the three phenotype groups and ranged from 134 in adult to 154 in childhood and 173 in classic infantile patients. A different result was obtained when the number of patients with a certain type of variant was plotted per disease onset group. The variants in all patients with a known phenotype were analyzed (Figure S2). We note that two GAA variants per patients were counted. A typical adult patient contains one null allele and one less severe allele, whereas a classic infantile patient contains two null alleles. In this case, the percentage of patients with a splicing variant (intron variant) increased from less than 1% in the classic infantile group to 18% in the childhood group and 44% in the adult group (Figure 2b). At the protein level, this effect was reflected at the level of the category deletion (translation initiation site), as the IVS1 variant causes skipping of the translation initiation site-containing exon 2. This effect can be attributed to the large number of childhood and adult Caucasian Pompe patients with the splicing (intron variant) variant c.-32-13T>G (IVS1). This variant leads to partial or complete skipping of exon 2 during GAA pre-mRNA splicing, which has been described previously (Bergsma et al., 2015; Boerkoel et al., 1995; Dardis et al., 2014; Huie, Chen, Brooks, Grix, & Hirschhorn, 1994; van der Wal, Bergsma, Pijnenburg, van der Ploeg, & Pijnappel, 2017; van der Wal, Bergsma, van Gestel et al., 2017).

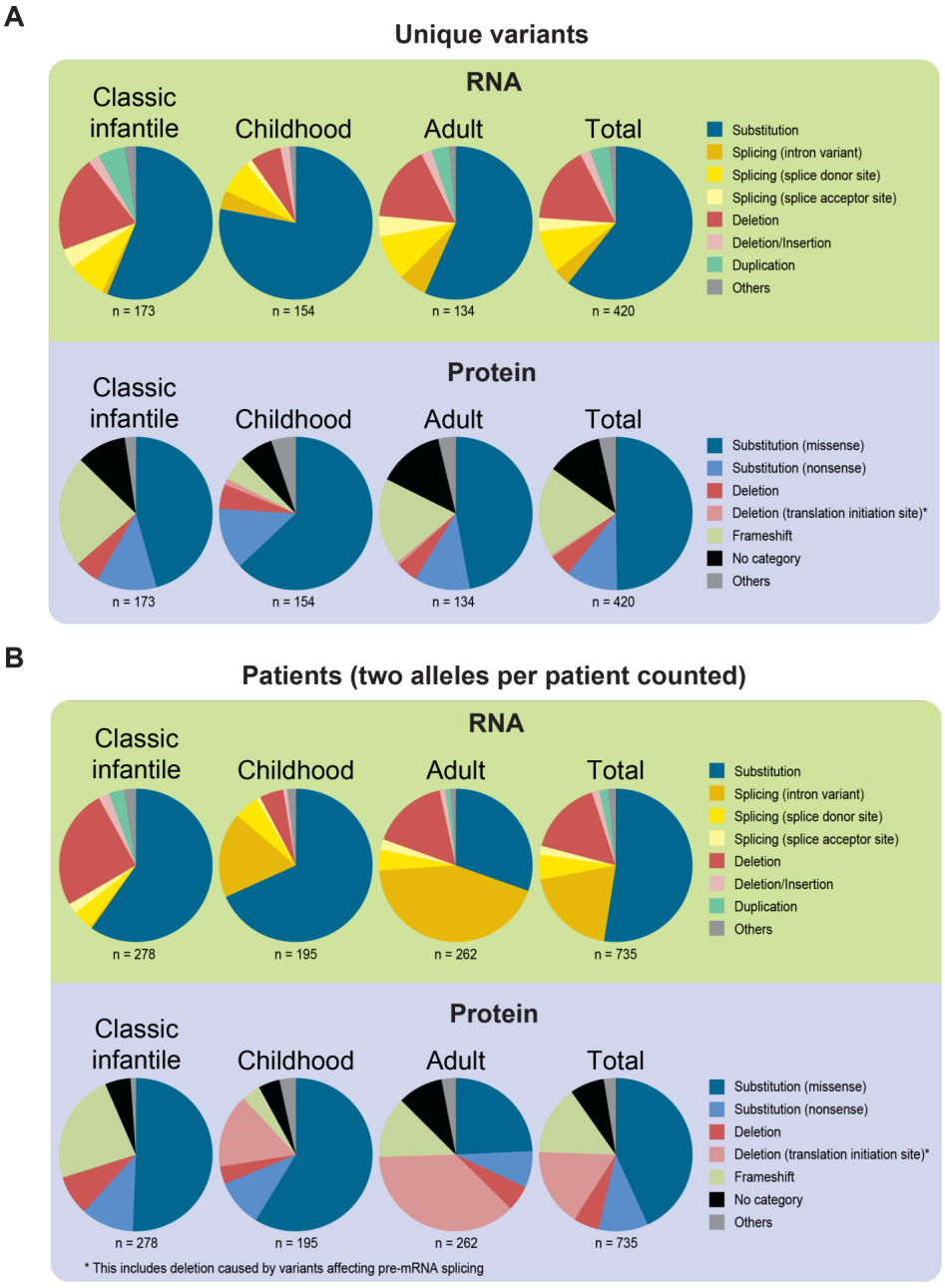


Figure 2: Frequencies of *GAA* variants and patients with specific *GAA* variants, stratified for disease severity. (A) Frequencies of unique variants, classified at the RNA level (upper panel) and protein level (lower panel). The consequences of variants that affect pre-mRNA splicing was only included in specific categories (e.g. deletion, insertion, frameshift) at the protein level if experimental evidence was available. A list of all included categories is shown in Supp. Table S6. (B) Frequencies of variant types in patients, classified at the RNA level (upper panel) and protein level (lower panel). Two *GAA* alleles per patient were counted. Please note that 1) only one *GAA* allele could be counted in patients with an unknown second *GAA* variant; and that 2) the allele was counted twice in patients with a homozygous genotype

CRIM status of classic infantile variants

The CRIM status of a Pompe patient is determined by the combined effect of both GAA variants: a CRIM negative patient has by definition two CRIM negative variants. A CRIM positive patient has either one CRIM positive (and one CRIM negative) or two CRIM positive variants. Only for a limited number of variants, experimental evidence of CRIM status, either by immunoblot or by RT-PCR analyses, has been published. This resulted in 46 CRIM negative, and 15 CRIM positive variants (Figure 3a,b). At the RNA level, predominant categories for CRIM negative variants with experimental evidence were substitutions, splicing (splice donor site), deletions, and duplications (Figure 3a). At the protein level, most of these variants were substitutions (nonsense) or frameshifts, both of which cause premature translation initiation sites, whereas substitutions (missense) were absent (Figure 3b). For CRIM positive variants with experimental evidence, most variants were substitutions or splicing variants. At the protein level, most variants were classified as substitutions (missense) or deletions. A small fraction of CRIM positive variants were frameshifts, suggesting that in these cases the nonsense-mediated decay pathway did not completely degrade the mRNA.

Prediction of CRIM status was performed on the basis of the variants' effects on the reading frame, but it should be noted that experimental validation is needed for these cases. 201 variants were predicted to be CRIM positive, and these were substitutions (missense) and deletions. Eighty-five GAA variants were predicted to be CRIM negative, the majority of which were frameshift and substitution (nonsense) variants. The total numbers of classified variants that is those determined experimentally and those predicted were 216 CRIM positive variants (62%) and 131 CRIM negative variants (38%).

Distribution of disease-associated variants along the GAA gene

Figure S3 shows that disease-associated variants are distributed throughout the GAA RNA (Figure S3A) and protein (Figure S3B) with a slight enrichment in exons 4-16 relative to the flanking exons. A similar pattern was obtained when the distribution of variants was plotted along exons according to clinical severity (Figure S3C). For introns, disease-associated variants were found in all introns except for intron 5, without a clear pattern (Figure S3D). We note that intronic variants are rarely investigated at the functional level and that only the exon flanking regions are sequenced in standard diagnostic analysis.

However, mapping of disease-associated substitution (missense) variants to functional GAA protein domains (Figure 4a), showed a strong enrichment (up to seven-fold) in the catalytic sites relative to the N-terminal signal peptide and the C-terminal distal β-sheet

(Figure 4b). Variants associated with the classic infantile phenotype showed a similar enrichment in the catalytic sites. When GAA variants in ExAC and dbSNP databases (which contains both diseaseassociated and non disease-associated variants) were plotted, no such enrichment in the catalytic domain was observed (Figure 4c). These results are in line with the critical function of the catalytic core for GAA enzyme activity.

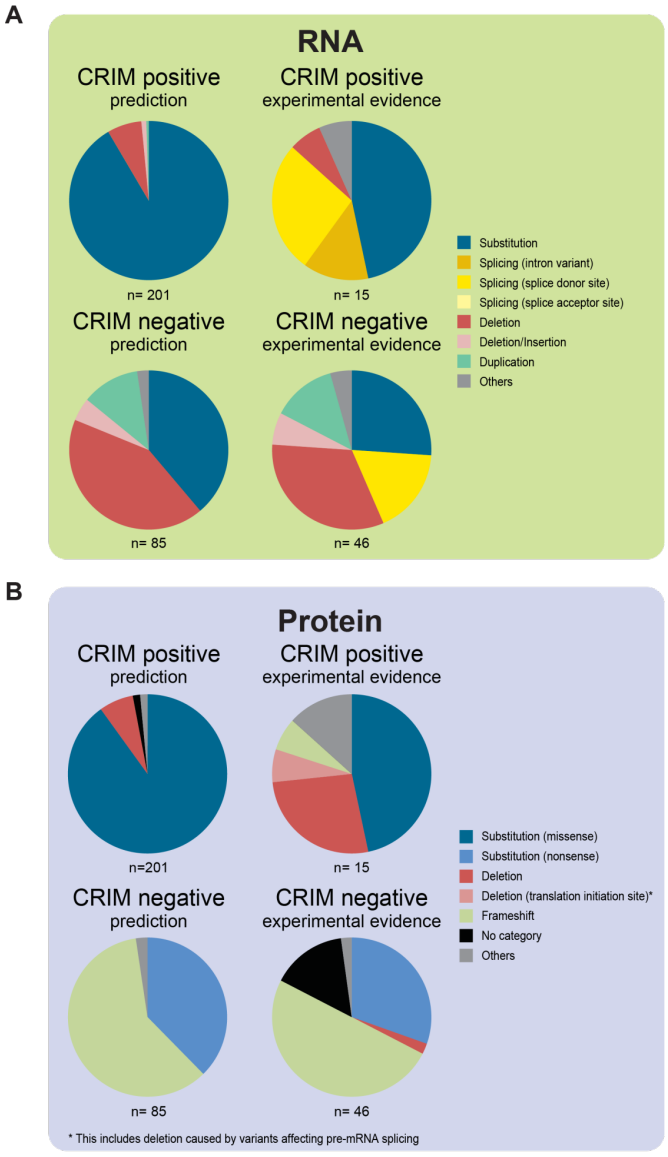


Figure 3: Distribution of CRIM positive and CRIM negative variants per variant type at the RNA level (A) and protein level (B). Left panels in (A) and (B): predicted variants; right panels: experimentally tested variants. CRIM, cross reactive immunologic material

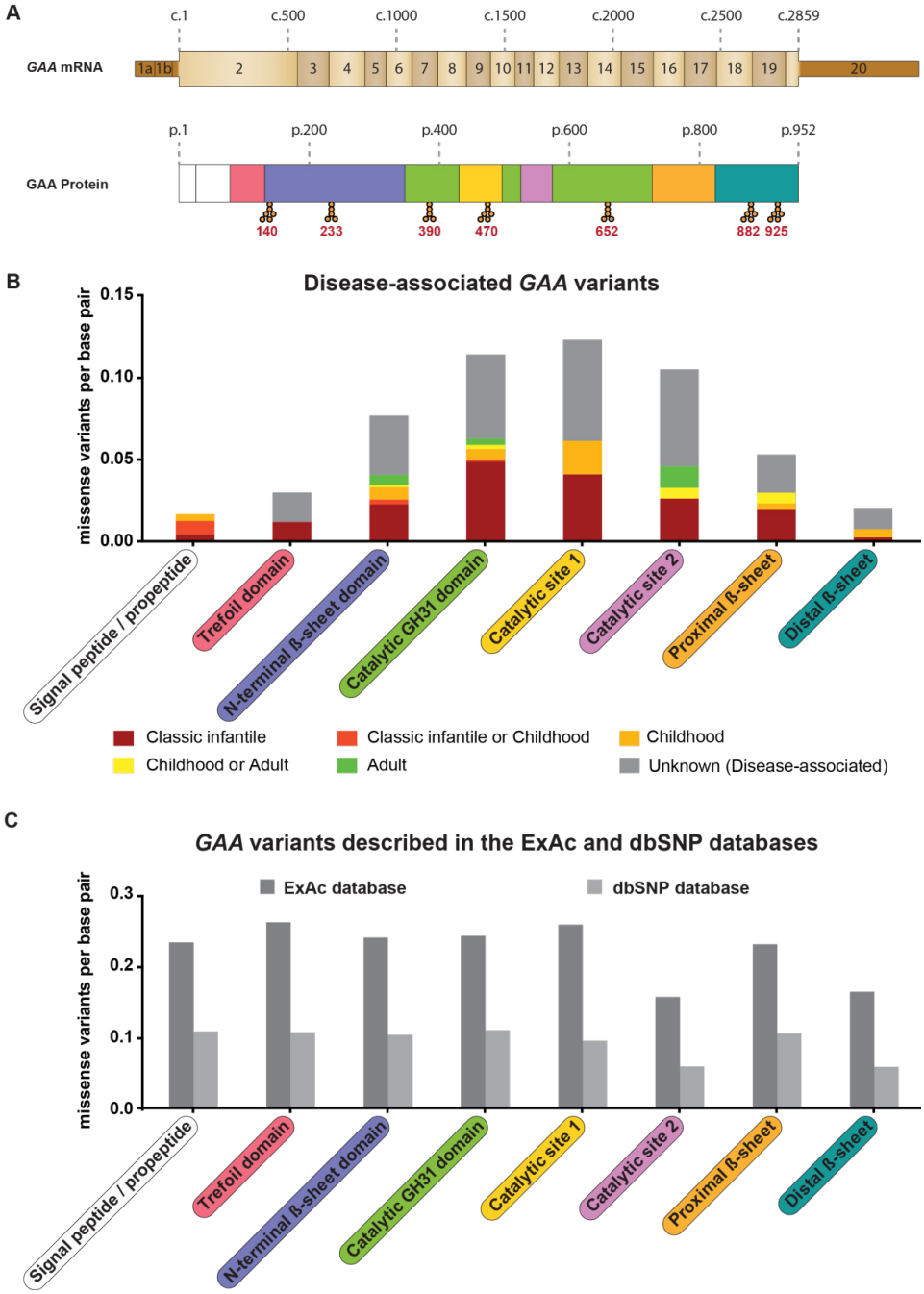


Figure 4 Distribution of missense variants along the GAA protein. (a) Cartoon of the *GAA* gene indicating the locations of exons. Dark brown regions indicate the 5′and 3′untranslated regions. The translation start codon is indicated as c.1 in exon 2. The second cartoon indicates the GAA protein domains and the sites modified by M6P residues. (b) The distribution of disease-associated variants in GAA protein domains according to frequency and clinical severity rating. Numbers are corrected for the length of each domain. (c) as (b), but now for all *GAA* variants, including both disease-associated and non-disease-associated missense variants, listed in the ExAc database

Geographic distribution of the most common disease-associated GAA variants

Table 1 shows the most common disease-associated variants per severity (when combined with a null allele), their geographical distributions, and their linked phenotypes. We note that two CRIM negative alleles are required to give rise to a CRIM negative phenotype. Variants associated with the classic infantile phenotype include the very common variants c.525del (Caucasian, CRIM negative), c.1935C>A (Asian, CRIM positive), and c.2481_102_2446+31del (delex18; Caucasian, CRIM positive). Other common variants associated with the classic infantile phenotype include the CRIM negative variants c.2560C>T (Latin American and African/African American), c.1411_1414del (Asian), and c.2237G>A (Caucasian), and the CRIM positive variants c.925G>A, c.1933G>A, c.784G>A, and c.1655T>C (all mostly Caucasian).

Three common variants that are classified as either associated with the classic infantile or childhood phenotypes include c.1064T>C (Caucasian), c.670C>T (Caucasian and Asian), and c.1561G>A (Asian). Variants associated with the childhood phenotype include the common variants c.1857C>G, c.796C>T (both Asian), c.-32-3C>A (Caucasian and Latin American), c.875A>G (Caucasian), and c.1082C>T (Asian).

Variants that are associated with the childhood or adult phenotypes are strongly dominated by the c.-32-13T>G (IVS1) variant (Caucasian and Latin American). This variant is present in 86% (253/294) of patients in this category. Other common variants associated with the childhood or adult phenotype include c.2238G>C (Asian), c.2014C>T and 2173C>T (both Caucasian), and c.1634C>T (Caucasian and Asian). Variants associated with the adult phenotype include the common variants c.2647-7G>A (Caucasian) and c.1585_1586delinsGT (Asian). This analysis demonstrates that common disease-associated variants can be found for each phenotype and geographical region.

Table 1. Frequency and geographical distribution of most common *GAA* variants

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Classic infantile (CRIM negative)	c.525del	29		1			30
	c.2560C>T	1	1	3	9	1	15
	c.1411_1414del		11				11
	c.2237G>A	5	1				6
	Total	35	13	4	9	1	**62**

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African/African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Classic infantile (CRIM positive)	c.1935C>A		68				68
	c.2481+102_2646+31del	20			3	1	24
	c.925G>A	7			1		8
	c.1933G>A	4	2		1		7
	c.784G>A	4	2		1		7
	c.1655T>C *	5			1		6
	Total	40	72		7	1	**120**

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African/African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Classic infantile or Childhood	c.1064T>C	9	3		2		14
	c.670C>T	3	3				6
	c.1561G>A	1	3				4
	Total	13	9		2		**24**

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African/African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Childhood	c.1857C>G		6				6
	c.796C>T		5				5
	c.-32-3C>A	1			3		4
	c.875A>G	3	1				4
	c.1082C>T	1	3				4
	Total	5	15		3		**23**

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African/African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Childhood or Adult	c.-32-13T>G	224	3		18	8	253
	c.2238G>C		21				21
	c.2014C>T	7	1				8
	c.1634C>T	3	2	1			6
	c.2173C>T	5			1		6
	Total	239	27	1	19	8	**294**

Phenotype with a null allele	DNA nomenclature	Caucasian (n)	Asian (n)	African /African-American (n)	Latin-American (n)	Undetermined (n)	Total (n)
Adult	c.2647-7G>A	9					9
	c.1585_1586delinsGT		4				4
	Total	9	4				**13**

*The c.1655T>C variant has been classified to be associated with the classic infantile phenotype, although we note that three patients have been reported that carry this variant and showed a childhood onset phenotype.

DISCUSSION

This study describes the extension of the Pompe mutation database (now named: Pompe disease *GAA* variant database) www.pompevariantdatabase.nl, which now includes a number of new features to improve its utility. First, by linking variants to the patient information reported in the literature, we have extended severity ratings with the clinical phenotype. Second, we have added information on the MAF, geographical distribution, onset and types of clinical symptoms; and have performed predictions for a) effects on splicing, b) severity of missense variants, and c) CRIM status. Third, we have provided separate lists of null alleles, variants with an 'unknown' clinical severity rating, and variants that can lower the activity of GAA without being capable of causing Pompe disease. An important asset of the database which is freely accessible online is that it provides comprehensive information on the genotypephenotype relationship. In the previous version of the database, the severity of the majority of GAA variants was predicted in silico. For a subset of variants, expression studies were available. The extended database now considers the clinical consequences of a GAA variant, and it enables a better judgment of the potential severity of variants. This is important for several aspects related to the diagnosis and treatment of Pompe disease as outlined below.

Novel aspects of the database

Clinical information

Pompe disease has a broad spectrum of disease onset and symptoms. It is important to be able to distinguish between classic infantile and late onset phenotypes, because the classic infantile phenotype is so severe that immediate treatment with ERT is required. This is reflected by severe muscle weakness and hypertrophic cardiomyopathy, which results in death within the first year of life if left untreated due tot cardiorespiratory insufficiency. It has also become evident that classic infantile patients slowly develop abnormalities in the central nervous system as detected using nuclear magnetic resonance, which can result in diminished intellectual performance (Ebbink et al., 2018). When patients have a late onset form of Pompe disease, at present it cannot be accurately predicted when symptoms will a rise and how severe these will be. It is therefore important to collect the clinical information on specific disease-associated variants to better understand the genotype–phenotype relationships. These aspects are relevant for newborn screening programs and genetic counseling.

In silico predictions

Whereas clinical information is crucial for judging the severity of GAA variants, in silico predictions can be helpful as they can provide part of the evidence that leads

to determination of the severity of variants. In addition, in silico predictions can provide evidence for the underlying mechanism involved in the disruption of GAA enzyme activity. For example, splicing predictions can point to possible effects on splicing of variants that would normally not be suspected of such effects. In particular, missense variants are often not evaluated for their effect on splicing, whereas several reports have suggested that missense variants can often affect splicing (Lim, Ferraris, Filloux, Raphael, & Fairbrother, 2011; Soukarieh et al., 2016). Also in Pompe disease, we recently described a missense variant (c.1075G>A; p.G359R) that appeared to be disease-associated mainly because it affects splicing rather than changing the amino acids sequence. The effect of this variant was predicted in silico, and this was experimentally confirmed (Bergsma et al., 2015). On the other hand, predictions can also fail, such as is the case for the IVS1 variant, which has an experimentally confirmed effect on splicing that cannot be predicted in silico (Boerkoel et al., 1995; Dardis et al., 2014; Huie et al., 1994; van der Wal, Bergsma, Pijnenburg et al., 2017; van der Wal, Bergsma, van Gestel et al., 2017). In addition, comparison of predictions of the severity of amino acid changes with expression studies revealed failure of at least two out of three prediction programs to predict the experimentally found deleterious effect on GAA enzyme activity. This was the case for the following GAA missense variants: c.307C>T, c.307C>G, c.380G>T, 701C>G, c.947A>T, c.1040C>G, c.1381G>A, c.1456G>C, c.1834C>T, c.1905C>A, c.2105G>T, and c.2210C>A. Conversely, there were also variants for which in silico programs predicted a deleterious effect, while expression studies showed no effect on enzyme activity. These variants included c. 664G>A, c.2132C>G. Finally, it is known that the strength of unnatural transcription termination signals depends on the adjacent sequences andthatsuchsignal canbe 'leaky', reinforcing theneed forexperimental confirmation (Dabrowski, Bukowy-Bieryllo, & Zietkiewicz, 2015). In all these cases, it is important to know whether a variant has some residual enzyme activity, whether the GAA protein is still expressed, or whether it is completelydeleterious. This is crucial tomake a distinction between the classic infantile and late onset forms of Pompe disease, and within the classic infantile population, to distinguish between CRIM positive and CRIM negative patients.

CRIM status

CRIM negative classic infantile patients generally have a poorer prognosis compared to CRIM positive classic infantile patients when treated with ERT. One of the explanations is the tendency of CRIM negative patients to form high antibody titers during ERT (Kishnani et al., 2010; van Gelder et al., 2016). This can be explained by the lack of any endogenous GAA protein in CRIM negative patients, which results in the recognition of rhGAA by the immune system and the generation of rhGAA-specific antibodies.

However, most CRIM positive patients with Pompe disease also develop anti-rhGAA antibodies, likely due to posttranslational differences between the endogenous GAA protein and rhGAA, although on average the antibody titers are lower in CRIM positive patients compared to CRIM negative patients (Kishnani et al., 2010; Poelman et al., 2018). Antibodies have the potential to interfere with ERT by binding to rhGAA and thereby neutralizing its activity and/or uptake by muscle cells. This has led to the exploration of immune modulating strategies before the start of ERT in classic infantile patients (Elder et al., 2013; Mendelsohn, Messinger, Rosenberg, & Kishnani, 2009; Messinger et al., 2012), in particular in CRIM negative patients. At least 32% of disease-associated variants listed in the database have been shown or predicted to be CRIM negative. Whereas the majority of missense variants is predicted to be CRIM positive, missense variants that affect splicing may be CRIM negative if they fully abrogate canonical splicing and change the reading frame to introduce a premature stop codon. For experimental validation of variants CRIM status, future efforts should focus on performing standardized assays such as immunoblot analysis combined with sensitive detection. As the generation of leaky wild type splicing can be detected by RT-PCR analysis and is indicative of a CRIM positive status, such validation should also include splicing analysis (Bergsma, van der Wal, Broeders, van der Ploeg, & Pim Pijnappel, 2018), such as the splicing assay for GAA that has been reported by us previously (Bergsma et al., 2015).

MAF

The extended database now also includes information on the MAF of variants. In general, a MAF of <1% is required for a variant to be considered disease-associated, because otherwise there would be many more patients then found in reality. Relatively high MAFs can be seen in the case of certain disease-associated GAA variants that are more frequent. The IVS1 variant, associated with late onset Pompe disease, is truly disease-associated but is clearly enriched in the Caucasian population with a MAF of up to 0.8%. Similarly, c.1935C>A, a disease-associated variant associated with classic infantile Pompe disease is enriched in the Asian population with a MAF of 0.17%. Novel variants for which the severity is unknown should first pass the MAF <1% rule, and this is a useful fast criterium to determine whether the variant could be considered disease-associated or not. This rule also applies to variants that can lower GAA activity but not sufficiently to be truly disease-associated (meaning they can cause Pompe disease) on their own. These include the GAA2 (c.271G>A) (MAF up to 4.9%) variant in the Caucasian population and c.1726G>A; c.2065G>A allele (MAF up to 14%) in the Asian population, both of which are variants that lower GAA enzyme activity to some extent (M. A. Kroos et al., 2008). Considering their MAF and modest effect on

GAA activity, they cannot be considered diseaseassociated variants with sufficient deleterious effect to cause Pompe disease. As screening and diagnostic procedures involve 4MU as substrate, these variants can cause false-positive results that can be excluded using second tier DNA analysis.

Geographical information

We have analyzed the geographical distribution of diseaseassociated variants and this indicated distinct sets of variants in Caucasian and Asian populations. For instance, the IVS1 variant is frequent in the Caucasian population but absent in people from Asian descent, whereas the opposite is true for c.1935C>A. In countries with mixed populations, it is difficult to attribute variants to a certain descent. For example, in Latin America, indigenous populations are mixed with Caucasians and people from African and Asian descent. Many populations including those from Africa, Russia, China, and more areas are currently not covered by the database as it mostly includes variants found in patients from Europe, North America, Australia, Taiwan, and Japan. For future diagnostic purposes in other countries, it will be important to extend the database beyond the current geographical regions.

Identification of disease-associated variants

There are a number of patients in which only one GAA variant has been identified. It is important to stress that this does not imply that Pompe disease can be caused by only a single diseaseassociated variant. This is evident from the lack of symptoms in carriers, such as the parents of Pompe patients. Most likely, lack of identification of a second disease-associated GAA variant is caused by variants that are not detected by standard DNA diagnostics. Currently, standard DNA diagnostic analysis involves Sanger sequencing of exonic regions. This means that intronic regions, 5′ and 3′ untranslated regions, and the GAA promoter are not examined, all of which could in principle harbor disease-associated variants. A first clue to the presence of such variant can be provided by analysis at the mRNA level, which can reveal splicing events or a lack of mRNA expression. We have shown that such analysis is useful for detection of deep intronic variants and for elucidating splicing events in fibroblasts (Bergsma et al., 2015), but in principle such analysis may also be performed in blood. Small deletions are analyzed using multiplex ligation-dependent probe amplification or DNA-qPCR at diagnostic departments, but other chromosomal aberrations such as deletions/duplications or uniparental disomy are currently not systematically analyzed. SNP arrays have recently been shown to be useful for the detection of such cases, which also occur in Pompe disease and may be useful to include in standard DNA diagnostics (Labrijn-Marks et al., 2019).

Utilization of the database for future personalized medicine

Like ERT with rhGAA, next-generation therapies are now under development. Clinical trials have started for modified versions of rhGAA, such as avaglucosidase alpha (former neoGAA; made by Sanofi Genzyme); and ATB200/AT2221 (made by Amicus), which combines rhGAA with a chaperone to improve the stability of ERT. When used as single treatment without ERT, chaperones could be used to treat a subpopulation of Pompe patients with certain missense variants by preventing degradation of endogenous GAA by the endoplasmatic reticulum-associated protein degradation pathway. The structure of rhGAA, which was solved recently (Roig-Zamboni et al., 2017), should aid the design of next generation chaperones for personalized treatment, a process that will be greatly supported by information in the Pompe mutation database. When disease-associated missense variants were mapped on the GAA structure, we found enrichment of deleterious variants in the catalytic domain, which is important for GAA enzyme activity. The signal peptide/propeptide is involved in protein targeting and activation (Choo, Tan, & Ranganathan, 2005), while the trefoil domain is found in a number of extracellular proteins (http://pfam.xfam.org/family/PF00088) and forms a secondary substrate binding site (Roig-Zamboni et al., 2017). Given their (potential) importance, it is surprising that a relatively low percentage of disease-associated missense variants was found in the signal peptide/propeptide and trefoil domains. Nevertheless, these domains also contained deleterious variants, which may be relevant for the idea that chaperones could be designed for trefoil domain variants that would have enhanced specificity compared to catalytic domain chaperones, because the catalytic domain is conserved in neutral glycosidases that would mediate undesired side effects when targeted with a chaperone (Roig-Zamboni et al., 2017). If approved, other potential therapies currently under preclinical development may, depending on the predicted phenotype, be the most suitable for certain types of Pompe patients. These include AAV-mediated gene therapy directed toward the diaphragm (Corti et al., 2017) or the liver (Puzzo et al., 2017), hematopoietic stem cell-mediated lentiviral gene therapy (van Til et al., 2010), and splice switching antisense oligonucleotides for the common c.-32-13T>G (IVS1) variant and for other more rare childhood/adult variants (van der Wal, Bergsma, Pijnenburg et al., 2017; van der Wal, Bergsma, van Gestel et al., 2017). It will be important to continue updating and improving the database as it forms an important reference point for research on genotype-phenotype relationships, diagnostics, treatment regimen, and clinical trial design.

Acknowledgements

We thank Dr. Arnold JJ Reuser for his critical review of the manuscript, and Ton Verkerk for his invaluable assistance with generating the database. This work was supported

by grants from the Administrative Department of Science, Technology and Innovation (Colciencias, Colombia) and Sanofi Genzyme. This project has received funding from the Ministry of Economic Affairs under TKI-allowance under the TKI-programme Life Sciences & Health (project number LSHM16008).

Research on Pompe disease at Erasmus MC is financially supported by "Prinses Beatrix Spierfonds" [project number W.OR13-21, W.OR15-10, W.OR16-07]; Tex Net; Sophia foundation for medical research (SSWO) [project number s17-32]; Metakids [project number 2016-063]; Departamento Administrativo de Ciencia, Tecnología e Innovación (Colciencias), Colombia (Administrative Department of Science, Technology and Innovation, Colciencias, Colombia).

Conflict of interest

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

ORCID

W.W.M. Pim Pijnappel http://orcid.org/0000-0002-7042-2482

REFERENCES

Bergsma, A. J., Kroos, M., Hoogeveen-Westerveld, M., Halley, D., van der Ploeg, A. T., & Pijnappel, W. W. (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. Human Mutation, 36(1), 57–68. https://doi.org/10. 1002/humu.22705

Bergsma, A. J., van der Wal, E., Broeders, M., van der Ploeg, A. T., & Pim Pijnappel, W. W. M. (2018). Alternative splicing in genetic diseases: Improved diagnosis and novel treatment options. International Review of Cell and Molecular Biology, 335, 85 –141.

Boerkoel, C. F., Exelbert, R., Nicastri, C., Nichols, R. C., Miller, F. W., Plotz, P. H., & Raben, N. (1995). Leaky splicing mutation in the acid maltase gene is associated with delayed onset of glycogenosis type II. American Journal of Human Genetics, 56(4), 887–897.

Broomfield, A., Fletcher, J., Davison, J., Finnegan, N., Fenton, M., Chikermane, A., & Vellodi, A. (2016). Response of 33 UK patients with infantile-onset Pompe disease to enzyme replacement therapy. Journal of Inherited Metabolic Disease (Dordrecht), 39(2), 261–271.

van Capelle, C. I., van der Beek, N. A., Hagemans, M. L., Arts, W. F., Hop, W. C., Lee, P., & van der Ploeg, A. T. (2010). Effect of enzyme therapy in juvenile patients with Pompe disease: A three-year open-label study. Neuromuscular Disorders: NMD, 20(12), 775–782. https://doi. org/S0960-8966(10)00572-9

Chakrapani, A., Vellodi, A., Robinson, P., Jones, S., & Wraith, J. E. (2010). Treatment of infantile Pompe disease with alglucosidase alpha: The UK experience. Journal of Inherited Metabolic Disease (Dordrecht), 33(6), 747–750. https://doi.org/10.1007/s10545-010-9206-3

Choo, K. H., Tan, T. W., & Ranganathan, S. (2005). SPdb--a signal peptide database. BMC Bioinformatics, 6, 249. https://doi.org/1471-2105-6249. [pii] 10.1186/1471-2105-6-249

Corti, M., Liberati, C., Smith, B. K., Lawson, L. A., Tuna, I. S., Conlon, T. J., & Byrne, B. J. (2017). Safety of intradiaphragmatic delivery of adenoassociated virus-mediated alphaglucosidase (rAAV1-CMV-hGAA) gene therapy in children affected by pompe disease. Human Gene Therapy. Clinical Development, 28(4), 208–218. https://doi.org/10. 1089/humc.2017.146

Dabrowski, M., Bukowy-Bieryllo, Z., & Zietkiewicz, E. (2015). Translational readthrough potential of natural termination codons in eucaryotes-The impact of RNA sequence. RNA Biology, 12(9), 950–958. https://doi.org/10.1080/15476286.2015.1068497

Dardis, A., Zanin, I., Zampieri, S., Stuani, C., Pianta, A., Romanello, M., & Buratti, E. (2014). Functional characterization of the common c.32–13T >G mutation of GAA gene: Identification of potential therapeutic agents. Nucleic Acids Research, 42(2), 1291–1302.

van Diggelen, O. P., Oemardien, L. F., van der Beek, N. A., Kroos, M. A., Wind, H. K., Voznyi, Y. V., & Reuser, A. J. (2009). Enzyme analysis for Pompe disease in leukocytes; superior results with natural substrate compared with artificial substrates. Journal of Inherited Metabolic Disease (Dordrecht), 32(3), 416–423. https://doi.org/10.1007/s10545009-1082-3

den Dunnen, J. T., Dalgleish, R., Maglott, D. R., Hart, R. K., Greenblatt, M. S., McGowan-Jordan, J., & Taschner, P. E. (2016). HGVS recommendations for the description of sequence variants: 2016 Update. Human

Mutation, 37(6), 564–569. https://doi.org/10.1002/humu.22981

Ebbink, B. J., Poelman, E., Aarsen, F. K., Plug, I., Régal, L., Muentjes, C. ... van den Hout, J. M. P. (2018). Classic infantile Pompe patients approaching adulthood: a cohort study on consequences for the brain. Developmental Medicine & Child Neurology, 60(6), 579–586. https://doi. org/10.1111/dmcn.13740. PMID: 29573408.

Elder, M. E., Nayak, S., Collins, S. W., Lawson, L. A., Kelley, J. S., Herzog, R. W., ... Byrne, B. J. (2013). B-cell depletion and immunomodulation before initiation of enzyme replacement therapy blocks the immune response to acid alpha-glucosidase in infantile-onset Pompe disease. Jornal de Pediatria, 163(3), 847–854. https://doi.org/10.1016/j.jpeds. 2013.03.002

van Gelder, C. M., Poelman, E., Plug, I., Hoogeveen-Westerveld, M., van der Beek, N., Reuser, A. J. J., & van der Ploeg, A. T. (2016). Effects of a higher dose of alglucosidase alfa on ventilator-free survival and motor outcome in classic infantile Pompe disease: An open-label singlecenter study. Journal of Inherited Metabolic Disease (Dordrecht), 39(3), 383–390.

Gungor, D., Kruijshaar, M. E., Plug, I., Rizopoulos, D., Kanters, T. A., Wens, S. C., & van der Ploeg, A. T. (2016). Quality of life and participation in daily life of adults with Pompe disease receiving enzyme replacement therapy: 10 years of international follow-up. Journal of Inherited Metabolic Disease (Dordrecht), 39(2), 253–260. https://doi.org/10. 1007/s10545-015-9889-6. [pii] 10.1007/s10545-015-9889-6.

Gungor, D., & Reuser, A. J. (2013). How to describe the clinical spectrum in Pompe disease? American journal of medical genetics. Part A, 161A(2), 399–400.

Gungor, D., de Vries, J. M., Brusse, E., Kruijshaar, M. E., Hop, W. C., Murawska, M., & van der Ploeg, A. T. (2013). Enzyme replacement therapy and fatigue in adults with Pompe disease. Molecular Genetics and Metabolism, 109(2), 174–178. https://doi.org/S1096-7192(13) 00105-4. [pii] 10.1016/j.ymgme.2013.03.016

Hahn, S. H., Kronn, D., Leslie, N. D., Pena, L. D. M., Tanpaiboon, P., Gambello, M. J., & Kishnani, P. S. (2018). Efficacy, safety profile, and immunogenicity of alglucosidase alfa produced at the 4,000-liter scale in US children and adolescents with Pompe disease: ADVANCE, a phase IV, open-label, prospective study. Genetics in Medicine, 20, 1284–1294. https://doi.org/gim20182. [pii] 10.1038/gim.2018.2

Reuser, A. J. J., Hirschhorn, R., & Kroos, M. A. (2018). Pompe disease: Glycogen storage disease type II, acid α-glucosidase (acid maltase) deficiency. In Beaudet, A. L., Vogelstein, B., Kinzler, K. W., Antonarakis, S. E., Ballabio, A., Gibson, K. M., & Mitchell, G. (Eds.), The Online Metabolic and Molecular Bases of Inherited Disease. Lysosomal Storage Disorders New York, NY: The McGraw-Hill Companies, Inc.

Van den Hout, H., Reuser, A. J., Vulto, A. G., Loonen, M. C., CrommeDijkhuis, A., & Van der Ploeg, A. T. (2000). Recombinant human alphaglucosidase from rabbit milk in Pompe patients. Lancet, 356(9227), 397–398. https://doi.org/S0140673600025332

Huie, M. L., Chen, A. S., Brooks, S. S., Grix, A., & Hirschhorn, R. (1994). A de novo 13 nt deletion, a newly identified C647W missense mutation and a deletion of exon 18 in infantile onset glycogen storage disease type II (GSDII). Human Molecular Genetics, 3(7), 1081–1087.

Kishnani, P. S., Corzo, D., Leslie, N. D., Gruskin, D., Van der Ploeg, A., Clancy, J. P., & Mandel, H. (2009). Early treatment with alglucosidase alpha prolongs long-term survival of infants with Pompe disease. Pediatric

Research, 66(3), 329–335. https://doi.org/10.1203/PDR. 0b013e3181b24e94

Kishnani, P. S., Goldenberg, P. C., DeArmey, S. L., Heller, J., Benjamin, D., Young, S., & Chen, Y. T. (2010). Cross-reactive immunologic material status affects treatment outcomes in Pompe disease infants. Molecular Genetics and Metabolism, 99(1), 26–33. https://doi.org/S10967192(09)00244-3. [pii] 10.1016/j.ymgme.2009.08.003

Kroos, M., Hoogeveen-Westerveld, M., Michelakakis, H., Pomponio, R., Van der Ploeg, A., Halley, D., & Consortium, G. A. A. D. (2012). Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. Human Mutation, 33(8).

Kroos, M., Pomponio, R. J., van Vliet, L., Palmer, R. E., Phipps, M., & Van der Helm, R. GA Database Consortium (2008). Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. Human Mutation, 29(6), E13–E26. https://doi.org/ 10.1002/humu.20745

Kroos, M. A., Mullaart, R. A., Van Vliet, L., Pomponio, R. J., Amartino, H., Kolodny, E. H., & Reuser, A. J. (2008). p.[G576S; E689K]: Pathogenic combination or polymorphism in Pompe disease? European Journal of Human Genetics, 16(8), 875–879. https://doi.org/10.1038/ejhg.2008. 34

Kuperus, E., Kruijshaar, M. E., Wens, S. C. A., de Vries, J. M., Favejee, M. M., van der Meijden, J. C., & van der Beek, N. (2017). Long-term benefit of enzyme replacement therapy in Pompe disease: A 5-year prospective study. Neurology, 89(23), 2365–2373. https://doi.org/ WNL.0000000000004711. [pii] 10.1212/WNL.0000000000004711

Labrijn-Marks, I., Somers-Bolman, G. M., In 't Groen, S. L. M., HoogeveenWesterveld, M., Kroos, M. A., Ala-Mello, S., & Halley, D. J. (2019). Segmental and total uniparental isodisomy (UPiD) as a disease mechanism in autosomal recessive lysosomal disorders: Evidence from SNP arrays. European Journal of Human Genetics, 27, 919–927. https://doi.org/10.1038/s41431-019-0348-y

Lim, K. H., Ferraris, L., Filloux, M. E., Raphael, B. J., & Fairbrother, W. G. (2011). Using positional distribution to identify splicing elements and predict pre-mRNA processing defects in human genes. Proceedings of the National Academy of Sciences of the United States of America, 108(27), 11093–11098. https://doi.org/10.1073/pnas.1101135108

Mendelsohn, N. J., Messinger, Y. H., Rosenberg, A. S., & Kishnani, P. S. (2009). Elimination of antibodies to recombinant enzyme in Pompe's disease. New England Journal of Medicine, 360(2), 194–195. https://doi. org/360/2/194. [pii] 10.1056/NEJMc0806809

Messinger, Y. H., Mendelsohn, N. J., Rhead, W., Dimmock, D., Hershkovitz, E., Champion, M., & Kishnani, P. S. (2012). Successful immune tolerance induction to enzyme replacement therapy in CRIM negative infantile Pompe disease. Genetics in Medicine, 14(1), 135–142. https:// doi.org/gim20114. [pii] 10.1038/gim.2011.4

Papadopoulos, C., Orlikowski, D., Prigent, H., Lacour, A., Tard, C., Furby, A., & French Pompe Study, G. (2017). Effect of enzyme replacement therapy with alglucosidase alfa (Myozyme(R)) in 12 patients with advanced late-onset Pompe disease. Molecular Genetics and Metabolism, 122(1-2), 80–85. https://doi.org/S1096-7192(17)30336-0. [pii] 10.1016/j.ymgme.2017.06.007

van der Ploeg, A. T., Clemens, P. R., Corzo, D., Escolar, D. M., Florence, J., Groeneveld, G. J., & Zivkovic, S. A. (2010). A randomized study of alglucosidase alfa in late-onset Pompe's disease. New England Journal of

Medicine, 362(15), 1396–1406. https://doi.org/362/15/1396

Poelman, E., Hoogeveen-Westerveld, M., Kroos-de Haan, M. A., van den Hout, J. M. P., Bronsema, K. J., van de Merbel, N. C., & Pijnappel, W. (2018). High sustained antibody titers in patients with classic infantile Pompe disease following immunomodulation at start of enzyme replacement therapy. Jornal de Pediatria, 195, 236–243.

Puzzo, F., Colella, P., Biferi, M. G., Bali, D., Paulk, N. K., Vidal, P., & Mingozzi, F. (2017). Rescue of Pompe disease in mice by AAVmediated liver delivery of secretable acid alpha-glucosidase. Science Translational Medicine, 9(418), eaam6375. https://doi.org/9/418/ eaam6375

Regnery, C., Kornblum, C., Hanisch, F., Vielhaber, S., Strigl-Pill, N., Grunert, B., & Schoser, B. (2012). 36 Months observational clinical study of 38 adult Pompe disease patients under alglucosidase alfa enzyme replacement therapy. Journal of Inherited Metabolic Disease (Dordrecht), 35(5), 837–845. https://doi.org/10.1007/s10545-0129451-8

Roig-Zamboni, V., Cobucci-Ponzano, B., Iacono, R., Ferrara, M. C., Germany, S., Bourne, Y. ... Sulzenbacher, G. (2017). Structure of human lysosomal acid α-glucosidase-a guide for the treatment of Pompe disease. Nature Communications, 8(1), 1111. https://doi.org/10. 1038/s41467-017-01263-3.

Schoser, B., Laforet, P., Kruijshaar, M. E., Toscano, A., van Doorn, P. A., van der Ploeg, A. T., & European Pompe, C. (2015). 208th ENMC International Workshop: Formation of a European Network to develop a European data sharing model and treatment guidelines for Pompe disease Naarden, The Netherlands, 26-28 September 2014. Neuromuscular Disorders: NMD, 25(8), 674–678. https://doi.org/ S0960-8966(15)00131-5

Shigeto, S., Katafuchi, T., Okada, Y., Nakamura, K., Endo, F., Okuyama, T., & Okumiya, T. (2011). Improved assay for differential diagnosis between Pompe disease and acid alpha-glucosidase pseudodeficiency on dried blood spots. Molecular Genetics and Metabolism, 103(1), 12–17. https:// doi.org/S1096-7192(11)00031-X

Soukarieh, O., Gaildrat, P., Hamieh, M., Drouet, A., Baert-Desurmont, S., Frebourg, T., & Martins, A. (2016). Exonic splicing mutations are more prevalent than currently estimated and can be predicted by using in silico tools. PLOS Genetics, 12(1), e1005756. https://doi.org/10.1371/ journal.pgen.1005756

Stepien, K. M., Whitby, J., Roberts, M., & Sharma, R. (2015). A 4-year observational clinical study of 24 adult pompe disease patients under alglucosidase alfa enzyme replacement therapy. Journal of Neuromuscular Diseases, S69(s1), 69.

Strothotte, S., Strigl-Pill, N., Grunert, B., Kornblum, C., Eger, K., Wessig, C., & Schoser, B. (2010). Enzyme replacement therapy with alglucosidase alfa in 44 patients with late-onset glycogen storage disease type 2: 12-Month results of an observational clinical trial. Journal of Neurology, 257(1), 91–97. https://doi.org/10.1007/s00415-0095275-3

Swallow, D. M., Kroos, M., Van der Ploeg, A. T., Griffiths, B., Islam, I., Marenah, C. B., & Reuser, A. J. (1989). An investigation of the properties and possible clinical significance of the lysosomal alpha-glucosidase GAA*2 allele. Annals of Human Genetics, 53(2), 177–184.

van Til, N. P., Stok, M., Aerts Kaya, F. S., de Waard, M. C., Farahbakhshian, E., Visser, T. P., & Wagemaker, G. (2010). Lentiviral gene therapy of murine hematopoietic stem cells ameliorates the Pompe disease phenotype. Blood, 115(26), 5329–5337. https://doi.org/blood-200911-252874

de Vries, J. M., van der Beek, N. A., Hop, W. C., Karstens, F. P., Wokke, J. H., de Visser, M., & van der Ploeg, A. T. (2012). Effect of enzyme therapy and prognostic factors in 69 adults with Pompe disease: An open-label single-center study. Orphanet Journal of Rare Diseases, 7(1), 73. https://doi.org/1750-1172-7-73. [pii] 10.1186/ 1750-1172-7-73

van der Wal, E., Bergsma, A. J., van Gestel, T. J. M., In 't Groen, S. L. M., Zaehres, H., Arauzo-Bravo, M. J., & Pijnappel, W. (2017). GAA Deficiency in Pompe disease is alleviated by exon inclusion in iPSCderived skeletal muscle cells. Molecular Therapy. Nucleic Acids, 7, 101–115. https://doi.org/S2162-2531(17)30140-3

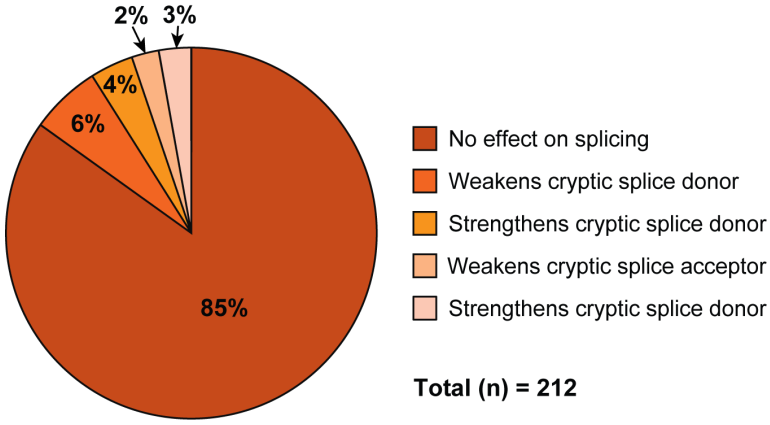
van der Wal, E., Bergsma, A. J., Pijnenburg, J. M., van der Ploeg, A. T., & Pijnappel, W. (2017). Antisense oligonucleotides promote exon inclusion and correct the common c.-32–13T>G GAA splicing variant in Pompe disease. Molecular Therapy. Nucleic Acids, 7, 90 –100. https:// doi.org/S2162-2531(17)30139-7

SUPPORTING INFORMATION

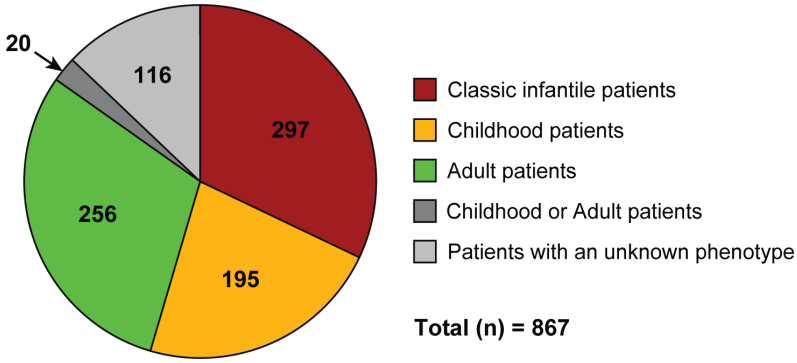
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Niño MY, in 't Groen SL, Bergsma AJ, et al. Extension of the Pompe mutation database by linking disease-associated variants to clinical severity. Human Mutation. 2019;40:1954–1967. https://doi.org/10.1002/humu.23854

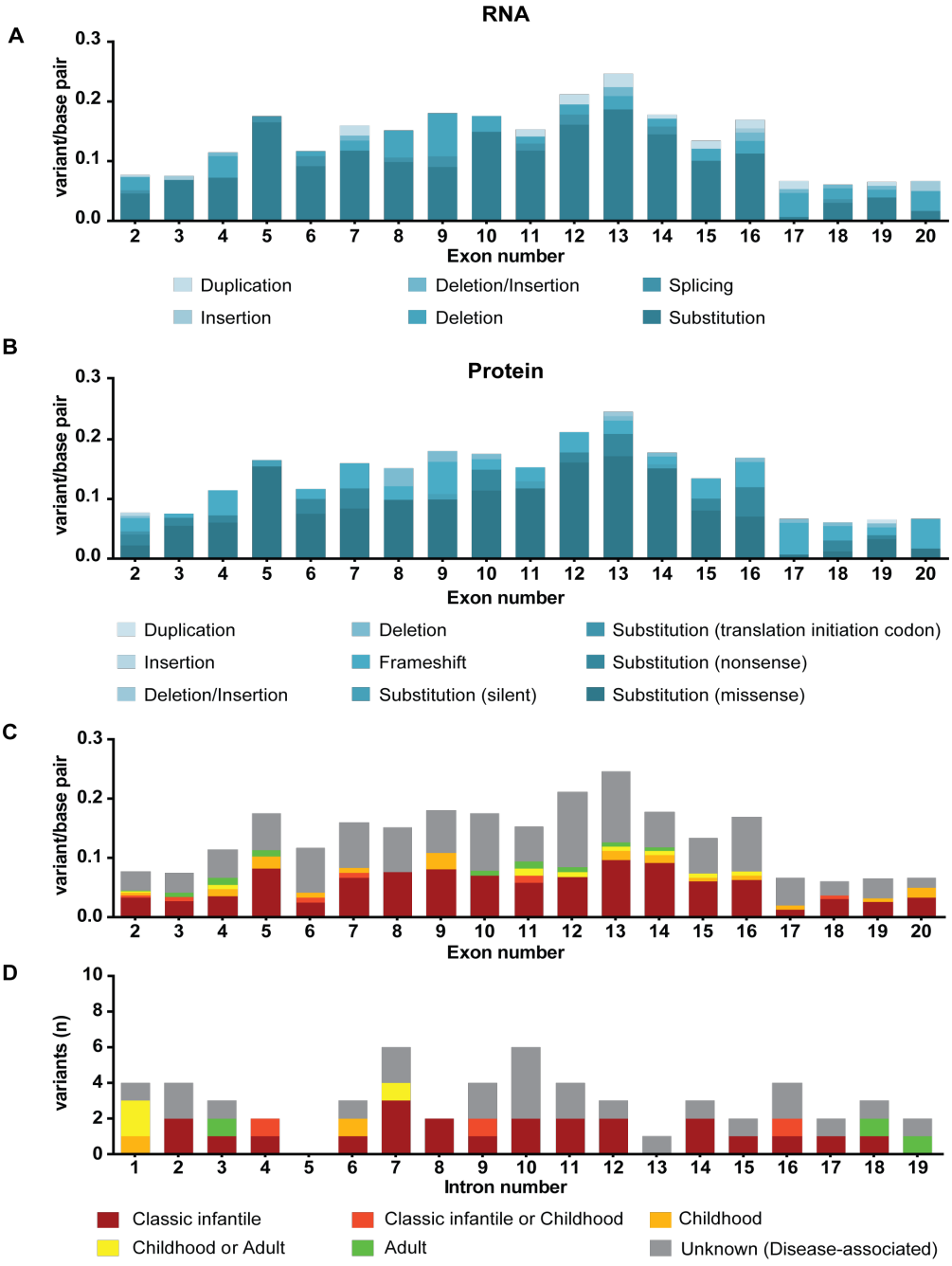
Supplementary data



Supplementary Figure S1: Splicing prediction of missense variants. Alamut® software was used, in which the cutoff for a predicted effect on splicing is met when two or more of the five prediction algorithms predict a difference in strength greater than 10%, as described previously (Bergsma, et al., 2016).



Supplementary Figure S2. Overview of the number of patients and their phenotypes used in this study.



Supplementary Figure S3: Distribution of variants in the *GAA* gene. (A) Frequencies of RNA variant types per exon. Numbers are corrected for the length of the exons. (B) As (A), but now at the level of protein. C) As (A) but now according to clinical severity rating. (D) Distribution of disease-associated variants within *GAA* intronic regions adjacent to the exons on the basis of clinical severity rating. Due to the lack of sequencing coverage of large introns, the frequency of variants within the *GAA* intronic regions is not corrected for intron size.

Supplementary Table S1, Algorithms used to predict impact on splicing

Prediction method	5' donor threshold	Range	3' acceptor threshold	Range	Website
Alamut 2.4.2	-	-	-	-	http:// http://www.interactive-biosoftware.com/doc/alamut-visual/2.4/splicing.html
Splice Site Finder-like	70	[0-100]	70	[0-100]	http://www.genet.sickkids.on.ca/~ali/splicesitefinder.html
Max Ent Scan	0	[0-12]	0	[0-16]	http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html
NNSPLICE	0,4	[0-1]	0,4	[0-1]	http://www.fruitfly.org/seq_tools/splice.html
Gene Splicer	0	[0-15]	0	[0-15]	http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml
Human Splicing Finder	65	[0-100]	65	[0-100]	http://www.umd.be/HSF/

Supplementary Table S2, List of *GAA* null alleles.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature
GAA and part of *CCDC40*	**Ch37/hg19:g.78,056,048_78,094,854delins14bp**	r.(-212_*551del)	p.(0)
exon 2	**c.18_25del**	r.(18_25del)	p.(Cys8Profs*24)
exon 2	**c.118C>T**	r.(118c>u)	p.(Arg40*)
exon 2	**c.136T>C**	r.(136u>c)	p.(Ser46Pro)
exon 2	**c.147_859-12del**	r.spl	p.?
exon 2	**c.172C>T**	r.(172c>u)	p.(Gln58*)
exon 2	**c.236_246del**	r.(236_246del)	p.(Pro79Argfs*13)
exon 2	**c.266G>A**	r.(266g>a)	p.(Arg89His)
exon 2	**c.307T>G**	r.(307u>g)	p.(Cys103Gly)
exon 2	**c.309C>A**	r.(309c>a)	p.(Cys103*)
exon 2	**c.340_341insT**	r.(340_341insu)	p.(Lys114Ilefs*32)
exon 2	**c.377G>A**	r.(377g>a)	p.(Trp126*)
exon 2	**c.378G>A**	r.(378g>a)	p.(Trp126*)
exon 2	**c.379_380del**	r.(379_380del)	p.(Cys127Leufs*18)
exon 2	**c.399C>A**	r.(399c>a)	p.(Tyr133*)
exon 2	**c.421C>A**	r.(421c>a)	p.(Leu141Met)
exon 2	**c.506T>C**	r.(506u>c)	p.(Leu169Pro)
exon 2	**c.525del**	r.(525del)	p.(Glu176Argfs*45)
exon 2	**c.525_526del**	r.(525_526del)	p.(Asn177Profs*11)
intron 2	**c.546+2T>C**	r.spl	p.?
intron 2	**c.546+2_546+5del**	r.spl	p.?
exon 3	**c.573C>A**	r.(573c>a)	p.(Tyr191*)
exon 3	**c.650C>T**	r.(650c>u)	p.(Pro217Leu)
exon 3	**c.655G>A**	r.(655g>a)	p.(Gly219Arg)
exon 3	**c.685_686insCGGC**	r.(685_686inscggc)	p.(Arg229Profs*102)
intron 3	**c.692+1G>A**	r.spl	p.?
exon 4	**c.716del**	r.(716del)	p.(Leu239Argfs*29)
exon 4	**c.722_723del**	r.(722_723del)	p.(Phe241Cysfs*88)
exon 4	**c.742del**	r.(742del)	p.(Leu248Profs*20)
exon 4	**c.784G>A**	r.(784g>a)	p.(Glu262Lys)
exon 4	**c.827_845del**	r.(827_845del)	p.(Ile276Thrfs*32)
exon 4	**c.829_851del**	r.(829_851del)	p.(Thr277Alafs*45)
intron 4	**c.859-2A>T**	r.spl	p.?
exon 5	**c.872T>C**	r.(872u>c)	p.(Leu291Pro)
exon 5	**c.877G>A**	r.(877g>a)	p.(Gly293Arg)
exon 5	**c.896T>G**	r.(896u>g)	p.(Leu299Arg)
exon 5	**c.896T>C**	r.(896u>c)	p.(Leu299Pro)
exon 5	**c.923A>C**	r.(923a>c)	p.(His308Pro)
exon 5	**c.925G>A**	r.(925g>a)	p.(Gly309Arg)
exon 5	**c.935T>G**	r.(935u>g)	p.(Leu312Arg)
exon 5	**c.953T>C**	r.(953u>c)	p.(Met318Thr)
exon 6	**c.988T>G**	r.(988u>g)	p.(Trp330Gly)
exon 6	**c.998C>A**	r.(998c>a)	p.(Thr333Lys)
exon 6	**c.1062C>G**	r.(1062c>g)	p.(Tyr354*)
intron 6	**c.1076-1G>C**	r.spl	p.?
exon 7	**c.1099T>C**	r.(1099u>c)	p.(Trp367Arg)
exon 7	**c.1101G>A**	r.(1101g>a)	p.(Trp367*)
exon 7	**c.1115A>T**	r.(1115a>u)	p.(His372Leu)

Supplementary Table S2 continued.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature
exon 7	**c.1120T>C**	r.(1120u>c)	p.(Cys374Arg)
exon 7	**c.1124G>T**	r.(1124g>u)	p.(Arg375Leu)
exon 7	**c.1128_1129delinsC**	r.(1128_1129delinsc)	p.(Trp376Cysfs*16)
exon 7	**c.1129G>C**	r.(1129g>c)	p.(Gly377Arg)
exon 7	**c.1157dup**	r.(1157dup)	p.(Val387Glyfs*119)
intron 7	**c.1194+2T>C**	r.spl	p.?
intron 7	**c.1195-19_2190-20del**	r.spl	p.?
intron 7	**c.1195-2A>G**	r.spl	p.?
exon 8	**c.1199_1210del**	r.(1199_1210del)	p.(Val400_Asn403del)
exon 8	**c.1202A>G**	r.(1202a>g)	p.(Gln401Arg)
exon 8	**c.1209C>G**	r.(1209c>g)	p.(Asn403Lys)
exon 8	**c.1209del**	r.(1209del)	p.(Asn403Lysfs*37)
exon 8	**c.1210G>A**	r.(1210g>a)	p.(Asp404Asn)
exon 8	**c.1211A>G**	r.(1211a>g)	p.(Asp404Gly)
exon 8	**c.1214T>C**	r.(1214u>c)	p.(Leu405Pro)
exon 8	**c.1222A>G**	r.(1222a>g)	p.(Met408Val)
exon 8	**c.1316T>A**	r.(1316u>a)	p.(Met439Lys)
exon 8	**c.1322_1326+9del**	r.spl	p.?
intron 8	**c.1326+1G>A**	r.spl	p.?
intron 8	**c.1327-2A>G**	r.spl	p.?
exon 9	**c.1356del**	r.(1356del)	p.(Ser454Alafs*23)
exon 9	**c.1377_1379del**	r.(1377_1379del)	p.(Asp459del)
exon 9	**c.1385T>C**	r.(1385u>c)	p.(Leu462Pro)
exon 9	**c.1396del**	r.(1396del)	p.(Val466Phefs*11)
exon 9	**c.1396G>T**	r.(1396g>u)	p.(Val466Phe)
exon 9	**c.1408_1410del**	r.(1408_1410del)	p.(Asn470del)
exon 9	**c.1411_1414del**	r.(1411_1414del)	p.(Glu471Profs*5)
exon 9	**c.1432G>A**	r.(1432g>a)	p.(Gly478Arg)
exon 9	**c.1437G>C**	r.(1437g>c), r.(spl?)	p.(Lys479Asn), p.(?)
intron 9	**c.1437+2T>C**	r.1327_1437del	p.(Asp443_Lys479del)
exon 10	**c.1441del**	r.(1441del)	p.(Trp481Glyfs*39)
exon 10	**c.1441T>C**	r.(1441u>c)	p.(Trp481Arg)
exon 10	**c.1456_1468del**	r.(1456_1468del)	p.(Ala486Serfs*30)
exon 10	**c.1465G>A**	r.(1465g>a)	p.(Asp489Asn)
exon 10	**c.1465G>T**	r.(1465g>u)	p.(Asp489Tyr)
exon 10	**c.1509_1511del**	r.(1509_1511del)	p.(Ala504del)
exon 10	**c.1540G>C**	r.(1540g>c)	p.(Gly514Arg)
exon 10	**c.1548G>A**	r.(1548g>a)	p.(Trp516*)
intron 10	**c.1551+1G>C**	r.spl	p.?
intron 10	**c.1551+2T>G**	r.spl	p.?
exon 11	**c.1555A>G**	r.(1555a>g)	p.(Met519Val)
exon 11	**c.1556T>C**	r.(1556u>c)	p.(Met519Thr)
exon 11	**c.1561G>C**	r.(1561g>c)	p.(Glu521Gln)
exon 11	**c.1564C>G**	r.(1564c>g)	p.(Pro522Ala)
exon 11	**c.1582_1583del**	r.(1582_1583del)	p.(Gly528Leufs*2)
intron 11	**c.1636+5G>C**	r.1636_1637ins957	p.(Gly546fs*145)
intron 11	**c.1637-2A>G**	r.spl	p.?
exon 12	**c.1650dup**	r.(1650dup)	p.(Thr551Aspfs*85)
exon 12	**c.1655T>C**	r.(1655u>c)	p.(Leu552Pro)
exon 12	**c.1669A>T**	r.(1669a>u)	p.(Ile557Phe)

Supplementary Table S2 continued.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature
exon 12	**c.1687C>T**	r.(1687c>u)	p.(Gln563*)
exon 12	**c.1696T>C**	r.(1696u>c)	p.(Ser566Pro)
exon 12	**c.1705dup**	r.(1705dup)	p.(tyr569Leufs*67)
exon 12	**c.1710C>G**	r.(1710c>g)	p.(Asn570Lys)
exon 12	**c.1735G>A**	r.(1735g>a)	p.(Glu579Lys)
intron 12	**c.1754+1G>A**	r.spl	p.?
intron 12	**c.1755-1G>A**	r.spl	p.?
exon 13	**c.1760T>C**	r.(1760u>c)	p.(Leu587Pro)
exon 13	**c.1798C>T**	r.(1798c>u)	p.(Arg600Cys)
exon 13	**c.1799G>A**	r.(1799g>a)	p.(Arg600His)
exon 13	**c.1799G>T**	r.(1799g>u)	p.(Arg600Leu)
exon 13	**c.1802C>A**	r.(1802c>a)	p.(Ser601*)
exon 13	**c.1802C>T**	r.(1802c>u)	p.(Ser601Leu)
exon 13	**c.1820G>A**	r.(1820g>a)	p.(Gly607Asp)
exon 13	**c.1824_1828dup**	r.(1824_1828dup)	p.(Ala610Aspfs*88)
exon 13	**c.1826dup**	r.(1826dup)	p.(Tyr609*)
exon 13	**c.1827C>G**	r.(1827c>g)	p.(Tyr609*)
exon 13	**c.1833_1847delinsACGGGGTAT**	c.(1833_1847delinsacgggguau)	p.(His612_ Asp616delinsArgGlylle)
exon 13	**c.1834C>T**	r.(1834c>u)	p.(His612Tyr)
exon 13	**c.1843G>A**	r.(1843g>a)	p.(Gly615Arg)
exon 14	**c.1905C>A**	r.(1905c>a)	p.(Asn635Lys)
exon 14	**c.1912G>T**	r.(1912g>u)	p.(Gly638Trp)
exon 14	**c.1913G>T**	r.(1913g>u)	p.(Gly638Val)
exon 14	**c.1924G>T**	r.(1924g>u)	p.(Val642Phe)
exon 14	**c.1927G>A**	r.(1927g>a)	p.(Gly643Arg)
exon 14	**c.1930_1936dup**	r.(1930_1936dup)	p.(Val646Glyfs*93)
exon 14	**c.1933G>C**	r.(1933g>c)	p.(Asp645His)
exon 14	**c.1933G>A**	r.(1933g>a)	p.(Asp645Asn)
exon 14	**c.1935C>A**	r.(1935c>a)	p.(Asp645Glu)
exon 14	**c.1941C>G**	r.(1941c>g)	p.(Cys647Trp)
exon 14	**c.1960T>C**	r.(1960u>c)	p.(Ser654Pro)
exon 14	**c.1987del**	r.(1987del)	p.(Gln663Serfs*33)
exon 14	**c.2012T>G**	r.(2012u>g)	p.(Met671Arg)
exon 14	**c.2024_2026del**	r.(2024_2026del)	p.(Asn675del)
intron 14	**c.2040+1G>T**	r.spl	p.?
intron 14	**c.2041-2A>C**	r.spl	p.?
exon 15	**c.2045A>G**	r.(2045a>g)	p.(Gln682Arg)
exon 15	**c.2078dup**	r.(2078dup)	p.(Ala694Glyfs*43)
exon 15	**c.2104C>T**	r.(2104c>u)	p.(Arg702Cys)
exon 15	**c.2105G>A**	r.(2105g>a)	p.(Arg702His)
exon 15	**c.2105G>T**	r.(2105g>u)	p.(Arg702Leu)
exon 15	**c.2171C>A**	r.(2171c>a)	p.(Ala724Asp)
exon 15	**c.2174G>C**	r.(2174g>c)	p.(Arg725Pro)
exon 15	**c.2185del**	r.(2185del)	p.(Leu729Trpfs*35)
exon 15	**c.2188G>T**	r.(2188g>u)	p.(Glu730*)
intron 15	**c.2189+459_3405del**	r.spl	p.?
exon 16	**c.2210C>A**	r.(2210c>a)	p.(Thr737Asn)
exon 16	**c.2236T>C**	r.(2236u>c)	p.(Trp746Arg)
exon 16	**c.2237G>A**	r.(2237g>a)	p.(Trp746*)

Supplementary Table S2 continued.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature
exon 16	**c.2238G>A**	r.(2238g>a)	p.(Trp746*)
exon 16	**c.2242dup**	r.(2242dup)	p.(Glu748Glyfs*48)
exon 16	**c.2269C>T**	r.(2269c>u)	p.(Gln757*)
exon 16	**c.2274dup**	r.(2274dup)	p.(Gly759Argfs*37)
exon 16	**c.2303C>G**	r.(2303c>g)	p.(Pro768Arg)
exon 16	**c.2303C>T**	r.(2303c>u)	p.(Pro768Leu)
intron 16	**c.2331+2T>C**	r.2316_2331del	p.(Tyr773fs*3)
exon 17	**c.2380del**	r.(2380del)	p.(Arg794fs*12)
exon 17	**c.2432del**	r.(2432del)	p.(Leu811fs*36)
intron 17	**c.2481+102_2646+31del**	r.2482_2646del	p.(Gly828_Asn882del)
exon 18	**c.2501_2502del**	r.(2501_2502del)	p.(Thr834Argfs*49)
exon 18	**c.2512C>T**	r.(2512c>u)	p.(Gln838*)
exon 18	**c.2560C>T**	r.(2560c>u)	p.(Arg854*)
exon 18	**c.2608C>T**	r.(2608c>u)	p.(Arg870*)
exon 18	**c.2639C>A**	r.(2639c>a)	p.(Ala880Asp)
intron 18	**c.2646+2T>A**	r.spl	p.?
exon 19	**c.2662G>T**	r.(2662g>u)	p.(Glu888*)
exon 19	**c.2707_2709del**	r.(2707_2709del)	p.(Lys903del)
exon 19	**c.2741delinsGAC**	r.(2741delinsgac)	p.(Gln914fs*30)
exon 19	**c.2758_2775dup**	r.(2758_2775dup)	p.(ly920_Asn925dup)
exon 20	**c.2815_2816del**	r.(2815_2816del)	p.(Val939Leufs*78)
exon 20	**c.2841dup**	r.(2841dup)	p.(Leu948Serfs*70)

Supplementary Table S3. List of new variants added to the Pompe disease *GAA* variant database.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature	Predicted severity	Phenotype with a null allele	CRIM status
exon 2	c.322T>G	r.(322u>g)	p.(Cys108Gly)	Potentially less severe	Unknown	Unknown
exon 2	c.364A>G	r.(364a>g)	p.(Met122Val)	Less severe	Unknown (disease-associated)	Unknown
exon 7	c.1101G>A	r.(1101g>a)	p.(Trp367*)	Very severe	Classic infantile	Negative
exon 8	c.1322_1326+9del	r.spl	p.?	Very severe	Classic infantile	Negative
exon 10	c.1468T>C	r.(1468u>c)	p.(Phe490Leu)	Less severe	Unknown	Positive
exon 11	c.1582_1583del	r.(1582_1583del)	p.(Gly528Leufs*2)	Very severe	Classic infantile	Negative
exon 13	c.1796C>T	r.(1796c>a)	p.(Ser599Tyr)	Less severe	Unknown (disease-associated)	Unknown
exon 18	c.2512C>T	r.(2512c>u)	p.(Gln838*)	Very severe	Classic infantile	Negative

Supplementary Table S4. List of variants with the clinical severity rating "Unknown".

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature	Minor Allele frequency
exon 1, 5' UTR	**c.-82G>C**	r.(-82g>c)	p.?	MAF not reported
exon 2	**c.1A>T**	r.(1a>u)	p.(0)	MAF not reported
exon 2	**c.32G>A**	r.(32g>a)	p.(Arg11Gln)	MAF is less than 1%
exon 2	**c.54C>T**	r.(54c>u)	p.(=)	MAF not reported
exon 2	**c.186_196dup**	r.(186_196dup)	p.(Arg66Hisfs*80)	MAF not reported
exon 2	**c.199G>A**	r.(199g>a)	p.(Asp67Asn)	MAF not reported
exon 2	**c.221G>A**	r.(221g>a)	p.(Arg74His)	MAF not reported
exon 2	**c.271G>A**	r.271g>a	p.(Asp91Asn)	MAF is over 1%
exon 2	**c.307T>C**	r.(307u>c)	p.(Cys103Arg)	MAF not reported
exon 2	**c.322T>G**	r.(322u>g)	p.(Cys108Gly)	MAF not reported
exon 2	**c.324T>C**	r.324u>c	p.(=)	MAF is over 5%
exon 2	**c.363G>A**	r.(363g>a)	p.(=)	MAF not reported
exon 2	**c.447G>A**	r.(447g>a)	p.(=)	MAF is over 1%
exon 2	**c.483dup**	r.(483dup)	p.(Lys162Glnfs*15)	MAF not reported
exon 2	**c.510C>T**	r.(510c>u)	p.(=)	MAF is less than 1%
exon 2	**c.532C>T**	r.(532c>u)	p.(Arg178Cys)	MAF is less than 1%
intron 2	**c.546+24G>A**	r.(=)	p.?	MAF is less than 1%
intron 2	**c.546+45G>C**	r.(=)	p.?	MAF not reported
intron 2	**c.547-4C>G**	r.(=)	p.?	MAF is over 5%
exon 3	**c.596A>G**	r.(596a>g)	p.(His199Arg)	MAF is less than 1%
exon 3	**c.642C>T**	r.642c>u	p.(=)	MAF is over 5%
exon 3	**c.658G>T**	r.(658g>u)	p.(Val220Leu)	MAF is less than 1%
exon 3	**c.664G>A**	r.(664g>a)	p.(Val222Met)	MAF is less than 1%
exon 3	**c.668G>A**	r.668g>a	p.(Arg223His)	MAF is over 5%
intron 3	**c.693-49C>T**	r.(=)	p.?	MAF is over 5%
exon 4	**c.701C>G**	r.(701c>g)	p.(Thr234Arg)	MAF not reported
exon 4	**c.725C>T**	r.(725c>u)	p.(Ala242Val)	MAF is less than 1%
exon 4	**c.743T>C**	r.(743u>c)	p.(Leu248Pro)	MAF not reported
exon 4	**c.768dup**	r.(768dup)	p.(Ile257Tyrfs*73)	MAF not reported
exon 4	**c.776G>T**	r.(776g>u)	p.(Gly259Val)	MAF not reported
exon 4	**c.852G>A**	r.(852g>a)	p.(=)	MAF is over 1%
intron 4	**c.858+5_858+6ins7**	r.(spl?)	p.?	MAF not reported
intron 4	**c.858+6G>A**	r.(=)	p.?	MAF not reported
intron 4	**c.858+17_858+23del**	r.(=)	p.?	MAF not reported
intron 4	**c.858+20dup**	r.(=)	p.?	MAF not reported
intron 4	**c.858+21C>G**	r.(=)	p.?	MAF not reported
intron 4	**c.858+17_858+23dup**	r.(=)	p.?	MAF not reported
intron 4	**c.858+30T>C**	r.(=)	p.?	MAF is over 5%
intron 4	**c.858+37C>T**	r.(=)	p.?	MAF not reported
exon 5	**c.868A>G**	r.(868a>g)	p.(Asn290Asp)	MAF is less than 1%
exon 5	**c.872T>A**	r.(872u>a)	p.(Leu291His)	MAF not reported
exon 5	**c.915G>A**	r.(915g>a)	p.(=)	MAF is less than 1%
exon 5	**c.921A>T**	r.921a>u	p.(=)	MAF is over 5%
exon 5	**c.929T>G**	r.(929u>g)	p.(Val310Gly)	MAF is less than 1%
exon 5	**c.947A>T**	r.(947a>u)	p.(Asn316Ile)	MAF not reported
intron 5	**c.955+2T>G**	r.spl	p.?	MAF not reported
intron 5	**c.955+12G>A**	r.(=)	p.?	MAF is over 5%
intron 5	**c.956-84C>T**	r.(=)	p.?	MAF is over 5%
intron 6	**c.1075+13C>T**	r.(=)	p.?	MAF is over 1%

Supplementary Table S4 continued.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature	Minor Allele frequency
exon 7	**c.1134C>G**	r.(1134c>g)	p.(Tyr378*)	MAF not reported
exon 7	**c.1190C>T**	r.(1190c>u)	p.(Pro397Leu)	MAF is less than 1%
intron 7	**c.1194+5G>A**	r.(spl?)	p.?	MAF not reported
intron 7	**c.1195-44C>T**	r.(=)	p.?	MAF is less than 1%
exon 8	**c.1203G>A**	r.1203g>a	p.(=)	MAF is over 5%
exon 8	**c.1204T>C**	r.(1204u>c)	p.(Trp402Arg)	MAF not reported
exon 8	**c.1229C>T**	r.(1229c>u)	p.(Ser410Phe)	MAF is less than 1%
exon 8	**c.1286A>G**	r.(1286a>g)	p.(Gln429Arg)	MAF is over 1%
intron 8	**c.1326+5G>A**	r.(spl?)	p.?	MAF not reported
intron 8	**c.1327-18A>G**	r.(=)	p.?	MAF is over 5%
exon 9	**c.1370C>A**	r.(1370c>a)	p.(Pro457His)	MAF not reported
exon 9	**c.1373A>G**	r.(1373a>g)	p.(Tyr458Cys)	MAF not reported
exon 9	**c.1374C>T**	r.1374c>u	p.(=)	MAF is over 5%
exon 9	**c.1381G>A**	r.(1381g>a)	p.(Gly461Ser)	MAF not reported
intron 9	**c.1437+4G>C**	r.(spl?)	p.?	MAF not reported
intron 9	**c.1438-19G>C**	r.(=)	p.?	MAF is over 5%
intron 9	**c.1438-1G>T**	r.spl	p.?	MAF is less than 1%
exon 10	**c.1445C>G**	r.(1445c>g)	p.(Pro482Arg)	MAF not reported
exon 10	**c.1448G>T**	r.(1448g>u)	p.(Gly483Val)	MAF not reported
exon 10	**c.1468T>C**	r.(1468u>c)	p.(Phe490Leu)	MAF not reported
intron 10	**c.1551+42G>A**	r.(=)	p.?	MAF is over 1%
intron 10	**c.1551+49C>A**	r.(=)	p.?	MAF is over 5%
intron 10	**c.1551+49C>T**	r.(=)	p.?	MAF not reported
exon 11	**c.1564C>T**	r.(1564c>u)	p.(Pro522Ser)	MAF not reported
exon 11	**c.1564C>T**	r.(1564c>u)	p.(Pro522Ser)	MAF not reported
exon 11	**c.1568C>A**	r.(1568c>a)	p.(Ser523Tyr)	MAF not reported
exon 11	**c.1581G>A**	r.1581g>a	p.(=)	MAF is over 5%
exon 12	**c.1672T>A**	r.(1672u>a)	p.(Cys558Ser)	MAF not reported
exon 12	**c.1717A>C**	r.(1717a>c)	p.(Asn573His)	MAF not reported
exon 12	**c.1724A>G**	r.(1724a>g)	p.(Tyr575Cys)	MAF not reported
exon 12	**c.1726G>A**	r.1726g>a	p.(Gly576Ser)	MAF is over 5%
exon 12	**c.1727G>A**	r.(1727g>a)	p.(Gly576Asp)	MAF not reported
intron 12	**c.1754+12G>A**	r.(=)	p.?	MAF is over 5%
intron 12	**c.1754+16C>T**	r.(=)	p.?	MAF is less than 1%
exon 13	**c.1804A>G**	r.(1804a>g)	p.(Thr602Ala)	MAF is less than 1%
exon 13	**c.1830C>T**	r.(1830c>u)	p.(=)	MAF is less than 1%
exon 13	**c.1850T>C**	r.(1850u>c)	p.(Val617Ala)	MAF not reported
exon 13	**c.1872C>T**	r.(1872c>u)	p.(=)	MAF is less than 1%
exon 13	**c.1879T>C**	r.(1879u>c)	p.(Ser627Pro)	MAF not reported
exon 13	**c.1886C>T**	r.(1886c>u)	p.(Pro629Leu)	MAF is less than 1%
intron 13	**c.1888+21G>A**	r.(=)	p.?	MAF is over 5%
intron 13	**c.1889-27_2040+23del**	r.spl	p.?	MAF not reported
exon 14	**c.1917G>A**	r.(1917g>a)	p.(=)	MAF not reported
exon 14	**c.1920T>G**	r.(1920u>g)	p.(=)	MAF is less than 1%
exon 14	**c.1923G>A**	r.(1923g>a)	p.(=)	MAF not reported
exon 14	**c.1951_1952delinsT**	r.(1951_1952delinsu)	p.(Gly651Serfs*45)	MAF not reported
exon 14	**c.1962_1964del**	r.(1962_1964del)	p.(Glu656del)	MAF not reported
exon 14	**c.1971G>A**	r.(1971g>a)	p.(=)	MAF not reported
exon 14	**c.1981T>G**	r.(1981u>g)	p.(Trp661Gly)	MAF not reported

Supplementary Table S4 continued.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature	Minor Allele frequency
intron 14	**c.2040+12G>A**	r.(=)	p.?	MAF not reported
intron 14	**c.2040+20A>G**	r.(=)	p.?	MAF is over 5%
intron 14	**c.2040+20A>T**	r.(=)	p.?	MAF not reported
intron 14	**c.2040+22G>T**	r.(=)	p.?	MAF not reported
intron 14	**c.2041-64G>A**	r.(=)	p.?	MAF is over 5%
intron 14	**c.2041-61del**	r.(=)	p.?	MAF not reported
exon 15	**c.2061C>T**	r.(2061c>u)	p.(=)	MAF not reported
exon 15	**c.2065G>A**	r.2065g>a	p.(Glu689Lys)	MAF is over 5%
exon 15	**c.2097_2102del**	r.(2097_2102del)	p.(Thr700_Leu701del)	MAF not reported
exon 15	**c.2133A>G**	r.2133a>g	p.(=)	MAF is over 5%
exon 16	**c.2152G>A**	r.(2152g>a)	p.(Val718Ile)	MAF is less than 1%
exon 15	**c.2154C>T**	r.(2154c>u)	p.(=)	MAF is less than 1%
intron 15	**c.2189+1G>A**	r.spl	p.?	MAF not reported
intron 15	**c.2190-53C>G**	r.(=)	p.?	MAF is less than 1%
exon 16	**c.2227C>A**	r.(2227c>a)	p.(Gln743Lys)	MAF not reported
exon 16	**c.2236T>G**	r.(2236u>g)	p.(Trp746Gly)	MAF not reported
exon 16	**c.2242G>T**	r.(2242g>u)	p.(Glu748*)	MAF not reported
exon 16	**c.2284G>A**	r.(2284g>a)	p.(Glu762Lys)	MAF is less than 1%
intron 16	**c.2331+20G>A**	r.(=)	p.?	MAF is over 5%
intron 16	**c.2331+24T>C**	r.(=)	p.?	MAF is over 5%
exon 17	**c.2338G>A**	r.2338g>a	p.(Val780Ile)	MAF is over 5%
exon 17	**c.2357dup**	r.(2357dup)	p.(Pro788Thrfs*8)	MAF not reported
exon 17	**c.2395C>G**	r.(2395c>g)	p.(His799Asp)	MAF not reported
exon 17	**c.2446G>A**	r.2446g>a	p.(Val816Ile)	MAF is over 5%
intron 17	**c.2481+1G>A**	r.spl	p.?	MAF not reported
exon 18	**c.2495_2496del**	r.(2495_2496del)	p.(Thr832Asnfs*51)	MAF not reported
exon 18	**c.2553G>A**	r.2553g>a	p.(=)	MAF is over 5%
intron 18	**c.2646+39G>A**	r.(=)	p.?	MAF is less than 1%
exon 19	**c.2724C>G**	r.(2724c>g)	p.(=)	MAF not reported
exon 19	**c.2770T>C**	r.(2770u>c)	p.(Ser924Pro)	MAF not reported
exon 19	**c.2780C>T**	r.2780c>u	p.(Thr927Ile)	MAF is over 5%
intron 19	**c.2800-60G>A**	r.(=)	p.?	MAF is less than 1%
exon 20	**c.2804T>C**	r.(2804u>c)	p.(Leu935Pro)	MAF not reported
exon 20	**c.2808C>T**	r.(2808c>u)	p.(=)	MAF not reported
exon 20, 3′ UTR	**c.*3G>A**	r.(*3g>a)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*16T>A**	r.(*16u>a)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*91G>A**	r.(*91g>a)	p.?	MAF is over 5%
exon 20, 3′ UTR	**c.*140del**	r.(*140del)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*143C>T**	r.(*143c>u)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*154_*155insG**	r.(*154_*155insg)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*223C>T**	r.(*223c>u)	p.?	MAF is over 5%
exon 20, 3′ UTR	**c.*227G>C**	r.(*227g>c)	p.?	MAF not reported
exon 20, 3′ UTR	**c.*418T>C**	r.(*418u>c)	p.?	MAF not reported

Supplementary Table S5. List of *GAA* variants that lower GAA activity but do not cause Pompe disease.

Location	DNA nomenclature	RNA nomenclature	Protein nomenclature	Predicted severity	Phenotype with a null allele	CRIM status	Common name	MAF
exon 2	c.271G>A	r.271g>a	p.(Asp91Asn)	Presumably non-pathogenic	Unknown	Positive	*GAA2*	2.0%
exon 12	c.1726G>A	r.1726g>a	p.(Gly576Ser)	Presumably non-pathogenic	Unknown	Unknown	Asian pseudo-deficiency allele	1.8%
exon 15	c.2065G>A	r.2065g>a	p.(Glu689Lys)	Non-pathogenic	Non-pathogenic	Positive	Asian pseudo-deficiency allele	8.8%

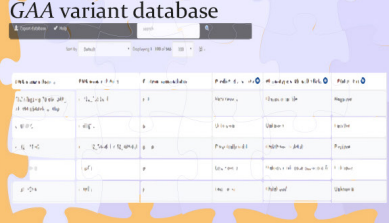
Supplementary Table S6. Definition of groups outlined in Figure 2.

	Type of variant (RNA)
	Substitution
	Splicing (intron variant)
	Splicing (splice donor site)
	Splicing (splice acceptor site)
	Deletion
	Deletion/Insertion
	Duplication
	Others:
	• Deletion (intron variant)
	• Deletion, Splicing (intron variant)
	• Deletion, Splicing (splice donor site)
	• Insertions
	• Splicing (splice acceptor site)
	• Substitution (intron variant)
	• Substitution, Splicing (exon variant)

	Type of variant (protein)
	Substitution (missense)
	Substitution (nonsense)
	Deletion
	Deletion (translation initiation site)
	Frameshift
	Others:
	• Deletion/Insertion
	• Duplication
	• Insertion
	• Substitution (missense), Deletion
	• Substitution (missense), Frameshift
	• Substitution (silent)
	• Substitution (silent), Insertion
	• Substitution (translation initiation codon)
	No category: includes all variants located in an intron or the UTR's without experimental data on pre-mRNA splicing

Pompe disease

GAA variant database



CHAPTER 4

THE EXTENDED POMPE MUTATION DATABASE REVEALS DISTINCT PHENOTYPIC SPECTRA OF POMPE DISEASE-ASSOCIATED *GAA* VARIANTS

Monica Y. Niño[1,2,3,#], Stijn L.M. in 't Groen[1,2,3,#], Atze J. Bergsma[1,2,3], Marianne Hoogeveen-Westerveld[2], Ans T. van der Ploeg[1,3] and W.W.M. Pim Pijnappel[1,2,3,*]

[1]Department of Pediatrics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

[2]Department of Clinical Genetics, Erasmus University Medical Center, 3015 GE Rotterdam, Netherlands

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

[#]equal contribution

*Correspondence: w.pijnappel@erasmusmc.nl

Submitted

4

The extended Pompe mutation database reveals distinct phenotypic spectra of Pompe disease-associated *GAA* variants

Monica Y. Niño, Stijn L.M. in 't Groen, Atze J. Bergsma, Marianne Hoogeveen-Westerveld, Ans T. van der Ploeg and W.W.M. Pim Pijnappel

ABSTRACT

Pompe disease is an autosomal recessive lysosomal storage disorder caused by variants in the acid alpha glucosidase (*GAA*) gene. Symptom onset depends on the *GAA* variants, and is influenced by putative disease-modifying factors. Phenotypic spectra of individual *GAA* variants remain largely unexplored. Here, we used the Pompe mutation database to compare phenotypic variation for common *GAA* genotypes. In patients with compound heterozygous genotypes, in which the second *GAA* alelle was fully deleterious, the childhood *GAA* variants c.503G>A, c.546G>T, and c.796C>T were associated with symptom onset between 1 and 17 years of age. For patients with the childhood or adult variants c.2238G>C, c.-32-2A>G, and c.2173C>T, onset ranged between 2 and 24 years of age. The adult variant c.2647-7G>A was associated with symptom onset between 35 and 53 years of age. The c.-32-13T>G (IVS1) variant was associated with an exceptionally broad onset between 0 and 67 years of age. Patients with homozygous *GAA* genotypes showed similar disease spectra but with later onsets compared to compound heterozygous patients. These results improve clinical prognosis of the progression of Pompe disease, and suggest that the influence of putative modifying factors is particularly large for the c.-32-13T>G variant as compared to other *GAA* variants.

Keywords: Pompe disease, www.pompecenter.nl; mutation database, genotype-phenotype relationship, GAA deficiency, Glycogenosis type II

INTRODUCTION

Pompe disease is a monogenic lysosomal storage disorder caused by disease associated variants in the acid α-glucosidase (*GAA*) gene (van der Ploeg and Reuser, 2008). A large number of disease associated *GAA* variants have been identified. These are listed at the open source website www.pompecenter.nl (Kroos, et al., 2012a; Kroos, et al., 2008) and are described in the accompanying manuscript (Niño, et al., 2019). Two disease associated *GAA* variants, one per allele, are required to cause Pompe disease. The clinical spectrum of Pompe disease ranges from a classic infantile form to an adult onset form (Gaeta, et al., 2015; Gungor and Reuser, 2013; Herzog, et al., 2012; Hirschhorn and Reuser, 2001; Kroos, et al., 2007; Laforet, et al., 2013; Schoser, et al., 2015; Van der Beek, et al., 2009; van der Ploeg and Reuser, 2008). The most severe classic infantile form of Pompe disease is caused by two very severe disease associated *GAA* variants that completely abrogate GAA enzymatic activity. Classic infantile patients have generalized skeletal muscle weakness and hypertrophic cardiomyopathy at birth and die without therapy within the first year of life. Patients with some residual GAA enzyme activity, due to the presence of at least one milder disease associated *GAA* variant, have variable onset of symptoms. These patients develop at some point in their life skeletal muscle weakness leading to mobility problems, and decreased pulmonary function resulting in wheelchair and/or ventilator dependency. Development of hypertrophic cardiomyopathy is rare. Symptoms including scoliosis/ kyphoscoliosis, scapular winging, ptosis, and cerebral vessel anomalies may variably occur.

While there are many different disease-associated *GAA* variants that are rare, certain variants are more frequent. Common variants are specific to geographical regions, for example the three Caucasian variants c.-32-13T>G (IVS1), c.2481+102_2646+31del (p.Gly828_Asn882del), c.525del (p.Glu176Argfs*45) and the two Asian variants c.1935C>A (p.Asp645Glu) and c.2238G>C (p.Trp746Cys) (Ausems, et al., 2001; Dagnino, et al., 2000; Hirschhorn and Reuser, 2001; Huie, et al., 1994; Kroos, et al., 1995; Laforet, et al., 2000; Lin and Shieh, 1996; Liu, et al., 2014; Muller-Felber, et al., 2007; Pittis and Filocamo, 2007; Shieh and Lin, 1998; Van der Kraan, et al., 1994; Wokke, et al., 1995). In the case of the IVS1 variant, a very broad range of symptom onset occurs (Herzog, et al., 2012; Kroos, et al., 2007). The phenotypic variation is also observed within and between families as is illustrated by a study on 22 families with two or three siblings with the IVS1 variant (Wens, et al., 2013). A large subset of unrelated patients witin this cohort also contained identical *GAA* variants on the second allele but varied widely in symptom onset, ruling out any effect of the second *GAA* variant on phenotypic variation. This variation has led to the idea that other factors -genetic,

epigenetic, or environmental- may modify the disease course of patients with the IVS1 variant (Herzog, et al., 2012; Ko, et al., 1999; Kroos, et al., 2012b; Kroos, et al., 2007; Shieh and Lin, 1998; Wens, et al., 2013). It it still unclear how the phenotypic variation is for other disease associated *GAA* variants is and how these relate to the variation observed in IVS1 patients. The accompanying manuscript describes how we have extended the Pompe mutation database at www.pompecenter.nl by linking genotype information and clinical information (Niño, et al., 2019). In the present study, we used the information included in the Pompe mutation database to compare the variation in symptom onset for common *GAA* genotypes in Pompe disease. The information is useful for the clinical prognosis of Pompe disease, and it provides a framework for investigating modifying factors that may contribute to the clinical phenotype of Pompe disease.

METHODS

NM_000152.3 was used as reference sequence for *GAA* mRNA and NP_000143.2 for GAA protein. Position c.1 represents the first nucleotide of the translation start codon located in *GAA* exon 2. We used the updated version of the Pompe mutation database (available at www.pompecenter.nl), in which *GAA* variants were linked to clinical phenotypes (Niño, et al., 2019). The patients' country or ethnic origin is stated in the database. For the analyses described in this manuscript, these origins were categorized in the four following groups: Caucasian (for patients from Europe, North America and Australia), Latin American (for patients from Central and South America), African (for patients from the African continent and also for African-American patients) and Asian (for patients from the Asian continent).

RESULTS

Phenotypic variation in compound heterozygous patients

The most frequent compound heterozygous genotypes were listed from patients in whom a classified disease associated *GAA* variant was combined with a null allele, which is defined as an allele that produces no detectable GAA enzyme activity. Their age of disease onset was plotted per GAA variant (Figure 1 and Table 1). We note that while other frequent compound heterozygous genotypes were present in the database, the second allele in these cases was not a null allele, and genotypic variation could not therefore be attributed to a specific disease associated variant.

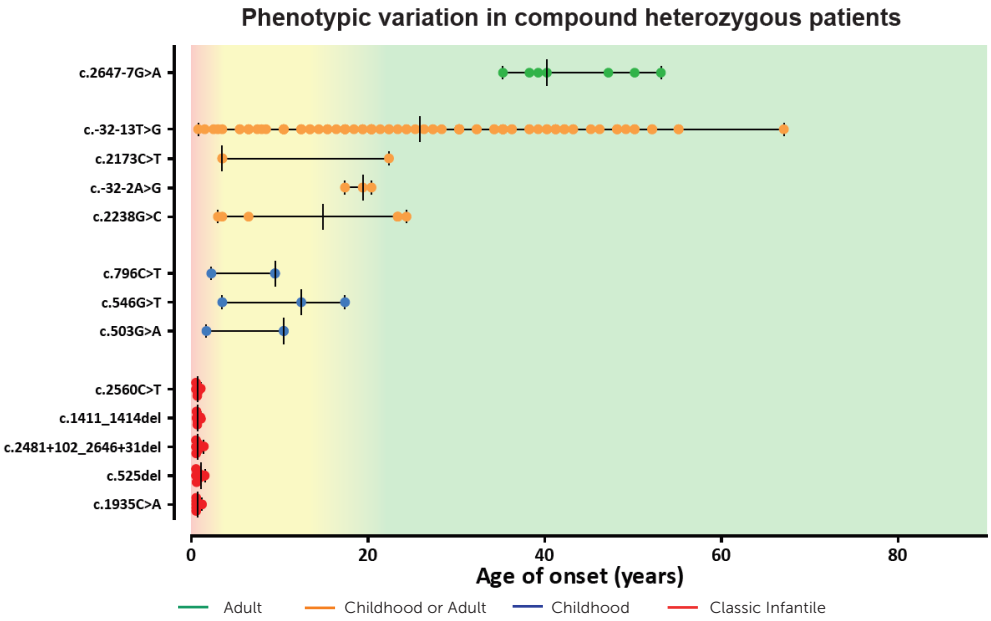


Figure 1: Variations in disease onset in compound heterozygous patients. Age of disease onset is indicated for the most common variants. The variant severity rating is shown for clarity. Patients carry the indicated variant exclusively in combination with a null allele. Per variant, the median age of onset is indicated by a vertical line. One exceptional case was excluded from the figure, this involved an adult patient with two very severe *GAA* variants (genotype: c.1935C>A/c.2560C>T, age of onset: 25 years, phenotype: adult), who was considered to be an enigmatic case (Hermans, et al., 1993).

Table 1. Phenotypic variation of most common *GAA* variants when present in compound heterozygous genotypes in combination with a *GAA* null allele.

DNA nomenclature	Predicted severity	Phenotype with a null allele	Total no. of patients	Age of onset	Population
c.2647-7G>A Combined with 1 null allele	Potentially mild	Adult	7	35 to 53 years (7)	Caucasian
c.-32-13T>G Combined with 40 different null alleles	Potentially mild	Childhood or Adult	141	3 months (1) 1 to 17 years (30) 18 to 67 years (63) ND (33) Asymptomatic (8) At adulthood (5) At childhood (1)	Caucasian, Latin-American
c.2173C>T Combined with 2 different alleles	Less severe	Childhood or Adult	3	3 years (2) 22 years (1)	Caucasian
c.-32-2A>G Combined with 1 null allele	Very severe	Childhood or Adult	3	17 years (1) 19 years (1) 20 years (1)	Latin-American
c.2238G>C Combined with 3 different null alleles	Potentially mild	Childhood or Adult	9	>1 and <3 years (1) 3 years (1) 6 years (1) 23 to 24 years (3) ND (3)	Asian

Table 1 continued.

DNA nomenclature	Predicted severity	Phenotype with a null allele	Total no. of patients	Age of onset	Population
c.796C>T Combined with 2 different null alleles	Potentially mild	Childhood	3	9 years (2) 1,75 years (1)	Asian
c.546G>T Combined with 2 different null alleles	Potentially mild	Childhood	3	3 years (1) 12 years (1) 17 years (1)	Asian
c.503G>A Combined with 1 null allele	Unknown	Childhood	3	10 years (2) 1,2 years (1)	Asian
c.2560C>T Combined with 8 different null alleles	Very severe	Classic infantile	8	<7 months (8)	African, Latin-American
c.1411_1414del Combined with 3 different null alleles	Very severe	Classic infantile	12	<7 months (9) N.D. (3)	Asian
c.2481+102_2646+31del Combined with 15 different null alleles	Very severe	Classic infantile	20	<1 year (10) N.D. (10)	Caucasian, Latin-American
c.525del Combined with 17 different null alleles	Very severe	Classic infantile	22	< 1year (15) 1 year (1) N.D. (6)	Caucasian
c.1935C>A Combined with 23 different null alleles	Potentially less severe	Classic infantile	39	<8 months (31) N.D. (8)	Asian

Genotypes containing at least one null allele were grouped according to the clinical severity of the disease associated *GAA* variant as decribed (Niño, et al., 2019). The groups are indicated on the Y axis of Figure 1 and are: (1) classic infantile, (2) childhood, (3) childhood or adult, and (4) adult. There was little phenotypic variation in patients with the classic infantile form of Pompe disease group: age of disease onset was below the age of 1 year by definition. Variants associated with the childhood phenotype included the missense variants c.796C>T (Asian) and c.503G>A (Asian), and the splicing variant c.546G>T (Asian). Despite the low number of patients, it was apparent that for all three variants the age of onset ranged from a minimum of 1.2 years to a maximum of 17 years (n=9 in total for this group). The group containing the most common variants associated with childhood or adult phenotypes was strongly dominated by the splicing variant c.-32-13T>G (IVS1) (Caucasian, Latin American). Age of onset in patients with the IVS1 variant in whom this was combined with a null allele differed exceptionally broadly (range 0.25-67 years). This contrasted with the age at onset of patients with other variants in this group including the missense variants c.2238G>C (Asian) and c.2173C>T (Caucasian), and the splice variant c.-32-2A>G (Latin American); this ranged from a minimum of 1 year to a maximum of 24 years-less broad than in those with the IVS1 variant but still substantial. In those with the non-coding splice variant c.2647-7G>A (Caucasian), which is associated with the adult phenotype, age of onset ranged from a minimum of 35 years to a maximum of 53 years. The information on the c.2647-7G>A variant is derived from a single Italian family and has been described previously (Sampaolo, et al., 2013).

Altogether, while age at onset in those with *GAA* variants associated with the classic infantile phenotype was similar (i.e., before the age of 1 year), it varied considerably in patients with other variants. Except for those with the IVS1 variant, whose ages at onset covered an 67-year range, the maximum range was 23 years.

Phenotypic variation in homozygous patients

Following an analysis similar to that in Figure 1, we analyzed differences in age of onset for four groups of homozygous patients (Figure 2 and Table 2. groups are indicated on the Y axis of Figure 2). As expected, age of onset varied little for patients who were homozygous for variants associated with the classic infantile phenotype (i.e. below 1 year of age in all cases). Age of onset varied from 0.5 to 6 years in patients who were homozygous for the missense variant c.1064T>C (Caucasian), associated with the classic infantile or childhood phenotype. In patients with the c.-32-3C>A splicing variant (Latin-American), associated with the childhood phenotype, age of onset varied from 31 to 43 years. Interestingly, the ages of onset of two patients who were compound heterozygous for c.-32-3C>A combined with a null allele varied from 0.5 to 7 years of age (data not shown). This suggests that onset had been delayed considerably from an early childhood to a mid-adult age by doubling of the residual enzyme activity in patients with two c.-32-3C>A alleles. Age of onset in patients who were homozygous for the c.-32-13T>G (IVS1) variant (Caucasian and Latin American), associated with the childhood or adult phenotype, ranged very broadly from 6 to 58 years. From this analysis of patients with a homozygous *GAA* genotype, we conclude that a broad range of disease onset in childhood and adult Pompe patients exists. The c.-32-13T>G (IVS1) variants shows an exceptionally broad range of disease onset.

Comparison of patients with compound heterozygous and homozygous c.-32-13T>G (IVS1) genotypes

We compared patients that carried the c.-32-13T>G (IVS1) *GAA* variant on one allele and different *GAA* variants on the second allele: 1) a null allele; 2) a non-null allele (i.e. any classified *GAA* variant except null alleles and c.-32-13T>G (IVS1) itself); and 3) c.-32-13T>G (IVS1) itself (Figure 3). While the ranges of disease onset were similarly broad in all three cases, the median age of onset depended on the severity of the second allele. In patients with an IVS1/null allele genotype, median age of onset was 25.5 years, in patients with an IVS1/non-null allele, this was 29, and in homozygous IVS1 patients, it was 39.5. These results suggests that in patients with the c.-32-13T>G (IVS1) variant, disease onset can be modulated by the severity of the second *GAA* allele. Because very few homozygous IVS1 patients have been reported among many patients diagnosed with Pompe disease, we speculate that a subset of homozygous IVS1 individuals remains

asymptomatic and have therefore not been included in the analysis.

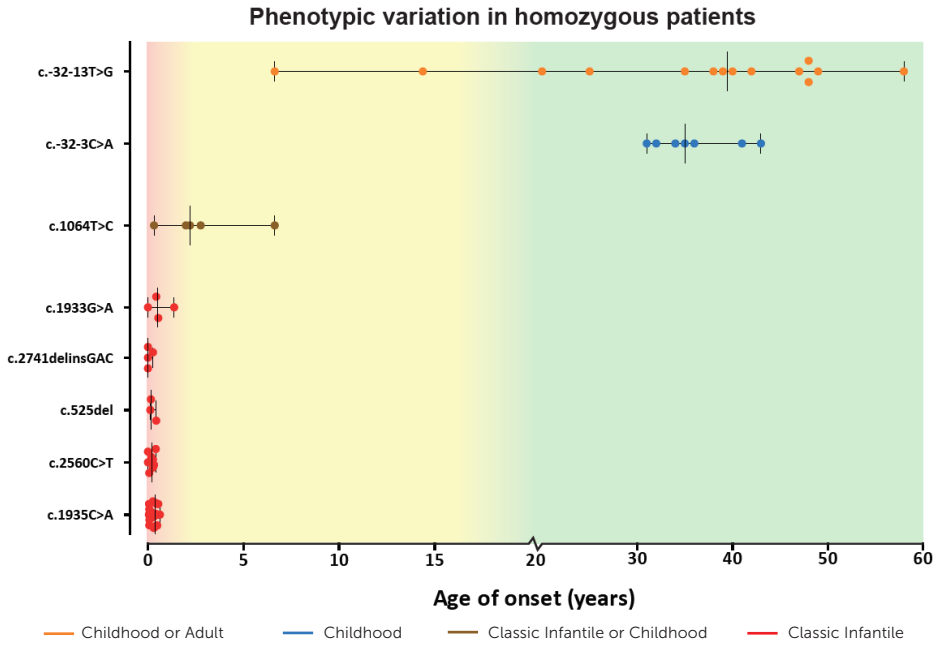


Figure 2: Variation in disease onset in homozygous patients. Disease onset is indicated for the most common variants. Per variant, the median age of onset is indicated by a vertical line.

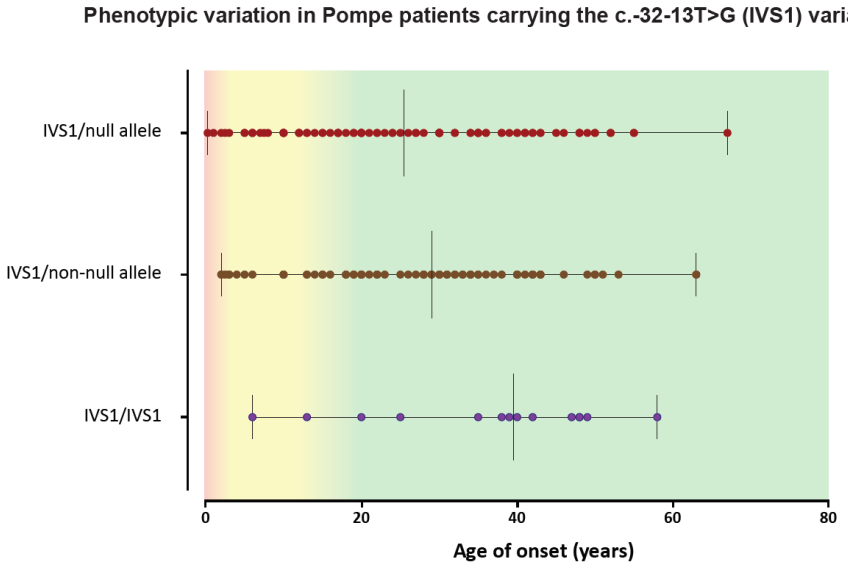


Figure 3: Variation in age of disease onset of Pompe patients carrying the c.-32-13T>G (IVS1) variant in combination with a null allele (top), a disease associated allele other than a null allele (non-null; middle), or a second c.-32-13T>G (IVS1) variant (bottom). The median age of onset is indicated by a vertical line.

Table 2. Phenotypic variation of most common *GAA* variants when present in homozygous *GAA* genotypes.

DNA nomenclature	Predicted severity	Phenotype with a null allele	Total no. of patients	Age of onset	Population
c.-32-13T>G	Potentially mild	Childhood or Adult	17	**Family 1:** 40 and 49 years **Others:** 6 years (1) 13 years (1) 20 to 58 years (10) N.D. (3)	Caucasian, Latin-American
c.-32-3C>A	Less severe	Childhood	7	**Family 1**: 34, 35 and 36 years **Family 2:** 31, 32 and 43 years **Other:** 41 years	Latin-American
c.1064T>C	Potentially less severe	Classic Infantile or Childhood	6	<4 months to 6 years (5) Childhood (1)	Caucasian, Latin-American, Asian
c.1933G>A	Potentially less severe	Classic Infantile	4	Birth (1) <1 year (2) 1 year, 3 months (1)	Caucasian, Asian
c.2741delinsGAC	Very severe	Classic Infantile	5	<2 years (4) N.D. (1)	Caucasian, Asian
c.525del	Very severe	Classic Infantile	8	<5 months (3) N.D. (5)	Caucasian
c.2560C>T	Very severe	Classic Infantile	13	<5 months (8) N.D. (5)	African, Asian, Latin-American
c.1935C>A	Potentially less severe	Classic Infantile	35	<7 months (29) N.D. (6)	Asian

DISCUSSION

Our analysis of the Pompe mutation database reveals that identical or similar *GAA* genotypes can result in substantial differences in age of disease onset. An exception is the homogeneous classic infantile phenotype that occurs in patients with two very severe *GAA* variants. Age of onset in patients with *GAA* variants associated with the childhood or adult phenotype differed by up to 23 years, except in patients with the IVS1 variant, in whom it varied by up to 67 years.

As suggested previously (Kroos, et al., 2007; Wens, et al., 2013), this strongly suggests that the progression of Pompe disease can be modified by additional factors. It is well known that monogenic diseases caused by the same genotypes can have variable phenotypes, and that additional (epi)genetic factors can affect disease progression (Kammenga, 2017). In the case of Pompe disease, very few disease-modifying factors—including *ACE* and *ACTN*—have been reported to date (Baek, et al., 2016; de Filippi, et al., 2010; De Filippi, et al., 2014). However, as only small differences in disease onset

have been observed, these cannot explain the large phenotypic differences in Pompe patients (Baek, et al., 2016; de Filippi, et al., 2010; De Filippi, et al., 2014).

From the current analysis, we propose that there are at least two sets of modifying factors. The first may be independent of the type of variant, and may modify the overall progression of Pompe disease. This is suggested by the observation that many different *GAA* variants have a broad phenotypic variation. The second set of factors may be variant specific. As the IVS1 variant stands out for its exceptionally broad phenotypic variation, it is a good candidate for such variant-specific modfying factor(s). It should be noted that far more patients with the IVS1 variant have been reported than patients with other variants. To more accurately establish the age range for these other variants, the group size for determining the phenotypic range should therefore be increased.

At this point, we can only speculate upon the identity of these putative modifying factors. Options include environmental, epigenetic, or genetic modifiers. It is interesting to note that the variability of patients with identical genotypes (IVS1/c.525del) was larger between families rather than within families (Wens, et al., 2013). Although this does not rule out epigenetic or environmental effects, one could speculate that modifiers of Pompe disease include genetic factors. One would expect candidate modifying genetic variants to be present in pathways involved in Pompe disease, such as skeletal muscle function and regeneration, glucose metabolism, autophagy, and lysosomal biogenesis. We speculate that *GAA* variant-specific modifiers may, in the case of splice variants, include splicing factors; or, in the case of missense variants, components of the endoplasmic reticulum-associated protein degradation (ERAD) machinery pathway involved in protein folding (Fan, 2008). The identification of such putative genetic factors should be aided by genome-wide analyses of large patient groups.

The identification of factors that can modify the progression of Pompe disease will be highly relevant for newborn screening programs for Pompe disease (Burton, et al., 2017; Johnson, et al., 2017; Kronn, et al., 2017; Savarese, et al., 2018). In the case that such factors will be found to be genetic, it will be straightforward to include genetic analysis in the program. Future genetic counseling of patients and their parents should be based not only on the severity of the disease associated *GAA* variant, but also on the presence of identified modifying factors that can accelerate or slow down disease progression. In current clinical practice, the decision on the timing of start of treatment with ERT is based on clinical parameters (Broomfield, et al., 2016; Echaniz-Laguna, et al., 2015; Kuperus, et al., 2017; Schoser, 2015; Schoser, et al., 2017; Stepien, et al., 2016; Strothotte, et al., 2010; van der Meijden, et al., 2018; van der Ploeg, et

al., 2017). For childhood or adult patiens, the timing at which treatment should be started is difficult to determine. In many cases, treatment will be started when disease symptoms start to develop. However, it is desirable to start treatment prior to the onset of pathology, because tissue damage may be difficult to revert. On the other hand, a too early treatment will impose uncessary burden on the patient and will increase costs. A better understanding of the variations in symptom onset that are associated with *GAA* variants should aid in a better prognosis of the progression of symptom development for patients with Pompe disease.

Acknowledgements
We thank Dr. Arnold JJ Reuser for his critical review of the manuscript, and Ton Verkerk for his invaluable assistance with generating the database. This work was supported by grants from the Administrative Department of Science, Technology and Innovation (Colciencias, Colombia) and Sanofi Genzyme. This project has received funding from the Ministry of Economic Affairs under TKI-allowance under the TKI-programme Life Sciences & Health (project number LSHM16008).

Research on Pompe disease at Erasmus MC is financially supported by "Prinses Beatrix Spierfonds" [project number W.OR13-21, W.OR15-10, W.OR16-07]; Tex Net; Sophia foundation for medical research (SSWO) [project number s17-32]; Metakids [project number 2016-063]; Conselho Nacional de Desenvolvimento Científico e Tecnológico - "National Counsil of Technological and Scientific Development", Brasil; Colciencias and Sanofi Genzyme. AvdP has provided consulting services for various industries in the field of Pompe disease under an agreement between these industries and Erasmus MC, Rotterdam, The Netherlands. The other authors declare that they have no conflict of interest.

REFERENCES

Ausems MG, ten Berg K, Sandkuijl LA, Kroos MA, Bardoel AF, Roumelioti KN, Reuser AJ, Sinke R, Wijmenga C. 2001. Dutch patients with glycogen storage disease type II show common ancestry for the 525delT and del exon 18 mutations. J Med Genet 38(8):527-9.

Baek RC, Palmer R, Pomponio RJ, Lu Y, Ma X, McVie-Wylie AJ. 2016. The influence of a polymorphism in the gene encoding angiotensin converting enzyme (ACE) on treatment outcomes in late-onset Pompe patients receiving alglucosidase alfa. Mol Genet Metab Rep 8:48-50.

Broomfield A, Fletcher J, Davison J, Finnegan N, Fenton M, Chikermane A, Beesley C, Harvey K, Cullen E, Stewart C and others. 2016. Response of 33 UK patients with infantile-onset Pompe disease to enzyme replacement therapy. J Inherit Metab Dis 39(2):261-71.

Burton BK, Kronn DF, Hwu WL, Kishnani PS, Pompe Disease Newborn Screening Working G. 2017. The Initial Evaluation of Patients After Positive Newborn Screening: Recommended Algorithms Leading to a Confirmed Diagnosis of Pompe Disease. Pediatrics 140(Suppl 1):S14-S23.

Dagnino F, Stroppiano M, Regis S, Bonuccelli G, Filocamo M. 2000. Evidence for a founder effect in Sicilian patients with glycogen storage disease type II. Hum Hered 50(6):331-3.

de Filippi P, Ravaglia S, Bembi B, Costa A, Moglia A, Piccolo G, Repetto A, Dardis A, Greco G, Ciana G and others. 2010. The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. Genet Med 12(4):206-11.

De Filippi P, Saeidi K, Ravaglia S, Dardis A, Angelini C, Mongini T, Morandi L, Moggio M, Di Muzio A, Filosto M and others. 2014. Genotype-phenotype correlation in Pompe disease, a step forward. Orphanet J Rare Dis 9(1):102.

Echaniz-Laguna A, Carlier RY, Laloui K, Carlier P, Salort-Campana E, Pouget J, Laforet P. 2015. Should patients with asymptomatic pompe disease be treated? A nationwide study in France. Muscle Nerve 51(6):884-9.

Fan JQ. 2008. A counterintuitive approach to treat enzyme deficiencies: use of enzyme inhibitors for restoring mutant enzyme activity. Biol Chem 389(1):1-11.

Gaeta M, Musumeci O, Mondello S, Ruggeri P, Montagnese F, Cucinotta M, Vinci S, Milardi D, Toscano A. 2015. Clinical and pathophysiological clues of respiratory dysfunction in late-onset Pompe disease: New insights from a comparative study by MRI and respiratory function assessment. Neuromuscul Disord 25(11):852-8.

Gungor D, Reuser AJ. 2013. How to describe the clinical spectrum in Pompe disease? Am J Med Genet A 161A(2):399-400.

Hermans MM, de Graaff E, Kroos MA, Wisselaar HA, Willemsen R, Oostra BA, Reuser AJ. 1993. The conservative substitution Asp-645-->Glu in lysosomal alpha-glucosidase affects transport and phosphorylation of the enzyme in an adult patient with glycogen-storage disease type II. Biochem J 289 (Pt 3)(Pt 3):687-93.

Herzog A, Hartung R, Reuser AJ, Hermanns P, Runz H, Karabul N, Gokce S, Pohlenz J, Kampmann C, Lampe C and others. 2012. A cross-sectional single-centre study on the spectrum of Pompe disease, German patients: molecular analysis of the GAA gene, manifestation and genotype-phenotype correlations. Orphanet J Rare Dis 7(1):35.

Hirschhorn R, Reuser AJJ. 2001. Glycogen Storage Disease Type II: acid alpha-glucosidase (acid maltase) deficiency. In: Scriver CR, Beaudet AL, Valle D, Sly WS, editors. The Metabolic and Molecular Bases of Inherited Disease. NY: McGraw-Hill. p 3389-3420.

Huie ML, Chen AS, Tsujino S, Shanske S, DiMauro S, Engel AG, Hirschhorn R. 1994. Aberrant splicing in adult onset glycogen storage disease type II (GSDII): molecular identification of an IVS1 (-13T->G) mutation in a majority of patients and a novel IVS10 (+1GT->CT) mutation. Human Molecular Genetics 3(12):2231-6.

Johnson K, Topf A, Bertoli M, Phillips L, Claeys KG, Stojanovic VR, Peric S, Hahn A, Maddison P, Akay E and others. 2017. Identification of GAA variants through whole exome sequencing targeted to a cohort of 606 patients with unexplained limb-girdle muscle weakness. Orphanet J Rare Dis 12(1):173.

Kammenga JE. 2017. The background puzzle: how identical mutations in the same gene lead to different disease symptoms. FEBS J 284(20):3362-3373.

Ko TM, Hwu WL, Lin YW, Tseng LH, Hwa HL, Wang TR, Chuang SM. 1999. Molecular genetic study of Pompe disease in Chinese patients in Taiwan. Human Mutation 13(5):380-4.

Kronn DF, Day-Salvatore D, Hwu WL, Jones SA, Nakamura K, Okuyama T, Swoboda KJ, Kishnani PS, Pompe Disease Newborn Screening Working G. 2017. Management of Confirmed Newborn-Screened Patients With Pompe Disease Across the Disease Spectrum. Pediatrics 140(Suppl 1):S24-S45.

Kroos M, Hoogeveen-Westerveld M, Michelakakis H, Pomponio R, Van der Ploeg A, Halley D, Reuser A, Consortium GAAD. 2012a. Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants. Hum Mutat 33(8):1161-5.

Kroos M, Hoogeveen-Westerveld M, van der Ploeg A, Reuser AJ. 2012b. The genotype-phenotype correlation in Pompe disease. Am J Med Genet C Semin Med Genet 160C(1):59-68.

Kroos M, Pomponio RJ, van Vliet L, Palmer RE, Phipps M, Van der Helm R, Halley D, Reuser A, Consortium GAAD. 2008. Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. Hum Mutat 29(6):E13-26.

Kroos MA, Pomponio RJ, Hagemans ML, Keulemans JL, Phipps M, DeRiso M, Palmer RE, Ausems MG, Van der Beek NA, Van Diggelen OP and others. 2007. Broad spectrum of Pompe disease in patients with the same c.-32-13T->G haplotype. Neurology 68(2):110-5.

Kroos MA, Van der Kraan M, Van Diggelen OP, Kleijer WJ, Reuser AJJ, Van den Boogaard MJ, Ausems MGEM, Ploos van Amstel HK, Poenaru L, Nicolino M and others. 1995. Glycogen storage disease type II: frequency of three common mutant alleles and their associated clinical phenotypes studied in 121 patients. J Med Genet 32(10):836-7.

Kuperus E, Kruijshaar ME, Wens SCA, de Vries JM, Favejee MM, van der Meijden JC, Rizopoulos D, Brusse E, van Doorn PA, van der Ploeg AT and others. 2017. Long-term benefit of enzyme replacement therapy in Pompe disease: A 5-year prospective study. Neurology 89(23):2365-2373.

Laforet P, Laloui K, Granger B, Hamroun D, Taouagh N, Hogrel JY, Orlikowski D, Bouhour F, Lacour A, Salort-Campana E and others. 2013. The French Pompe registry. Baseline characteristics of a cohort of 126 patients with adult Pompe disease. Rev Neurol (Paris) 169(8-9):595-602.

Laforet P, Nicolino M, Eymard PB, Puech JP, Caillaud C, Poenaru L, Fardeau M. 2000. Juvenile and adult-onset acid maltase deficiency in France: genotype-phenotype correlation. Neurology 55(8):1122-8.

Lin CY, Shieh JJ. 1996. Molecular study on the infantile form of Pompe disease in Chinese in Taiwan. Chung-Hua Min Kuo Hsiao Erh Ko i Hsueh Hui Tsa Chih 37(2):115-21.

Liu X, Wang Z, Jin W, Lv H, Zhang W, Que C, Huang Y, Yuan Y. 2014. Clinical and GAA gene mutation analysis in mainland Chinese patients with late-onset Pompe disease: identifying c.2238G > C as the most common mutation. BMC Med Genet 15(1):141.

Muller-Felber W, Horvath R, Gempel K, Podskarbi T, Shin Y, Pongratz D, Walter MC, Baethmann M, Schlotter-Weigel B, Lochmuller H and others. 2007. Late onset Pompe disease: clinical and neurophysiological spectrum of 38 patients including long-term follow-up in 18 patients. Neuromuscul Disord 17(9-10):698-706.

Niño, MY, In 't Groen, SLM, Bergsma, AJ, van der Beek, N, Kroos, M, Hoogeveen-Westerveld, M, *et al.* (2019). Extension of the Pompe mutation database by linking disease-associated variants to clinical severity. *Hum Mutat* ***40****: 1954-1967*.

Pittis MG, Filocamo M. 2007. Molecular genetics of late onset glycogen storage disease II in Italy. Acta Myol 26(1):67-71.

Sampaolo S, Esposito T, Farina O, Formicola D, Diodato D, Gianfrancesco F, Cipullo F, Cremone G, Cirillo M, Del Viscovo L and others. 2013. Distinct disease phenotypes linked to different combinations of GAA mutations in a large late-onset GSDII sibship. Orphanet J Rare Dis 8:159.

Savarese M, Torella A, Musumeci O, Angelini C, Astrea G, Bello L, Bruno C, Comi GP, Di Fruscio G, Piluso G and others. 2018. Targeted gene panel screening is an effective tool to identify undiagnosed late onset Pompe disease. Neuromuscul Disord.

Schoser B. 2015. A Troublesome Debate: When to Start Treatment in Adult Pompe Patients? J Neuromuscul Dis 2(s1):S8.

Schoser B, Laforet P, Kruijshaar ME, Toscano A, van Doorn PA, van der Ploeg AT, European Pompe C. 2015. 208th ENMC International Workshop: Formation of a European Network to develop a European data sharing model and treatment guidelines for Pompe disease Naarden, The Netherlands, 26-28 September 2014. Neuromuscul Disord 25(8):674-8.

Schoser B, Stewart A, Kanters S, Hamed A, Jansen J, Chan K, Karamouzian M, Toscano A. 2017. Survival and long-term outcomes in late-onset Pompe disease following alglucosidase alfa treatment: a systematic review and meta-analysis. J Neurol 264(4):621-630.

Shieh JJ, Lin CY. 1998. Frequent mutation in Chinese patients with infantile type of GSD II in Taiwan: evidence for a founder effect. Human Mutation 11(4):306-12.

Stepien KM, Hendriksz CJ, Roberts M, Sharma R. 2016. Observational clinical study of 22 adult-onset Pompe disease patients undergoing enzyme replacement therapy over 5years. Mol Genet Metab 117(4):413-8.

Strothotte S, Strigl-Pill N, Grunert B, Kornblum C, Eger K, Wessig C, Deschauer M, Breunig F, Glocker FX, Vielhaber S and others. 2010. Enzyme replacement therapy with alglucosidase alfa in 44 patients with late-onset glycogen storage disease type 2: 12-month results of an observational clinical trial. J Neurol

257(1):91-7.

Van der Beek NA, Hagemans ML, Reuser AJ, Hop WC, Van der Ploeg AT, Van Doorn PA, Wokke JH. 2009. Rate of disease progression during long-term follow-up of patients with late-onset Pompe disease. Neuromuscul Disord 19(2):113-7.

Van der Kraan M, Kroos MA, Joosse M, Bijvoet AG, Verbeet MP, Kleijer WJ, Reuser AJ. 1994. Deletion of exon 18 is a frequent mutation in glycogen storage disease type II. Biochemical & Biophysical Research Communications 203(3):1535-41.

van der Meijden JC, Kruijshaar ME, Harlaar L, Rizopoulos D, van der Beek N, van der Ploeg AT. 2018. Long-term follow-up of 17 patients with childhood Pompe disease treated with enzyme replacement therapy. J Inherit Metab Dis.

van der Ploeg AT, Kruijshaar ME, Toscano A, Laforet P, Angelini C, Lachmann RH, Pascual Pascual SI, Roberts M, Rosler K, Stulnig T and others. 2017. European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. Eur J Neurol 24(6):768-e31.

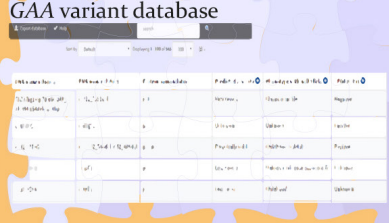
van der Ploeg AT, Reuser AJ. 2008. Pompe's disease. Lancet 372(9646):1342-53.

Wens SC, van Gelder CM, Kruijshaar ME, de Vries JM, van der Beek NA, Reuser AJ, van Doorn PA, van der Ploeg AT, Brusse E. 2013. Phenotypical variation within 22 families with Pompe disease. Orphanet J Rare Dis 8(1):182.

Wokke JH, Ausems MG, van den Boogaard MJ, Ippel EF, van Diggelene O, Kroos MA, Boer M, Jennekens FG, Reuser AJ, Ploos van Amstel HK. 1995. Genotype-phenotype correlation in adult-onset acid maltase deficiency. Ann Neurol 38(3):450-4.

Pompe disease

GAA variant database



CHAPTER 5

THE ACE I/D POLYMORPHISM DOES NOT EXPLAIN HETEROGENEITY OF NATURAL COURSE AND RESPONSE TO ENZYME REPLACEMENT THERAPY IN POMPE DISEASE

Kuperus E[1,2,‡], van der Meijden JC[1,3,‡], In 't Groen SLM[1,3,4], Kroos MA[1,3,4], Hoogeveen-Westerveld M[1,3,4], Rizopoulos D[5], **Niño MMY[1,3,4]**, Kruijshaar ME[1,3], van Doorn PA[1,2], van der Beek NAME[1,2,*], van der Ploeg AT[1,3], Pijnappel WWMP[1,3,4,*].

Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

Department of Neurology, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

Department of Pediatrics, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

Department of Biostatistics, Erasmus MC University Medical Center, Rotterdam, the Netherlands.

‡ These authors contributed equally; *Corresponding authors

PLoS One. 2018 Dec 7;13(12):e0208854.

5

The ACE I/D polymorphism does not explain heterogeneity of natural course and response to enzyme replacement therapy in Pompe disease

Kuperus E, van der Meijden JC, In 't Groen SLM, Kroos MA, Hoogeveen-Westerveld M, Rizopoulos D, **Niño MMY**, Kruijshaar ME, van Doorn PA, van der Beek NAME, van der Ploeg AT, Pijnappel WWMP.

ABSTRACT

The majority of children and adults with Pompe disease in the population of European descent carry the leaky splicing *GAA* variant c.-32-13T>G (IVS1) in combination with a fully deleterious *GAA* variant on the second allele. The phenotypic spectrum of this patient group is exceptionally broad, with symptom onset ranging from early infancy to late adulthood. In addition, the response to enzyme replacement therapy (ERT) varies between patients. The insertion/deletion (I/D) polymorphism of the angiotensin I-converting enzyme (*ACE*) has been suggested to be a modifier of disease onset and/or response to ERT. Here, we have investigated the effect of the *ACE* I/D polymorphism in a relatively large cohort of 131 children and adults with Pompe disease, of whom 112 were followed during treatment with ERT for 5 years. We assessed the use of wheelchair and mechanical ventilation, muscle strength assessed via manual muscle testing and hand-held dynamometry (HHD), distance walked on the six-minute walk test (6MWT), forced vital capacity (FVC) in sitting and supine position and daily-life activities assessed by R-PAct. Cross sectional analysis at first visit showed no differences between the genotypes with respect to age at first symptoms, diagnosis, wheelchair use, or ventilator use. Also response to ERT over 5 years assessed by linear mixed model analyses showed no significant differences between *ACE* groups for any of the outcome measures. The patient cohort contained 24 families with 54 siblings. Differences in ACE genotype could neither explain inter nor intra familial differences.

We conclude that the *ACE* I/D polymorphism does not explain the large variation in disease severity and response to ERT observed among Pompe patients with the same c.-32-13T>G *GAA* variant.

Keywords: *ACE* I/D polymorphism, clinical variation, enzyme remplacemt therapy

INTRODUCTION

Pompe disease (OMIM 232300) is a metabolic myopathy caused by disease-associated variants in the acid α-glucosidase *(GAA)* gene (OMIM 606800). This results in deficiency of the lysosomal enzyme GAA, leading to an impaired breakdown of glycogen [1]. Clinically, a broad disease spectrum can be observed, ranging from a rapidly progressive classic infantile phenotype to a slower progressing disease course in children and adults. Classic infantile patients present shortly after birth with hypertrophic cardiomyopathy and generalized muscle weakness. Without treatment these patients die within the first year of life due to cardiorespiratory insufficiency [2, 3]. Children and adults present with a slower progressive limb girdle muscle weakness, while cardiac involvement is rare. Most of these patients become wheelchair and ventilator dependent. Survival is reduced compared to the general population [4–6].

The majority of children and adults of European descent with Pompe disease carry the common c.-32-13T>G (IVS1) variant on one *GAA* allele. The IVS1 variant causes aberrant *GAA* pre-mRNA splicing by inducing partial or complete skipping of *GAA* exon 2. A small percentage (10–15%) of leaky wild type splicing occurs [7–12]. The second *GAA* allele in children and adults with Pompe disease is in many cases a 'null' allele, which is defined as an allele that does not generate any detectable *GAA* enzymatic activity, for example an allele that carries the frequent variant c.525delT [13]. Interestingly, the natural disease course in patients with the IVS1 variant shows an exceptionally broad spectrum: disease onset can vary from early infancy to late adulthood and symptoms can vary. Such differences are even observed among patients with the same IVS1/c.525delT *GAA* genotype [14–17]. The variation between siblings is less broad, suggesting that disease progression can be modified by genetic background factors [17].

Since 2006, enzyme replacement therapy (ERT) with recombinant human *GAA* (rhGAA), is available for Pompe disease. In children and adults with Pompe disease ERT has shown to improve muscle function and strength and to stabilize pulmonary function.

However, individual responses can vary considerably [18], which is highlighted in our recent study on the clinical response to ERT during 5 year follow up [19]. This is observed irrespective of formation of anti-rhGAA antibodies, suggesting that other factors exist that can modify the response to ERT in this patient group [19–21].

It has been suggested that a polymorphism in the angiotensin converting enzyme (*ACE*) gene -the insertion (I) or deletion (D) of an alu repeat in intron 16- may affect phenotypic variation and response to ERT in patients with Pompe disease [22–25]. So far, four studies were performed, but the outcomes of the studies were different [22–25]. The DD genotype was associated with an earlier disease onset in some studies [22, 23], but not in another [25], and it was associated with less favorable response to ERT with regard to muscle mass in one study [24], and with regard to FVC and 6MWT in another study [25]. The different outcomes may be explained by small group sizes, an overrepresentation of adult patients, and/or a short follow up of maximum 2 years. This was reason to perform the current nationwide study in a group of 131 children and adults representing the full phenotypic spectrum of patients with the c.-32-13T>G/ 'null' gene type. The aim was to further explore the potential influential effect of the *ACE* polymorphism on age of disease onset, disease severity and outcome of patients when treated with ERT by using a relatively large patient cohort and a longer follow up of 5 years of ERT.

5

MATERIALS AND METHODS

Patients and study design

This study was part of an ongoing single-center prospective, open-label study, in which all Dutch children and adults with a confirmed diagnosis of Pompe disease by enzyme analysis in leucocytes or fibroblasts, and by DNA analysis participated [15, 26–28]. Only patients that carried the c.-32-13 T>G (IVS1) *GAA* variant on one allele and a fully deleterious ("null") *GAA* variant on the other allele were included in this study. Information was collected with respect to onset of symptoms, age of diagnosis and wheelchair- or ventilator dependency. Clinical assessments were performed with 3–6 months intervals. Daily life activities were assessed via the Rasch-built Pompe-specific Activity (R-PAct) scale [29]. Data were collected from January 1, 1999 through January 1, 2016. All data available during this period were used in this analysis. The study was conducted according to the Declaration of Helsinki, the Medical Ethical Committee at Erasmus MC University Medical Center approved the study protocol, and all patients, or their parents or legal guardians, provided written informed consent.

ACE polymorphism

ACE genotyping was performed based on the methods described by Al-Awadhi et al [30]. In short, genomic DNA from blood or fibroblasts was used in a first PCR flanking the alu insertion in intron 16 using the following primers: fw1: 5'-CTGGAGACCACTCCCATCCTTTCT-3' and rev1: 5'-GATGTGGCCATCACATTCGTCAGAT-3'. PCR was performed using FastStart PCR (Roche) in which 6% DMSO was included in the reaction mixture. Reactions were run on a Bio-rad C1000 Touch Thermal Cycler. When no I allele was found, a second PCR was performed using an alu insertion-specific internal primer fw2: 5'-TGGGATTACAGGCGTGATACAG-3', which was used in a PCR with the same primer rev1 as above. Positive, negative, and blank controls for the *ACE* genotype were included in all analyses.

Clinical outcome measures

The use of a wheelchair and mechanical ventilation was registered at each visit. Skeletal muscle strength was assessed using the Medical Research Council (MRC) grading scale and hand-held dynamometry (HHD; Cytec dynamometer, Groningen, The Netherlands) [31–33]. The following muscle groups were tested for either method: neck flexors, shoulder abductors, elbow flexors, elbow extensors, hip flexors, hip abductors, knee flexors, and knee extensors. Additionally, an MRC grade was determined for the hip extensors and hip adductors. This was expressed as percentage of the maximum possible score for MRC sum scores. HHD sum scores were expressed as percentage of the median strength of healthy males and females. Sum scores were calculated if no more than two muscle groups were missing. Muscle function was assessed using the Quick Motor Function Test (QMFT), consisting of 16 motor skills related to daily life [34]. Muscle endurance was assessed using the six-minute walk test (6MWT) in which the distance walked in 6 minutes was recorded [35]. Forced vital capacity (FVC) was measured in upright and supine positions. Results were expressed as percentage of the predicted FVC [36]. The R-PAct scale was used to assess patients' self-reported ability to perform daily life activities. A score was calculated as described [29], only when all items had been answered. Only adult patients performed this test.

Statistical analysis

Differences in characteristics between the three different *ACE* genotypes (II, DD and ID) at first visit were calculated as follows. First we tested if any of the *ACE* groups differed from the others using the Kruskal-Wallis test for numerical data and the chi-square test (2x3) for categorical data. When significant, the Mann-Whitney and chi-square tests (2x2) were used to identify which of the groups (II vs ID, II vs DD or ID vs DD) differed. We corrected for multiple testing using the Holm method [37].

Longitudinal analyses of the effects of ERT were performed using linear mixed effect models to account for repeated measurements per patient as described [19]. Models were fitted for each outcome measure using the nlme package of the statistical program R (version 3.2.5) [38, 39]. Time was expressed as years after start of ERT. To account for potential non-linear profiles we used natural cubic splines in the fixed-effects and random-effects parts of the model. In the specification of the splines, boundary knots were placed at 0 (i.e. start of ERT) and 5 years, and internal knots were placed at 1 and 3.5 years. For the random-effects part of the model an unstructured covariance matrix was used. Likelihood-ratio-tests were used to asses if there was an interaction between time and *ACE* polymorphism; e.g. if outcome measures progressed differently during treatment for each *ACE* polymorphism or if the different polymorphisms had different intercepts; e.g. overall disease severity. Obtained p-values were corrected for multiple testing using the Holm method [37]. Plots for the group means were generated for the different outcome measures during 5 years of ERT as described [19]. Siblings from various families were identified. For each sibling the age at first symptoms, diagnosis, the start of ERT, the start of wheelchair/ventilator use, and their last follow-up or death were plotted to study if different *ACE* genotypes explained variation within families. Siblings were classified as having varying outcomes when an event occurred in one sibling and did not occur within 10 years in the other sibling.

Power analysis

5

We conducted a post hoc power analysis based on our cross-sectional data using the outcome measure age of onset, as this is highly variable in patients with the same IVS-1/null variants. We assumed that the ACE genotype would fully explain all phenotypical variation between patients. In this scenario, the medians of the three ACE genotype groups (II, ID, and DD) are the same as the 25th, 50th and 75th percentiles of the age of onset of the overall population. We calculated how many subjects are needed to detect a difference between ACE groups at a power of >0.8 by simulating a thousand 'sample populations' with the given number of patients, and performing a Kruskal-Wallis test.

RESULTS

Patients

A total of 146 Dutch patients with the IVS1/"null" genotype were known in our center at data closure. DNA was available for 131 patients to determine the *ACE* genotype (Table 1), of whom 112 had started with ERT. The characteristics of the total patient group (n = 131) were as follows. Gender was evenly distributed (50% females) in the total group

and the three different *ACE* groups (II, DD and ID). Most patients (86%) started ERT at adulthood (>18 years of age). Median ages (in years, with ranges indicated in brackets) at symptom onset, diagnosis, first visit, and start of ERT were 31 (0–62); 38 (0–72); 46 (0–75); and 49 (1–76) years, respectively. Further testing of the age range at start of ERT revealed that these did not differ between the ACE groups. At first visit, 31% of patients used a wheelchair, while the median age at which patients started to use a wheelchair was 49 (11–76) years. For usage of a ventilator, these numbers were 22% and 52 years of age (6–72). The age ranges for all parameters listed above were broad, highlighting the heterogeneity of disease onset and progression of Pompe patients with the IVS1 variant.

Effect of ACE I/D polymorphism at first visit: Cross sectional analysis

The ACE polymorphism genotype was normally distributed within the total patient group (II: 24%; ID: 44%; DD: 31%) (Table 1). Outcomes were compared between the three groups in a cross sectional analysis at first visit (Table 2A) and at start of ERT (Table 2B). This showed that none of the parameters were different between the groups. The use of a wheelchair at first visit was initially found to be different with a p-value of 0.047 with the lowest number of wheelchair users in the DD group, but post-hoc testing showed that this was not significant. Slightly more II patients started ERT during childhood, but the differences between ACE groups were also not significant.

Table 1. Patient characteristics at first visit.

	Total (n = 131/112[#])	ACE polymorphism			p-value
		II (n = 32; 24%)	DD (n = 41; 31%)	ID (n = 58)	
Gender, No. of patients (%)					
- Male	65 (50%)	15 (47%)	20 (49%)	30 (52%)	
- Female	66 (50%)	17 (53%)	21 (51%)	28 (48%)	n.s.
Start ERT during childhood (<18y), n (%)[#]					
- Yes	13 (12%)	6 (22%)	2 (6%)	5 (10%)	
- No	99 (88%)	20 (77%)	32 (94%)	46 (90%)	n.s.
Median age (range), at:					n.s.
Onset of symptoms	31 (0–62)	28 (0–54)	33 (4–61)	30 (0–62)	n.s.
Diagnosis	38 (0–72)	35 (0–69)	42 (0–67)	38 (0–72)	n.s.
First visit	46 (0–75)	41 (2–69)	47 (6–71)	47 (0–75)	n.s.
Start ERT[#§]	49 (1–76)	42 (1–68)	50 (14–73)	50 (1–76)	n.s.
Wheelchair use at first visit, n (%)					
- No	91 (69)	22 (69)	23 (56)	46 (79)	
- Yes	40 (31)	10 (31)	18 (44)	12 (21)	n.s.*
Wheelchair age, median (range)	49 (11–76)	43 (11–60)	51 (24–64)	55 (33–76)	n.s.
Ventilation use at first visit, n (%)					
- No	102 (78)	28 (88)	31 (76)	43 (74)	
- Yes	29 (22)	4 (12)	10 (24)	15 (26)	n.s.
Ventilation age, median (range)	52 (6–72)	53 (33–61)	48 (6–72)	51 (13–69)	n.s.

*Null hypothesis II = ID = DD rejected at the p<0.05 level. Post-hoc testing (II vs ID, II vs DD or ID vs DD) did not show significant differences between the ACE
groups.

#All parameters were tested for 131 patients, except for start ERT during childhood and median age at start ERT for which a total of 112 patients were analyzed.

§ age ranges were n.s.

https://doi.org/10.1371/journal.pone.0208854.t001

Table 2, Cross-sectional evaluation of muscle force, function and lung function at first visit (A) and start of ERT (B).

	Total (n = 131)	A. Outcomes at first visit			
		II (n = 32)	DD (n = 41)	ID (n = 58)	p-value
HHD; % of maximal score (range)	73 (32–99)	75 (33–95)	73 (3–97)	73 (32–99)	n.s.
MRC; % of maximal score (range)	83 (33–100)	86 (55–100)	81 (51–100)	84 (33–100)	n.s.
QMFT; % of maximal score (range)	67 (14–100)	80 (22–100)	64 (14–100)	65 (17–100)	n.s.
R-PAct; R-Pact score (range)	54 (7–100)	58 (7–100)	50 (7–94)	55 (17–83)	n.s.
6MWT; meters walked (range)	436 (48–650)	455 (75–645)	347 (82–544)	448 (48–650)	n.s.
FVC sitting; % of predicted (range)	73 (10–117)	84 (10–117)	73 (15–107)	72 (15–107)	n.s.
FVC supine; % of predicted (range)	61 (17–107)	71 (18–107)	68 (24–104)	48 (17–105)	n.s.
	Total (n = 112)	B. Outcomes at start ERT			
		II (n = 26)*	DD (n = 35)*	ID (n = 51)*	p-value
HHD; % of maximal score (range)	70 (26–100)	70 (33–100)	66 (26–95)	73 (26–95)	n.s.
MRC; % of maximal score (range)	82 (47–99)	84 (55–99)	79 (53–96)	82 (47–99)	n.s.
QMFT; % of maximal score (range)	74 (28–100)	57 (14–94)	57 (14–94)	63 (13–97)	n.s.
R-PAct; R-Pact score (range)	52 (7–86)	59 (7–86)	44 (7–75)	56 (17–83)	n.s.
6MWT; meters walked (range)	417 (41–650)	436 (75–645)	353 (82–626)	435 (41–650)	n.s.
FVC sitting; % of predicted (range)	57 (15–111)	84 (41–111)	57 (15–110)	62 (18–105)	n.s.
FVC supine; % of predicted (range)	52 (16–111)	65 (16–111)	58 (24–98)	44 (17–96)	n.s.

Abbreviations: ERT = enzyme replacement therapy; HHD = handheld dynamometry; MRC = Medical Research Council; QMFT = Quick Motor Function Test;
R-PAct = Rasch-Built Pompe-Specific Activity scale; 6MWT = 6-minute walk test; FVC = forced vital capacity percentage predicted; n.s. = not significant.
*The age ranges of patients that had started with ERT were 1–68 for the II group, 14–73 for the DD group and 1–76 for the ID group, and these were not significantly
different between the ACE genotype groups.
https://doi.org/10.1371/journal.pone.0208854.t002

Effect of ACE I/D polymorphism on response to enzyme replacement therapy

A total of 112 patients within the cohort received ERT and were included in longitudinal analysis of 5 years follow-up (Fig 1). We followed the response to ERT by assessment of muscle strength (MRC sumscore, HHD), muscle function (6-minute walking test, QMFT), respiratory function (FVC in sitting and supine positions), and daily life activities (R-Pact scale). The results were analyzed using linear mixed effects models. For some parameters, a difference was observed at the start of ERT between the II and the DD genotype groups, notably for R-Pact scale, QMFT, and FVC in sitting position, with a more favorable value for the II group. However, following multiple testing correction, these differences were not significant. The responses to ERT (i.e. the slopes of the curves during ERT treatment) were not different between the ACE genotype groups for all parameters tested. Individual responses to ERT varied within the ACE genotype groups. We found both good responders and non-responders to ERT in all three ACE genotype groups.

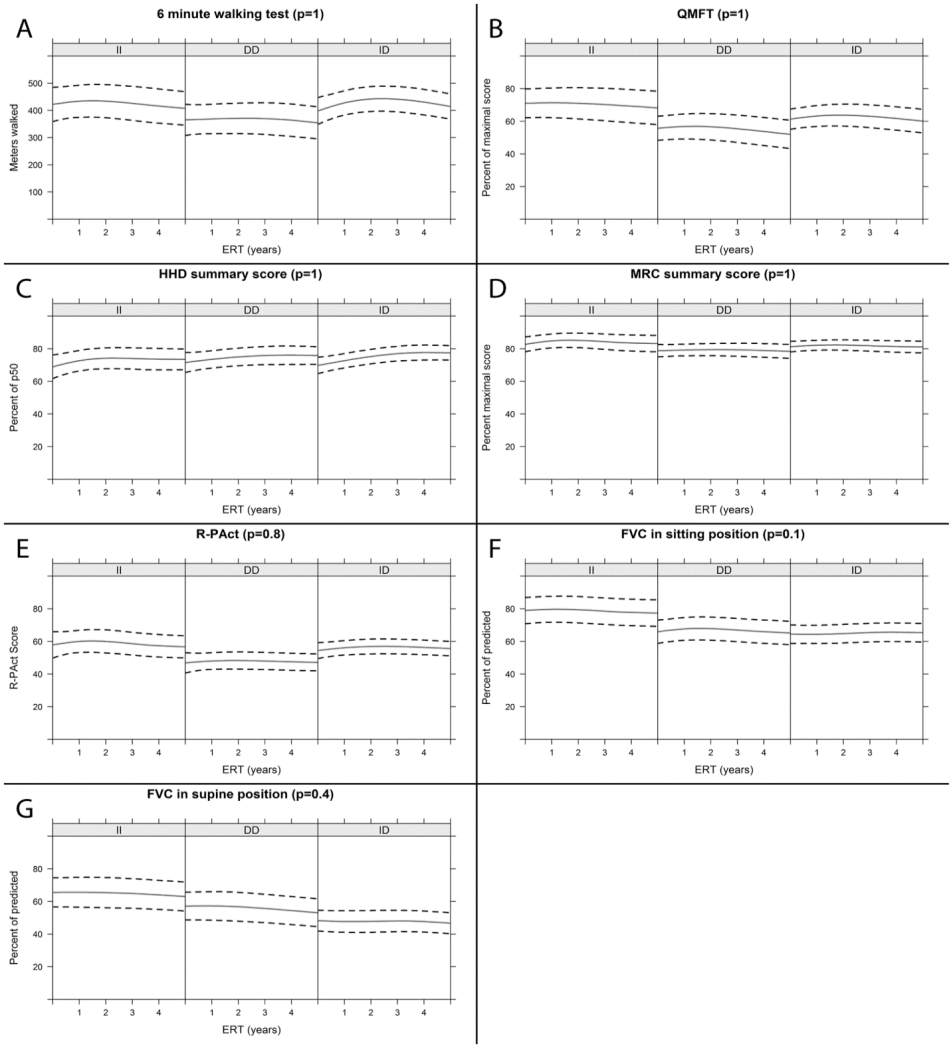


Fig 1. Predicted group means for outcome measures over 5 years of ERT treatment. Group means (continuous line) of the outcome measures and 95% prediction interval (area between the dotted lines) obtained for the II, ID and DD genotypes using linear mixed models. P values are indicated in the titles above the graphs.

ACE polymorphisms within families

Our cohort of patients included 54 siblings from 24 different families (2–3 siblings per family). Age at first symptoms, diagnosis, start of wheelchair use, ventilator use, ERT, death (if applicable), and ACE genotype are shown for each patient in Fig 2. Families are ordered by the onset of disease of the youngest family member. In 14 of the 24 families, siblings had discordant ACE polymorphisms, while 10 had the same ACE polymorphism. We found siblings with discordant ACE genotypes but with similar disease courses (8 siblings from 4 families; families 14, 19, 20 and 22), while we also

found siblings with the same ACE genotype but very different disease courses (21 siblings from 10 families; families 5, 6, 9, 10, 11, 13, 15, 16, 18 and 24). We conclude that no clear influence of the ACE genotype on onset of disease symptoms can be detected within these families.

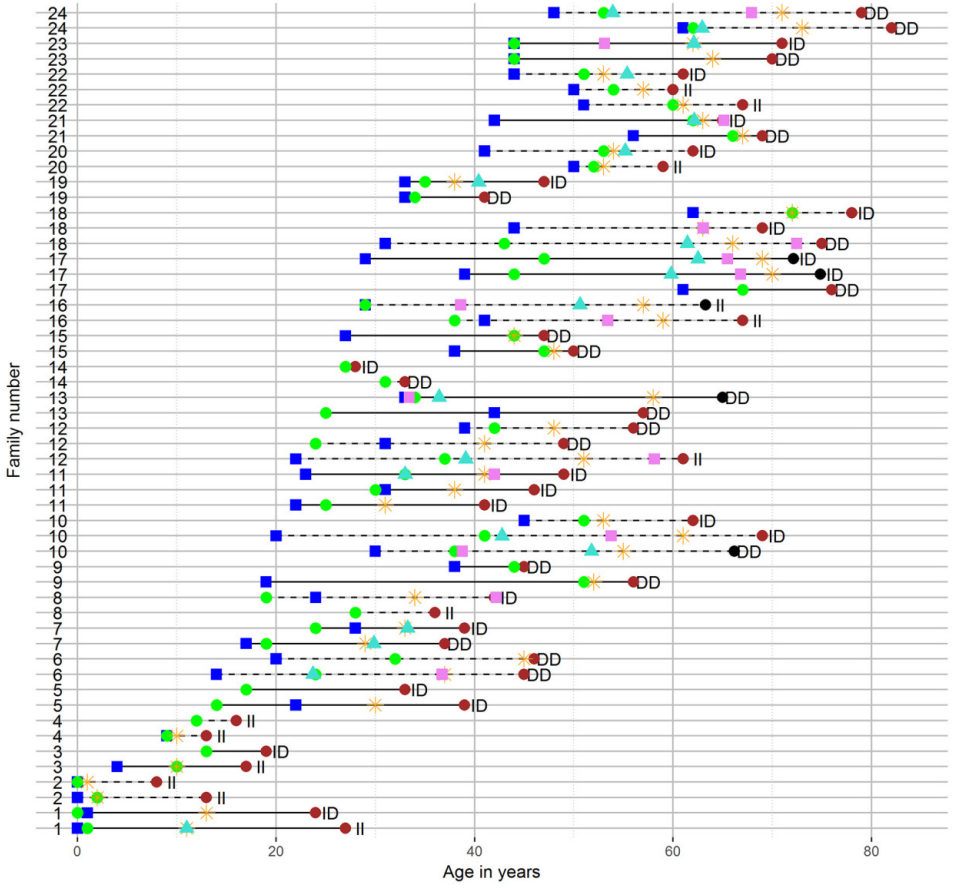


Fig 2. ACE polymorphisms and clinical parameters in family members with Pompe disease. Each line represents one patient. Families are numbered and visualized using alternating straight and dotted lines. The ACE genotype of individual patients is indicated at the right side of each patient's line. Onset of clinical events is plotted on the X axis and indicated with the symbols indicated on the right.

DISCUSSION

Pompe patients with the IVS1 variant represent the largest patient group of European descent with childhood or adult disease onset. These patients have a particularly large variation in phenotype, with symptom onset ranging from 0 to 62 years of age [5, 15, 40]. Pompe patients in general have a variable response to ERT, ranging from good to moderate to non-responders [19, 21, 41]. Here, we investigated the potential contribution of the ACE I/D genotype as a modifier of Pompe disease and response to ERT. Our results indicate that such contribution, if it exists, is too small to explain the phenotypic heterogeneity in this patient group. This applies both to differences in disease severity and to the response to ERT. Analysis of siblings with the same IVS1 genotype also did not point to a contribution of ACE genotype to clinical variation within families. These results are discussed in the light of other reports and with respect to the statistical power needed to identify a genetic modifier for Pompe disease.

A priori, at least two categories of genetic factors can be envisioned as potential modifiers of the disease course in Pompe disease. The first is a modifier of splicing, because the IVS1 variant is a splicing variant. For example, in theory it is possible that polymorphisms in splicing factors affect the amount of leaky wild type splicing by the IVS1 variant. If such polymorphism would exist, it would likely be present in a general splicing factor that is shared between fibroblasts and skeletal muscle cells, because these two cell types show similar aberrant splicing patterns [42]. Another category might be a genetic factor that modifies skeletal muscle function, as skeletal muscle is affected in Pompe disease. The ACE I/D genotype falls in this category, as it has been associated with performance of top athletes. Depending on the type of sport and its requirement for either endurance or strength, groups of top athletes are either enriched in the I or the D ACE alleles, respectively [43–49]. Considering the exceptionally broad clinical phenotype of patients with the IVS1 variant, a strong modifier is expected to explain the large variation in disease severity. Alternatively, modulation of the phenotype may occur via a combination of genetic factors with small effects that, when combined, have a strong effect on muscle function. In this scenario, the ACE polymorphism might be one of multiple factors that, when combined, affect the clinical phenotype.

The first study on the effect of the ACE I/D genotype on the natural course of Pompe patients was published in 2010. This study included 38 patients, 36 of whom contained the IVS1 GAA variant [22]. Sixteen parameters were tested in cross sectional analyses. The results suggested significant differences for some parameters, notably an association between the DD genotype and a worse Walton score (which scores muscle function),

an earlier disease onset, more muscle pain, and higher CK levels were found. In 2014, the initial study was extended to 85 patients that contained the IVS1 variant [23]. Associations were found between the DD genotype and pain, but in contrast to the previous study, no associations between ACE genotype and Walton score, CK levels or other clinical parameters were found in cross sectional analyses. Another study was published in 2016, in which 58 patients that had previously been included in the treatment arm of the late onset placebo controlled multicenter enzyme replacement study (LOTS study) were analyzed [25]. In this study, very mildly affected and severely affected patients were excluded. This showed no association between ACE genotype and any parameter including onset of first symptoms, disease duration, 6MWT, or FVC rating disease severity. In the present study, no inclusion criteria were applied other than the presence of the IVS1 variant combined with a null allele and a confirmed diagnosis of Pompe disease. Given the conflicting results published so far, we aimed to test the effect of the ACE genotype in a relatively large patient cohort of 131 patients including both children and adults of various ages and different disease severities. This revealed no significant effects of the ACE genotype on any of the parameters tested. It should be noted that initial statistical analysis suggested some significant differences between different ACE genotype groups, but that these turned into non-significant p-values after multiple testing correction, a method that has not always been applied in previous studies. When we only included adult patients, there were no significant effects of the ACE genotype on any parameter tested (data not shown). Taken together, the initial idea that the ACE DD genotype may be associated with faster symptom onset and more severe muscle symptoms could not be confirmed.

Two previous studies have investigated the influence of the ACE I/D genotype on the response to treatment with ERT. In a study on 16 patients with the IVS1 allele that were treated for > 2 years with ERT, the DD genotype (n = 3) was associated with reduced muscle mass over the course of treatment, while no associations were found for muscle strength, FVC, or 6MWT [24]. In the previously mentioned LOTS study, in which 58 moderately affected patients were treated with ERT for 78 weeks, the DD genotype (n = 17) was associated with a poorer response to ERT of FVC in sitting position [25]. One other parameter was tested, namely the 6MWT, and this showed a better response in patients with the ID genotype and a trend toward a better response in patients with the II genotype compared to the DD genotype. Other parameters for muscle strength and function were not reported in this study. The overall conclusion from the LOTS study was therefore that the II genotype may be associated with a better response to ERT. In the present study, 112 patients were treated with ERT with a follow-up of 5 years after start of ERT, and severely affected patients were also included. Patients with the DD

genotype in general did worse compared to the II genotype for several parameters, including 6MWT, QMFT , FVC, and R-Pact scale. However, the effects were small and not statistically significant. Altogether, the data available to date do not support the idea that the ACE genotype can explain the heterogeneous response to ERT in juvenile and adult Pompe patients.

The question arises whether we would have been able to detect an effect of the ACE genotype in our patient cohort, given the number of patients tested in relation to the heterogeneity of clinical outcome. To address this, we performed a post-hoc power analysis. We performed the calculation for the hypothetical situation in which the ACE genotype is the only modifying factor that is responsible for phenotypic variation in patients with the IVS1 variant. Although we realize that this may not be the case, the calculation has been performed in this way to give a sense of the number of patients required. As a primary outcome measure, we used symptom onset, because this outcome measure shows a documented large variation among IVS1 patients [17]. In our patient cohort, we calculated that 12 patients would be sufficient to demonstrate a significant effect on natural course (power = 0.96). In this study, we included 131 patients with Pompe disease, arguing that we should have detected an effect of the ACE polymorphism in our cohort if it would exist. We conclude that a possible effect of the ACE I/D genotype on natural course or the clinical response to ERT, if any, is very small, and remains undetected in our patient cohort. The search for modifying factors that can explain phenotypic variation in Pompe disease should continue.

Data Availability

Data cannot be shared publicly because of privacy rules concerning patients' data. Data are available from the Erasmus MC Ethics Committee (e-mail: metc@erasmusmc.nl, phone: +31 10 70 34428, address: Erasmus MC, Dr. Molewaterplein 40, 3015 GD, Rotterdam, Room Ae-337) for researchers who meet the criteria for access to confidential data.

Funding

Research on Pompe disease at Erasmus MC is financially supported by "Prinses Beatrix Spierfonds" [projectnumber W.OR13-21, W.OR15-10, W.OR16-07]; Tex Net; Sophia Foundation for Medical Research (SSWO) [projectnumber S17-32]; Metakids [projectnumber 2016-063]; Conselho Nacional de Desenvolvimento Científico e Tecnológico—"National Counsel of Technological and Scientific Development", Brasil (PI); and Colciencias. This project has received funding from the Ministry of Economic Affairs under TKI-Allowance under the TKI-programme Life Sciences & Health (project

number LSHM16008). The research described here has been funded in part through a collaboration agreement between Erasmus MC and Sanofi Genzyme. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests

ATvdP has provided consulting services for various industries in the field of Pompe disease (Sanofi Genzyme, Amicus, Spark therapeutics, Pharming, Biomarin, Audentes) under an agreement between these industries and Erasmus MC, Rotterdam, the Netherlands. All the other authors declare no conflict of interest. This does not alter our adherence to PLOS ONE policies on sharing data and materials.

Acknowledgments

We would like to thank the patients (and their parents) for participating in the study and the medical, paramedical and research staff who assisted with the analysis and assessments.

REFERENCES

1. Hirschhorn R R A. Glycogen storage disease type II: acid alpha-glucosidase (acid maltase) deficiency In: Scriver CR B A, Sly WS, Valle D., editor. The Metabolic and Molecular Bases of Inherited Disease. 8th edition ed New York: McGraw-Hill; 2001. p. 3389–420.
2. van den Hout HM, Hop W, van Diggelen OP, Smeitink JA, Smit GP, Poll-The BT, et al. The natural course of infantile Pompe's disease: 20 original cases compared with 133 cases from the literature. Pediatrics. 2003;112(2):332–40. Epub 2003/08/05.
3. Kishnani PS, Hwu WL, Mandel H, Nicolino M, Yong F, Corzo D, et al. A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease. J Pediatr. 2006;148(5):671–6. Epub 2006/06/02. 10.1016/j.jpeds.2005.11.033.
4. Beek NA, Hagemans ML, Reuser AJ, Hop WC, Ploeg AT, Doorn PA, et al. Rate of disease progression during long-term follow-up of patients with late-onset Pompe disease. Neuromuscul Disord. 2009;19 10.1016/j.nmd.2008.11.007.
5. Beek NA, Vries JM, Hagemans ML, Hop WC, Kroos MA, Wokke JH, et al. Clinical features and predictors for disease natural progression in adults with Pompe disease: a nationwide prospective observational study. Orphanet J Rare Dis. 2012;7.
6. Laforet P, Nicolino M, Eymard PB, Puech JP, Caillaud C, Poenaru L, et al. Juvenile and adult-onset acid maltase deficiency in France: genotype-phenotype correlation. Neurology. 2000;55(8):1122–8. Epub 2000/11/09.
7. Boerkoel CF, Exelbert R, Nicastri C, Nichols RC, Miller FW, Plotz PH, et al. Leaky splicing mutation in the acid maltase gene is associated with delayed onset of glycogenosis type II. Am J Hum Genet. 1995;56(4):887–97.
8. van der Wal E, Bergsma AJ, Pijnenburg JM, van der Ploeg AT, Pijnappel W. Antisense Oligonucleotides Promote Exon Inclusion and Correct the Common c.-32-13T>G GAA Splicing Variant in Pompe Disease. Mol Ther Nucleic Acids. 2017;7:90–100. Epub 2017/06/19. 10.1016/j.omtn.2017.03.001.
9. Chen L, Bush SJ, Tovar-Corona JM, Castillo-Morales A, Urrutia AO. Correcting for differential transcript coverage reveals a strong relationship between alternative splicing and organism complexity. Mol Biol Evol. 2014;31(6):1402–13. Epub 2014/04/01. 10.1093/molbev/msu083.
10. Jangi M, Sharp PA. Building robust transcriptomes with master splicing factors. Cell. 2014;159(3):487–98. Epub 2014/11/25. 10.1016/j.cell.2014.09.054.
11. Lee Y, Rio DC. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. Annu Rev Biochem. 2015;84:291–323. Epub 2015/03/19. 10.1146/annurev-biochem-060614-034316.
12. Bergsma AJ, Kroos M, Hoogeveen-Westerveld M, Halley D, van der Ploeg AT, Pijnappel WW. Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. Hum Mutat. 2015;36(1):57–68. 10.1002/humu.22705.
13. Kroos MA, Van der Kraan M, Van Diggelen OP, Kleijer WJ, Reuser AJ, Van den Boogaard MJ, et al. Glycogen storage disease type II: frequency of three common mutant alleles and their associated clinical

phenotypes studied in 121 patients. J Med Genet. 1995;32(10):836–7. Epub 1995/10/01.

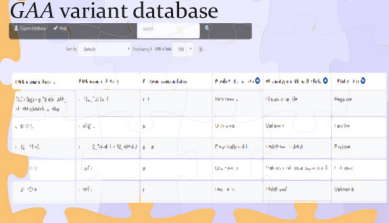
14. Kroos M, Hoogeveen-Westerveld M, van der Ploeg A, Reuser AJ. The genotype-phenotype correlation in Pompe disease. Am J Med Genet C Semin Med Genet. 2012;160C(1):59–68. Epub 2012/01/19. 10.1002/ajmg.c.31318.
15. Kroos MA, Pomponio RJ, Hagemans ML, Keulemans JL, Phipps M, DeRiso M, et al. Broad spectrum of Pompe disease in patients with the same c.-32-13T->G haplotype. Neurology. 2007;68(2):110–5. 10.1212/01.wnl.0000252798.25690.76.
16. van Capelle CI, van der Meijden JC, van den Hout JMP, Jaeken J, Baethmann M, Voit T, et al. Childhood Pompe disease: clinical spectrum and genotype in 31 patients. Orphanet Journal of Rare Diseases. 2016;11(1):65 10.1186/s13023-016-0442-y.
17. Wens SC, Gelder CM, Kruijshaar ME, Vries JM, Beek NA, Reuser AJ, et al. Phenotypical variation within 22 families with Pompe disease. Orphanet J Rare Dis. 2013;8.
18. van der Ploeg AT, Kruijshaar ME, Toscano A, Laforet P, Angelini C, Lachmann RH, et al. European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. Eur J Neurol. 2017;24(6):768–e31. Epub 2017/05/10. 10.1111/ene.13285.
19. Kuperus E, Kruijshaar ME, Wens SCA, de Vries JM, Favejee MM, van der Meijden JC, et al. Long-term benefit of enzyme replacement therapy in Pompe disease: A 5-year prospective study. Neurology. 2017;89(23):2365–73. 10.1212/WNL.0000000000004711.
20. de Vries JM, Kuperus E, Hoogeveen-Westerveld M, Kroos MA, Wens SC, Stok M, et al. Pompe disease in adulthood: effects of antibody formation on enzyme replacement therapy. Genet Med. 2017;19(1):90–7. Epub 2016/07/01. 10.1038/gim.2016.70.
21. van der Meijden JC, Kruijshaar ME, Harlaar L, Rizopoulos D, van der Beek N, van der Ploeg AT. Long-term follow-up of 17 patients with childhood Pompe disease treated with enzyme replacement therapy. J Inherit Metab Dis. 2018. 10.1007/s10545-018-0166-3.
22. de Filippi P, Ravaglia S, Bembi B, Costa A, Moglia A, Piccolo G, et al. The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease. Genet Med. 2010;12(4):206–11. Epub 2010/03/24. 10.1097/GIM.0b013e3181d2900e.
23. De Filippi P, Saeidi K, Ravaglia S, Dardis A, Angelini C, Mongini T, et al. Genotype-phenotype correlation in Pompe disease, a step forward. Orphanet J Rare Dis. 2014;9:102 Epub 2014/08/12. 10.1186/s13023-014-0102-z.
24. Ravaglia S, De Filippi P, Pichiecchio A, Ponzio M, Saeidi Garaghani K, Poloni GU, et al. Can genes influencing muscle function affect the therapeutic response to enzyme replacement therapy (ERT) in late-onset type II glycogenosis? Mol Genet Metab. 2012;107(1–2):104–10. Epub 2012/06/19. 10.1016/j.ymgme.2012.05.016.
25. Baek RC, Palmer R, Pomponio RJ, Lu Y, Ma X, McVie-Wylie AJ. The influence of a polymorphism in the gene encoding angiotensin converting enzyme (ACE) on treatment outcomes in late-onset Pompe patients receiving alglucosidase alfa. Mol Genet Metab Rep. 2016;8:48–50. Epub 2016/08/05. 10.1016/j.ymgmr.2016.07.005.

26. van Diggelen OP, Oemardien LF, van der Beek NA, Kroos MA, Wind HK, Voznyi YV, et al. Enzyme analysis for Pompe disease in leukocytes; superior results with natural substrate compared with artificial substrates. J Inherit Metab Dis. 2009;32(3):416–23. Epub 2009/04/24. 10.1007/s10545-009-1082-3.
27. Kroos M, Pomponio RJ, van Vliet L, Palmer RE, Phipps M, Van der Helm R, et al. Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating. Hum Mutat. 2008;29(6):E13–26. Epub 2008/04/22. 10.1002/humu.20745.
28. Okumiya T, Keulemans JL, Kroos MA, Van der Beek NM, Boer MA, Takeuchi H, et al. A new diagnostic assay for glycogen storage disease type II in mixed leukocytes. Mol Genet Metab. 2006;88(1):22–8. Epub 2005/12/20. 10.1016/j.ymgme.2005.10.016.
29. van der Beek NA, Hagemans ML, van der Ploeg AT, van Doorn PA, Merkies IS. The Rasch-built Pompe-specific activity (R-PAct) scale. Neuromuscul Disord. 2013;23(3):256–64. 10.1016/j.nmd.2012.10.024.
30. Al-Awadhi AM, Hasan EA, Sharma PN, Haider MZ, Al-Saeid K. Angiotensin-converting enzyme gene polymorphism in patients with psoriatic arthritis. Rheumatol Int. 2007;27(12):1119–23. Epub 2007/04/19. 10.1007/s00296-007-0349-y.
31. Council MR. Aids to examination of the peripheral nervous system. London: Her Majesty's Stationary Office; 1976.
32. Beenakker EA, van der Hoeven JH, Fock JM, Maurits NM. Reference values of maximum isometric muscle force obtained in 270 children aged 4–16 years by hand-held dynamometry. Neuromuscul Disord. 2001;11(5):441–6. Epub 2001/06/19.
33. van der Ploeg RJ, Fidler V, Oosterhuis HJ. Hand-held myometry: reference values. J Neurol Neurosurg Psychiatry. 1991;54(3):244–7. Epub 1991/03/01.
34. van Capelle CI, van der Beek NA, de Vries JM, van Doorn PA, Duivenvoorden HJ, Leshner RT, et al. The quick motor function test: a new tool to rate clinical severity and motor function in Pompe patients. J Inherit Metab Dis. 2012;35(2):317–23. 10.1007/s10545-011-9388-3.
35. A. T. S. Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. Am J Respir Crit Care Med. 2002;166(1):111–7. Epub 2002/07/02. 10.1164/ajrccm.166.1.at1102.
36. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012;40(6):1324–43. 10.1183/09031936.00080312.
37. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 1979;6(2):65–70.
38. Pinheiro J BD, Debroy S, Sarkar D, {nlme}. Linear and Nonlinear Mixed Effects Models. 2015.
39. A Language and Environment for Statistical Computing [Internet]. R Foundation for Statistical Computing. 2015.
40. Ausems MG, Verbiest J, Hermans MP, Kroos MA, Beemer FA, Wokke JH, et al. Frequency of glycogen storage disease type II in The Netherlands: implications for diagnosis and genetic counselling. Eur J Hum Genet. 1999;7(6):713–6. 10.1038/sj.ejhg.5200367.

41. Van Der Ploeg AT, Kruijshaar ME, Toscano A, Laforet P, Angelini C, Lachmann R, et al. European recommendations for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a ten-year experience. Submitted. 2017.
42. van der Wal E, Bergsma AJ, van Gestel TJM, In 't Groen SLM, Zaehres H, Arauzo-Bravo MJ, et al. GAA Deficiency in Pompe Disease Is Alleviated by Exon Inclusion in iPSC-Derived Skeletal Muscle Cells. Mol Ther Nucleic Acids. 2017;7:101–15. Epub 2017/06/19. 10.1016/j.omtn.2017.03.002.
43. Tsianos G, Eleftheriou KI, Hawe E, Woolrich L, Watt M, Watt I, et al. Performance at altitude and angiotensin I-converting enzyme genotype. Eur J Appl Physiol. 2005;93(5–6):630–3. Epub 2004/12/04. 10.1007/s00421-004-1284-1.
44. Tsianos G, Sanders J, Dhamrait S, Humphries S, Grant S, Montgomery H. The ACE gene insertion/deletion polymorphism and elite endurance swimming. Eur J Appl Physiol. 2004;92(3):360–2. Epub 2004/05/13. 10.1007/s00421-004-1120-7.
45. Woods D, Hickman M, Jamshidi Y, Brull D, Vassiliou V, Jones A, et al. Elite swimmers and the D allele of the ACE I/D polymorphism. Hum Genet. 2001;108(3):230–2. Epub 2001/05/17.
46. Williams AG, Rayson MP, Jubb M, World M, Woods DR, Hayward M, et al. The ACE gene and muscle performance. Nature. 2000;403(6770):614 Epub 2000/02/25. 10.1038/35001141.
47. Woods DR, Montgomery HE. Angiotensin-converting enzyme and genetics at high altitude. High Alt Med Biol. 2001;2(2):201–10. Epub 2001/07/10. 10.1089/152702901750265305.
48. Folland J, Leach B, Little T, Hawker K, Myerson S, Montgomery H, et al. Angiotensin-converting enzyme genotype affects the response of human skeletal muscle to functional overload. Exp Physiol. 2000;85(5):575–9. Epub 2000/10/20.
49. Vaughan D, Brogioli M, Maier T, White A, Waldron S, Rittweger J, et al. The Angiotensin Converting Enzyme Insertion/Deletion Polymorphism Modifies Exercise-Induced Muscle Metabolism. PLoS One. 2016;11(3):e0149046 10.1371/journal.pone.0149046.

Pompe disease

GAA variant database



CHAPTER 6

CLINICAL DIVERSITY OF POMPE PATIENTS WITH c.-32-13T>G GENOTYPES INVESTIGATED BY GENE EXPRESSION PROFILING

Niño Martinez MY[1, 2, 3], de Faria, DOS[1,2, 3], Kroos M[1], Hoogeveen-Westerveld M [1], in 't Groen SLM[1,2,3], van den Hout JMP[4], van IJcken W[4], van der Ploeg AT[2,3], Pijnappel WWMP[1,2,3]

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

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

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

[4]Erasmus Center for Biomics, Erasmus MC University Medical Center, Rotterdam, The Netherlands.

Correspondence: WWMP Pijnappel, Department of Clinical Genetics, Erasmus University Medical Center, P.O.Box 2040, 3000 CA Rotterdam. The Netherlands E-mail: w.pijnappel@erasmusmc.nl

Manuscript in preparation

6

Clinical diversity of Pompe patients with c.-32-13T>G genotypes investigated by gene expression profiling

Niño Martinez MY, de Faria, DOS, Kroos M, Hoogeveen-Westerveld M, in 't Groen SLM, van den Hout JMP, van IJcken W, van der Ploeg AT, Pijnappel WWMP.

ABSTRACT

Background: Pompe disease is an autosomal recessive neuromuscular disorder caused by acid α-glucosidase (GAA) deficiency. The disease presents as a clinical spectrum with regard to age of onset, symptom severity and rate of disease progression. Notably, clinical diversity occurs also among patients with *GAA* genotypes that exert the same functional effect on GAA activity. This observation points to the interaction of factors co-determining the clinical course of Pompe disease. We hypothesized that differences in gene expression profiles result in modulation of secondary responses and outcomes of disease progression while the level of GAA deficiency is the same. The differentially expressed genes (DEGs) provide opportunities for a better understanding of the molecular pathogenesis of Pompe disease and may at the same time serve as markers of disease progression.

Objective: The aim of our present study was to identify transcriptional differences in cultured fibroblasts from patients with childhood and adult onset Pompe disease.

Materials and methods: RNA sequencing (RNA-seq) was applied to cultured fibroblasts from patients with *GAA* genotypes (c.-32-13T>G / 'null', in 35/39 cases). The patients were selected on the basis of ages of onset. Seventeen had childhood onset (<16 years) and 22 adult onset (>35 years) Pompe disease. Patients and controls were analyzed in 2 separate analyses. Study 1 was the initial study, while study 2 was an independent study using individuals to validate the results of study 1. For comparison, we investigated gene expression also in fibroblasts from 19 childhood and 18 adult,

age-matched, unaffected individuals. Gene ontology analysis was performed using Panther and Ingenuity software.

Results: In study 1, the median age (years) at symptom onset was 1.25 in childhood onset patients, and 47 years in adulthood onset patients. These numbers were 7 and 33 in study 2, respectively. RNA-seq revealed 569 differentially expressed genes (DEGs) between childhood and adult Pompe patients in study 1. A substantial number of these DEGs were linked to known metabolic and cellular processes. The number of DGEs identified in the replicate study 2 was only 7, but a core set of 3 disease related DEGs could be established between childhood and adult phenotypes: *MAOA*, *Stc-1*, and *NRK-1*. When comparing expression profiles between Pompe patients and healthy controls, we found 6607 DEGs for adults and 4426 DEGs for children in study 1. For study 2 these numbers were 53 and 89, respectively. Importantly, only one gene overlapped when comparing genes that were differentially expressed in adults plus children with Pompe disease compared to age matched controls, which was *GAA*. This result strongly supports the validity of the chosen methodology and the data analysis.

Conclusion: Our study shows that fibroblasts can be used for detection of genome-wide transcriptional differences between patients with Pompe disease and unaffected individuals but also between Pompe disease phenotypes. However, large individual differences in gene expression within both healthy controls and patients prevents the reproducible identification of most disease-regulated genes. *MAOA*, *Stc-1*, and *NRK-1* were the exceptions and might be associated with the onset of symptoms in Pompe disease. This should be confirmed in larger cohorts.

Keywords: Pompe disease, clinical heterogeneity, disease-marker, differential gene expression, RNA sequencing, c.-32-13T>G (IVS1)

INTRODUCTION

Pompe disease (MIM# 232300) or glycogen storage type II (GSDII) is an autosomal recessive disorder caused by acid α-glucosidase (GAA; NP_000143.2) deficiency, a lysosomal enzyme degrading glycogen to glucose. The combination of disease-associated variants in the two copies of the *GAA* gene causes deficiency or total absence of GAA enzyme activity resulting in intra-lysosomal accumulation of glycogen with profound degenerative effects on muscle cells. Continuous accumulation of glycogen prohibits the normal functioning of lysosomes, disturbs the autophagic

transport pathways, interferes with cellular metabolism, and may even lead to rupture of lysosomes, all together resulting in cellular damage and progressive degeneration of skeletal muscles. Pompe disease presents as a spectrum of phenotypes. Apart from the rapidly fatal classic infantile phenotype, characterized by death within the first year of life due to cardiorespiratory failure [1], there are more slowly progressive phenotypes exemplified by children and adults presenting with skeletal and/or respiratory muscles weakness at older age [2-6].

In 2006, enzyme replacement therapy (ERT) using recombinant DNA technology was approved for the treatment of Pompe disease based on the outcome of clinical trials in infants [7, 8]. Currently, ERT is commercially available as the only treatment for Pompe disease. Studies in infants have demonstrated that human recombinant GAA (rhGAA) is capable of improving cardiac and skeletal functions and longer-term survival by reducing lysosomal glycogen accumulation [7, 9, 10]. Studies focusing on affected children and adults have shown that ERT, at the group level, improves muscle function (distance walked during 6 minutes), muscle strength and pulmonary function, and has long-term benefit [8, 11-13]. However, patients' response to ERT with rhGAA is highly variable [13, 14]. This variation could be due to a multitude of factors among which, obviously, the severity of symptoms at age of diagnosis and the disease duration at start of treatment. Also, other factors play a role such as antibody formation against the administered rhGAA, which is generally considered as a negative outcome predictor and in most cases but not all associated with a CRIM negative status [15-17].

Optimal management of patients with Pompe disease depends on early diagnosis, intelligent prediction of disease progression, coordinated follow up, and timely initiation of ERT. However, other than the patients' *GAA* genotype, their level of residual GAA activity, and their clinical signs, there are hardly any biochemical/molecular disease-markers to be guided by. Currently, Glucose tetrasaccharide (Glc4) is used as biomarker in Pompe disease; patients with Pompe disease typically excrete increased amounts of Glc4 in urine, which can be analyzed by high performance liquid chromatography (HPLC). A decrease in the level of Glc4 has been reported in patients responding to ERT. However, raised levels of Glc4 have been observed also in patients with glycogen storage diseases (GSDs) type Ia and type III suggesting that Glc4 is not a highly sensitive marker for Pompe disease [18]. Myostatin and insulin- like growth factor 1 (*IGF-1*) have been proposed as potential therapeutic biomarkers in Pompe disease as well. Those molecules have shown lower serum levels in Pompe patients than in healthy controls. After ERT, the expression levels of both myostatin and IGF-1 reached normal levels, suggesting possible utility for disease monitoring [19].

Recent studies have suggested a group of miRNAs as potential biomarkers of Pompe disease. The validation of the results identified three miRNAs, suggesting that may represent additional biomarkers for the follow-up of adult onset Pompe disease [20, 21]. At present there are no established methods that can facilitate and provide a better understanding of decision-making on therapeutic intervention. This difficult task has prompted the need to develop methods for prediction of disease progression and start of therapeutic intervention.

In our search for genes and associated pathways that could potentially affect the clinical course of Pompe disease and serve as biochemical / molecular disease-marker we analyzed the transcriptome of cultured skin fibroblasts from two groups of patients with functionally 'the same' *GAA* genotype but quite different age of onset of clinical symtoms, patients with childhood onset and patients with adult onset. Age matched healthy individuals were analyzed as controls. To enable the collection of a sufficiently large patient cohort with functionally 'the same' *GAA* genotypes we have chosen in our study for affected children and adults with the common c.-32-13T>G/disease-associated variant genotypes (-32-13T>G/ null, in 35/39 cases) Patients with the IVS1/ null genotype may present with symptoms and any age from early childhood to late adulthood [22].

The c.-32-13T>G transversion upstream of the splice acceptor site of exon 2 results in aberrant splicing whereby only 5-15% of the normal amount of structurally and functionally normal acid α-glucosidase is produced. c.-32-13T>G is the most frequent *GAA* variant among Caucasians with Pompe disease and is observed in more than two-thirds of the patients worldwide [23-25]. In the Netherlands, about 95% of adults with Pompe disease and 68% of affected children under 18 years have the c.-32-13T>G / 'null' genotype [22, 23].

MATERIAL AND METHODS

Study population

Inclusion criteria and study design

An overview of the entire experimental setting performed in study 1 and study 2 is provided in Figure 1.

Patients had to be diagnosed with Pompe disease during childhood or adulthood. For genome-wide expression analysis, all patients (39) had to be compound heterozygote for

the common c.-32-13T>G *GAA* gene variant in combination with a disease-associated variant on the second *GAA* allele (which was null in 35 cases and presumably null in 4 cases). Patients with substantially different ages of onset (<8 years vs >35 years) were selected for the study. Cohort 1 comprised 17 patients with childhood onset of Pompe disease (13 male, 4 female) with a median age at symptom onset of 2 years (range 0.5-13). Cohort 2 consisted of 22 patients with adult onset Pompe disease (8 male, 14 female) with a median age at onset of symptoms of 41.5 years (range 20-62). In addition, a total of 37 unaffected individuals were included in the study: 19 children (12 males, 7 females) with a median age at time of skin biopsy of 9 years (range 1.2-14.2), and 18 adults (9 males, 9 females) with a median age at time of biopsy of 52.5 years (range 35-76.9). Biopsies were taken from either the skin of the proximal part of the leg (patients) or form the distal part of the leg or from the foreskin (controls). Characteristics of study subjects and study materials including ages at biopsy) are provided in Table 1 and Figure 2A. To further validate the results obtained a skin biopsy, a third cohort of Pompe patients (3 childhood, 3 adult and 1 classic infantile) were analyzed. The Medical Ethical Committee at Erasmus University MC approved the study protocol, and all patients, or their parents or legal guardians provided written informed consent.

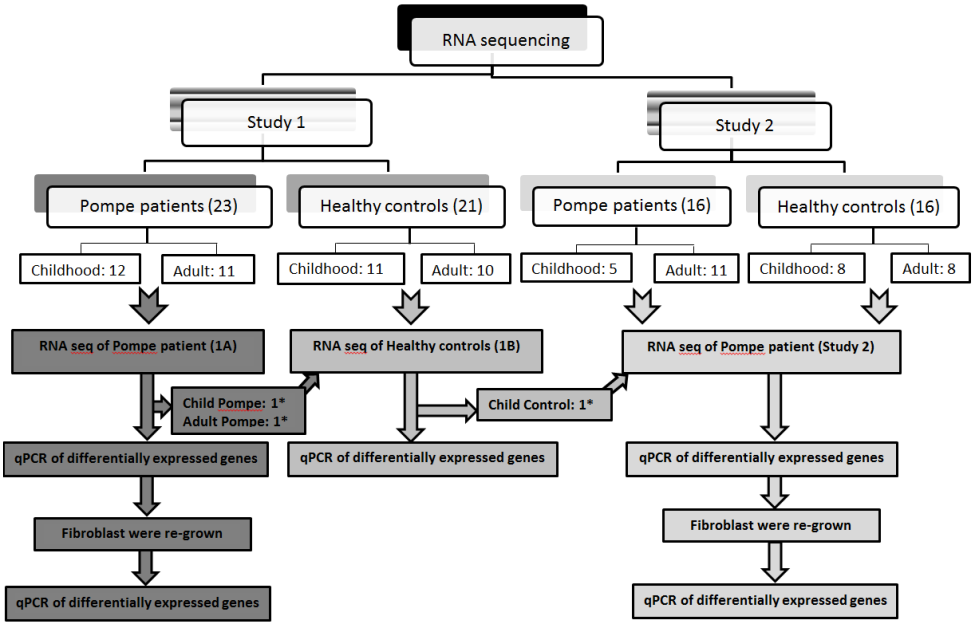


Figure 1. Workflow of the generic approach for the RNAsequencing analysis in Pompe patients and healthy controls included in studies 1 and 2. The differentially expressed genes (DEGs) were analyzed by qPCR. Afterwards, skin fibroblasts derived from Pompe patients were re-grown (1A and 2), and qPCR of DEGs as performed to confirm the RNAseq findings. *Samples used as internal control.

Table 1. Demographics of Patients and Controls in studies 1 and 2.

Criteria	Childhood Pompe	Adult Pompe	Childhood control	Adult control
Age:				
at onset (y)	<8*	>35#		
at biopsy (y)	<16	>35	<16	>35
Gender:				
male	13	8	12	9
female	4	14	7	9
Biopsy:				
leg	17	22	13	18
foreskin			6	
Total	17	22	19	18

*Patients seen by doctors at indicated age (<8 years) with clear signs of Pompe disease. #Patients with first signs at indicated age (>35 years) and seen by doctors at that age.

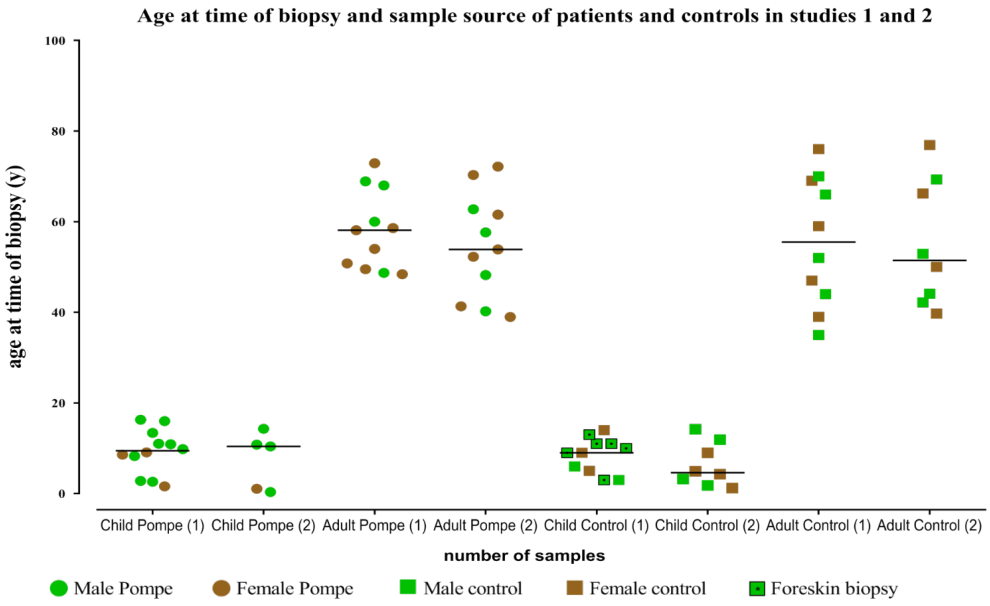


Figure 2A. Patient's age at time of biopsy and sample source. Skin biopsies were taken from the upper lower limb or the lower limb except for 6 samples whereby served as sample source (foreskin fibroblasts are represented as dotted green square). The study number is given within brackets. The horizontal line presents the median value.

For several reasons it was not possible to incorporate all RNAseq measurements in a single study. Therefore, the analyses were divided as follows: study 1 (all from cohort 1) was performed in two experiments, study 1A (12 children with evident symptoms under the age of 8 years and 11 adults with symptoms over 35 years); and study 1B (11 unaffected children and 10 unaffected adults with skin biopsies taken under the age

of 16 or over the age of 35, respectively), 3 samples were chosen for measuring RNA profiles in both study 1A and study 1B to determine a potential batch effect, which was absent. Study 2 (all from cohort 2) included 8 unaffected and 5 affected children, and 8 unaffected and 11 affected adults with skin biopsies taken under the age of 16 (childhood) or over the age of 35 (adult), respectively (Fig. 1 and Fig. 2A).

Cell culture, GAA enzyme activity and glycogen content

Skin fibroblasts were cultured according to routine procedures in medium containing DMEM supplemented with 10% fetal calf serum and 1% penicillin/streptomycin, in a humidified atmosphere containing 5% CO_2/95% air, at 37º C. The cells were harvested with trypsin/EDTA (ethylenediaminetetraacetic acid) at day 3 after the last subculture and stored as pellet at -80°C until use. Cell lysates were prepared and assayed for GAA activity and protein, as mentioned below, and for measuring the glycogen content according to Umapathysivan et al, 2005; whereby the cells were kept for 3 to 5 days on glucose-free medium before harvesting [26].

The standard GAA enzyme activity assay mixture consisted of 10 µL of cell homogenate (10 µg of protein) and 20 µL of 2.2 mmol/L 4-methylumbelliferone (4-MU)–α-D-glucopyranoside (4MUG, Sigma) in 0.2 mol/L Na-acetate buffer (pH 4.0) with 0.02% (wt / vol) sodium azide. The reaction mixtures were incubated for 1 hour at 37°C, and the reaction was stopped by the addition of 200 µL 0.5 mol/L Na_2CO_3/$NaHCO_3$ (pH 10.7). The fluorescence of released 4MU was measured with a fluorimeter (Thermo Electron corporation). The protein concentration of the samples was determined using Pierce™ BCA Protein Assay Kit (Thermo Scientific™) following the instructions provided by manufacturer and previously described [27]. The GAA enzyme activity assays and the protein assays of study 1 (see Results) were performed at time of this study. The GAA enzyme activities per mg protein of study 2 were obtained using exactly the same routine procedures, but were derived from the patients' medical records.

RNA sample collection for RNA sequencing and qPCR

Skin fibroblasts were cultured according to routine procedures in DMEM medium supplemented with 10% fetal calf serum and 1% penicillin/streptomycin. Fully confluent tissue culture flasks were split and harvested 3 days later for RNA isolation using the RNAeasy miniprep kit with on-column DNase treatment according to the manufacturer's recommendations (Qiagen). For gene expression analysis, 800 ng of RNA was used for generation of cDNA using the iScript cDNA synthesis kit (Biorad) and real time qPCR. The cDNA solution was diluted 10 times before use. To determine the relative concentration of each sample, 4 µl of each cDNA sample (10 times diluted

in H_2O) was processed in a 15 µl PCR reaction containing IQ Mastermix (Biorad) and 0.333 µM of each primer. To account for the efficiency of each specific primer set, all samples were related to a standard curve from the healthy control sample. All samples were measured in triplicate and normalized using β-actin.

RNA isolation, cDNA library construction, Illumina deep sequencing and Data processing of RNA-Seq experiments

Total RNA was extracted using the RNAeasy miniprep kit following the manufacturer's protocol (Qiagen). A total of ≥ 100 ng RNA per sample was sent for library preparation using the TruSeq RNA sample preparation kit and sequenced according to the Illumina TruSeq v3 protocol on the HiSeq2000 with a single read 43 bp and 7bp index, and mapped against the requested reference using Tophat (version 2.0.10). We called gene expression values using Cufflinks (version 2.1.1). More than 10 million aligning reads per sample were generated for each sample. The gene/transcript annotation hg19_transcriptome was used. Raw data in the fastq format were processed. We applied three methods for differential gene expression analysis: T-test per gene on cufflinks FPKM (fragments per kilobase of transcript per million mapped reads) values, Limma on read counts, and EdgeR with a generalized linear model. RNA sequencing was performed by the Erasmus Center for Biomics at Erasmus MC (www.biomics.nl, Rotterdam, The Netherlands).

Differential Analysis: Pompe patients Child *vs* Adult and Controls

Differential expression analysis was performed using edgeR (R package) with a generalized linear mode and we called gene expression values using Cufflinks (version 2.1.1). Hierarchical cluster analysis and multidimensional scaling (MDS) plots were used to identify DEGs. DEGs were filtered with false discovery rate (FDR) <0.05, fold change (FC) >1.5 or <−1.5 in each pairwise comparison.

GO analysis

Gene ontology (GO) analysis was performed to determine the biological implications of the expression of unique genes in significant or representative profiles of genes that were differentially expressed using The PANTHER database (Protein Analysis Through Evolutionary Relationships) at http://pantherdb.org). The Bonferroni correction for multiple testing was applied to identify significant GO categories, and p-values of <0.05 were considered significant. The categories are described as GO molecular function (GOTERM_MF), GO biological process (GOTERM_BP), and GO cellular component (GOTERM_CC), GO panther pathways, and Reactome pathways. All enriched biological processes were ranked from top to bottom according to the

p-values for each GO term.

Pathway analysis was used to determine the significant pathway(s) of the DEGs according to the PANTHER database and Ingenuity software (IPA; Qiagen, Valencia, CA; http://www.ingenuity.com). The analysis was determined by adjusting p-values for multiple testing, including the Bonferroni and Benjamini-Hochberg correction methods (P<0.05), in order to identify the most significant pathways and causal relationships associated with the experimental data.

Statistical Analysis

We called gene expression values using Cufflinks (version 2.1.1).

The data were analyzed using IBM SPSS (version 25), and GraphPad prism (version 8.0.0). The category percentage and the basic statistics value were calculated for each group and/or study. Linear regression analysis was used to modeling the relationship between gene expression level and age of onset/biopsy. We performed the one-way and 2-way ANOVA: Turkey's multiple comparisons test for testing groups to see if there is a difference among and between them. Differences were regarded to be significant in all tests at a significance level of "P < 0.05"

The pathway and GO terms associated with the obtained DEGs were tested for significance using the Bonferroni correction and Benjamini-Hochberg correction for multiple testing (P < 0.05) and were subjected to gene ontology analysis

RESULTS

6

Figure 1 provides an overview of the experimental setting.

Clinical features and disease severity

Table 2 provides the characteristics of the patients in *cohort 1* (onset of symptoms under the age of 16 years) and *2* (onset of symptoms over the age of 35 years). The majority (76.5%) of affected children in cohort 1 was male. The median age at first symptoms was 2 years (0.5-13), and the median age at time of biopsy was 9.8 years (0.35-16.3). Among the adult patients in cohort 2, 36.4% are males. The median age at first symptoms was 41.5 years (20-62), and the median age at time of biopsy was 55.8 years (38.9-72.9). Table 2 also lists the number of *GAA* genotypes of all patients. With regard to the latter, all patients carry the common c.-32-T>G variant that combines in 6 (35%) of the childhood cases and in 12 (55%) of the adult cases with the c.525del

allele and in all other cases of childhood (53%) and adult onset (36%) with an equally deleterious "null" other disease-associated *GAA* variant, except for 4 patients (2 child (12%) and 2 adult (9%) that contain a 'non-null' *GAA* variant.

Table 2. Characteristics of patients reported in studies 1 and 2.

	Childhood Pompe	Adult Pompe	all patients
	Total	Total	Total
Patients with Pompe disease	17	22	39
Gender: males (%)[a]	13 (76.5)	8 (36.4)	21 (54)
Median age at first symptoms (y)[b]	2 (0.5-13)	41.5 (20-62)	22 (0.5-62)
Median diagnosis age (y)[b]	2.5 (0-14.1)	46.5 (26-72)	34 (0-72)
Median age at time of biopsy (y)[b]	9.8 (0.35-16.3)	55.8 (39-72.9)	41 (0.35-72.9)
Mobility (%)[a]			
Ambulant	16 (94)	14 (63.6)	30 (76.9)
Wheelchair dependent		2 (9.1)	2 (5.1)
Partially wheelchair dependent	1(5.9)	5 (22.7)	6 (15.4)
Walking aid		1 (4.5)	1 (2.6)
Ventilation dependent (%)[a]	1 (5.9)	6 (27.3)	7 (17.9)
GAA activity[b,c]	9.1 (6.2-29)	12 (5.2-18.7)	11.5 (5.2-29)
Patients and genotypes (%)[a]			
c.-32-13T>G/c.525del p.(Glu176Argfs*45)	6 (35)	12 (55)	18 (46)
c.-32-13T>G/ other 'null' *GAA* variant	9 (53)	8 (36)	17 (44)
c.-32-13T>G/ 'non-null', c.2135T>C p.(Leu712Pro)	2 (12)		2 (5)
c.-32-13T>G/ 'non null', c.1076-22T>G p.?		2 (9)	2 (5)

The figures are [a] percentages or [b] medians with ranges. [c] The control range for GAA activity in fibroblast is: 40–180 nmol MU/mg.h.

Figures 2B and 2C present the age at first symptoms, the age at diagnosis, the age at time of biopsy, and the disease duration for each individual patient, and Suppl. Figure 12 provides the ages at biopsy of controls. If known, the figures also include the age at which the patients became wheelchair or ventilator dependent and/or started to receive ERT. Patients sharing exactly the same *GAA* genotype are marked. The youngest of all patients developed symptoms in the first year of life, while the oldest developed symptoms at the age of 62 years. The age at diagnosis varies between 0 and 72 years (median 34 years). The age at time of skin biopsy varies between 0.35 and 72.9 years (median 41 years). Table 3 shows the median age at first symptoms of the childhood and adulthood onset patients with Pompe disease per study.

Figure 2B. Clinical course of patients in study 1. C1-C12 are patients with childhood onset Pompe patients and A1-A11 are adult onset Pompe patients. Dots represent an event during the course of the disease. Patients marked with * or **#** share the same *GAA* genotype. Patient C11 is partly wheelchair dependent, but can walk for 200 m. Patient Ch1 was wheelchair dependent but acquired walking ability during ERT.

GAA enzyme activity and glycogen accumulation

The GAA enzyme activity and glycogen accumulation assays were performed in fibroblasts of Pompe patients and healthy controls to investigate potential associations between residual GAA activity, glycogen accumulation, and clinical phenotype. Figure 2D showed no statistically significant differences between the GAA enzyme activity in fibroblasts from childhood vs adult patients that could explain differences in phenotype (P>0.05). For healthy controls a significant difference was observed between GAA activities of children and adults (P < 0.001). As expected, patients showed GAA deficiency compared to controls (P<0.001), with the following values: childhood patients: 9.1 [6.2-29], adult patients: 12 [5.2-18.7], childhood control: 98.95 [79.1-168], adult control: 70.3 [30.2-116] (control range of GAA activity in fibroblasts: 40-180 nmol MU/mg.h) [22].

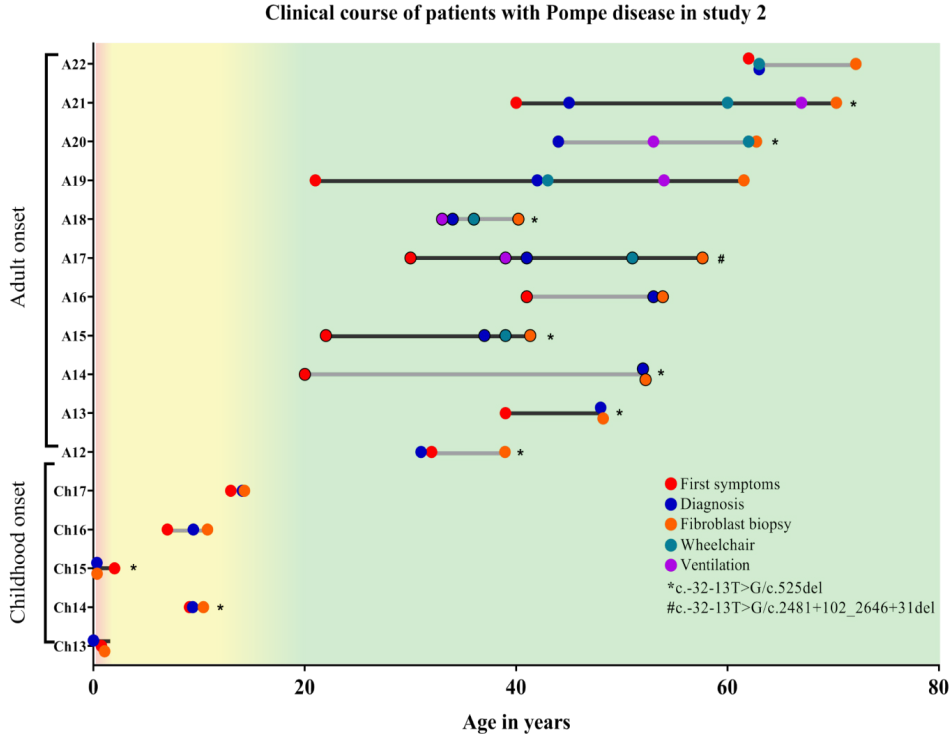


Figure 2C. Clinical course of patients in study 2. Patients Ch13-Ch17 had childhood onset and A12-A22 during adulthood. Dots represent an event during the course of the disease. Patients marked with * or **#** share the same *GAA* genotype. Patients A15, A18- A20 and A22 are partially wheelchair dependent.

In order to measure the lysosomal glycogen content of the fibroblast, a culture condition without glucose had to be applied for 3-5 days to deplete cytoplasmic glycogen. Before embarking on this assay, we investigated whether changing the culture conditions had any effect on the GAA activity (Suppl. Fig S1). There were no significant differences between the GAA activity of fibroblasts from affected children and adults (cohorts 1 and 2), neither between childhood controls and adulthood controls cultured in media with or without glucose for either 3 or 5 days' time (P>0.05). We measured the glycogen content of the fibroblasts from children with Pompe disease and unaffected children and adults on day 3 after culture in glucose free media and did not find statistically significant differences between the four sample types (Suppl. Fig. S2). The cells did not store glycogen, as opposed to fibroblasts from a single patient with classic infantile Pompe disease, which was included in the experiment as positive control and which did store glycogen with a 5-10 fold increase. Supplementary Figure S2 summarizes the results per individual sample and shows the median glycogen content per sample type in µg glycogen per mg protein (childhood patient: 37 [8.1-143], adult patient: 25 [9.5-60], childhood control: 58.4 [15-86], adult control: 56.5 [11-88]).

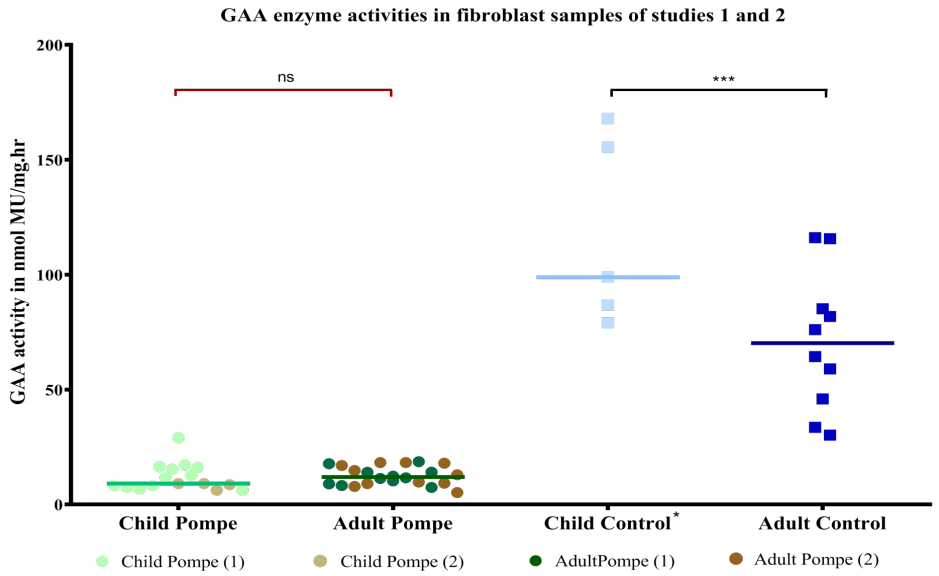


Figure 2D. GAA enzyme activities in fibroblast samples of studies 1 and 2. The horizontal line presents the median value. The study number is given within brackets. The difference between the childhood and adult onset groups is non-significant (ns; (P>0.05), but the difference between the patients (childhood plus adult onset) and controls were significantly different P<0.05 (**** <0.0001, *** indicate p 0.0001). Normal range of GAA activity in fibroblasts is 40–180 nmol MU/mg.h (4MUG as substrate was used). *Foreskin fibroblasts are not included in this figure. Enzyme activities for controls in study 2 were not performed. The one-way ANOVA test was used for the analysis.

Searching for DEGs: RNA-Seq analysis

Differentially expressed genes in childhood vs adult onset Pompe disease compared to childhood versus adult controls

RNA-Seq analyses were conducted to generate transcriptome profiles from fibroblasts of Pompe patients and healthy controls in order to identify genetic denominators that could possibly explain the difference between childhood and adult onset Pompe disease. We conducted pairwise comparison of materials from adult and childhood Pompe patients (study 1A) and adult and childhood controls (study 1B).

The Multidimensional scaling (2-D MDS) plot (Fig. 3A) revealed distinct expression profiles of Pompe patients and healthy controls included in the studies 1A-B. Adult Pompe patients (dark green) had different gene expression levels than childhood Pompe patients (light green), whereas adult (dark blue) compared to childhood (light blue) control samples did not. Quite unexpectedly, six childhood control samples (yellow) had completely different gene expression profiles than all other samples from either patients or controls. It then turned out that the deviant samples were from foreskin, while all

others were derived from skin biopsies of the upper leg. Without clear understanding, we were confronted with the fact that foreskin may have other characteristics than other types of skin fibroblasts. This was reason to exclude the results obtained with foreskin fibroblasts were excluded from all further analyses. The DEGs were filtered based on level fold change (FC) >1.5 or < –1.5, and false discovery rate (FDR) <0.05, in each pairwise comparison. The outcome of hierarchical cluster analysis of study 1A plus 1B (foreskin excluded) is depicted in Supplementary Figure S3A and showed a group-wise clustering of patients in the left branches of the tree and controls in the right branches. Healthy children and adults were randomly represented in the various sub-branches. Within the patient branches were sub-branches with adults or children only, but there was no complete separation of phenotypes. Similar results were obtained in study 2 (Supp. Fig. S3B) albeit that the branching pattern was not precisely the same as in study 1, and the separation between adult and childhood cases of Pompe disease even less evident, which suggested a very large variation in study 2 compared to study 1.

Pairwise comparisons of adult vs childhood cases of Pompe disease (study 1A) compared to adult vs childhood healthy controls (study 1B) revealed that 569 genes were significantly differentially expressed between the childhood compared with the adult patient cohorts and 6 between the childhood and adult controls (Fig. 3B[A]). The Venn diagram shows that only 1 gene, *TMEM26*, was shared between patients and controls. Of the 569 DEG's between fibroblasts of affected children compared to affected adults, 385 were down regulated (Fig. 3B[B] and 184 upregulated in adults (Fig. 3B-[C])

All DEG's that were identified in study 1 were subjected to further analysis using the Panther database and Ingenuity software. The pathway and GO terms associated with the obtained DEGs were subjected to gene ontology analysis. The GO enrichment of the GO terms suggested that the transcriptional levels of genes related to molecular function (such as aminoacyl-tRNA ligase activity catalyzing the formation of aminoacyl-tRNA and related compounds), and biological processes (such as the type I interferon signaling pathway; cellular response to type I interferon, response to type I interferon, response to virus, defense response to virus) were altered (Supp. Figs. S4A-B). Similar results were obtained when analyzing the data in the context of GO pathway category. In fact, the results of the GO enriched pathway analysis supported our findings obtained with ingenuity pathway analysis (Supp. Figs. S5A-B). Pathway terms that came out as being significant included several major processes such as i) amino acid metabolism and biosynthesis (tRNA charging and the super-pathway of serine and glycine biosynthesis I), ii) cell and organ morphology (axonal

guidance signaling, nitric oxide signaling in the cardiovascular system), iii) immune system response/processes (role of pattern recognition receptors in recognition of bacteria and viruses), and iv) hormone metabolism (thyronamine and iodothyronamine metabolism, thyroid hormone metabolism I via de-iodination).

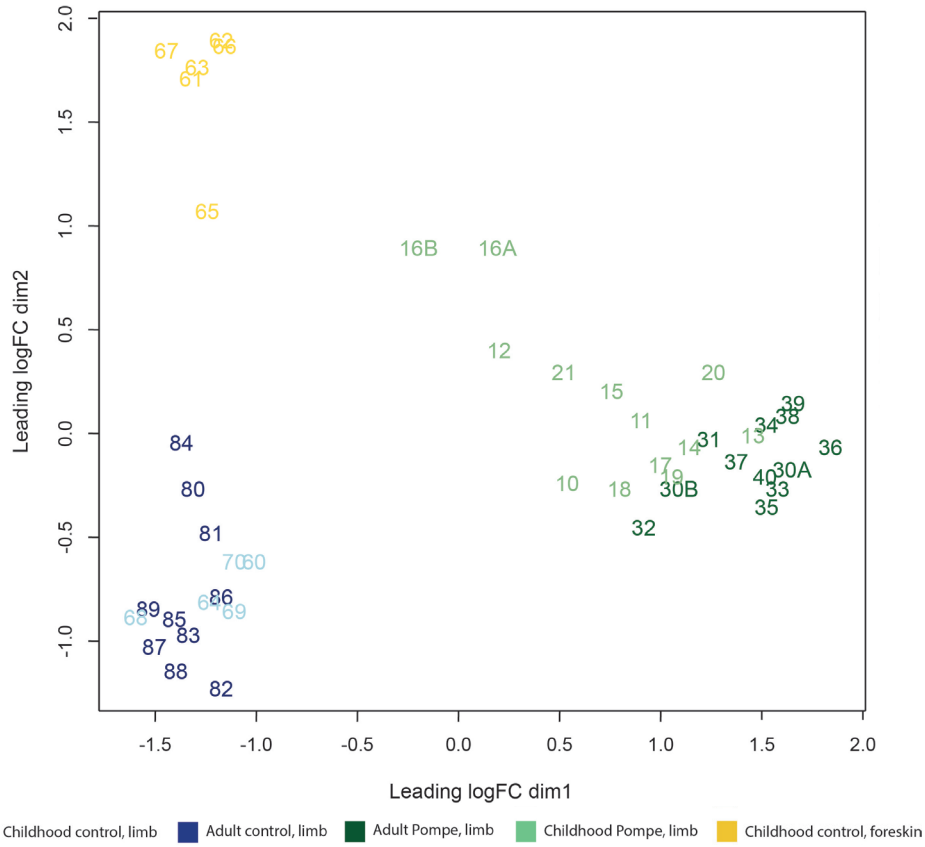


Figure 3A. Multidimensional Scaling Plot (2-D MDS) of the RNA-seq data for the patients and controls of study 1. The MDS plot shows different gene expression profiles between Pompe patients and healthy controls. The profiles of adult onset patients (dark green, n= 11: 30A, 30B, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40) differ from those of childhood onset patients (light green, n=12: 10, 11, 12, 13, 14, 15 ,16A, 16B, 17, 18, 19, 20, 21). In contrast, unaffected children (light blue, n=5: 60, 64, 68, 69, 70) and adult controls (dark blue, n=10: 80, 81, 82, 83, 84, 85, 86, 87, 88, 89) controls show similar expression profiles except for samples from 6 unaffected children (gold: 61, 62, 63, 65, 66, 67) that revealed a completely different profile. These were fibroblasts derived from foreskin. The stress for this plot is 13.5%. Of note, the patient numbers shown in this figure do not correspond with those in the tables (see Suppl. Table S5. for comparison).

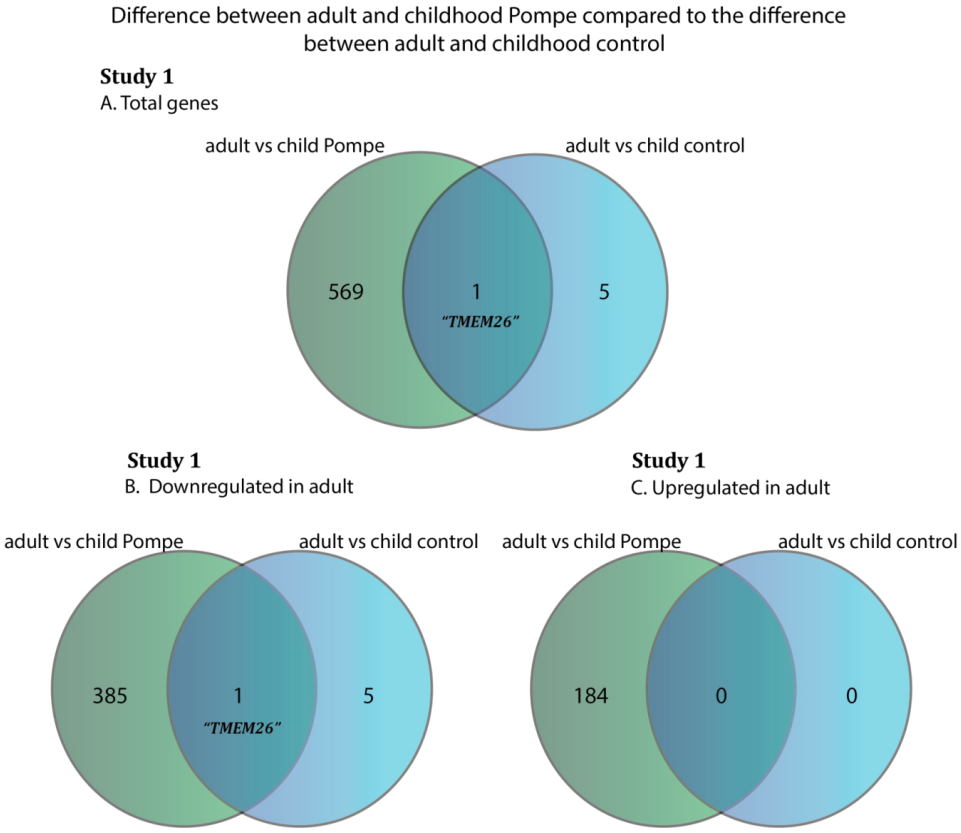


Figure 3B. Venn diagram showing the numerical difference between genes expressed in adult *versus* childhood onset Pompe disease (green) compared to the difference between adult *versus* childhood control (blue) of study 1. (A) A total of 569 genes are differentially expressed in childhood *versus* adult onset Pompe disease (corrected for controls), whereas only 5 in the comparison of unaffected children *versus* unaffected adults. **(B)** Of the 569, 385 are downregulated in adult onset Pompe disease, and **(C)** 184 upregulated. Foreskin fibroblasts were excluded. The DEGs were filtered using a cut off value of < -1.5 - >1.5 fold change (FC) (LogFC: +/- 0.584) and false discovery rate (FDR) of <0.05

6

qRT-PCR analysis was used to validate the quality of the RNA-Seq data. For this purpose, we selected 10 downregulated and 10 upregulated genes that were identified in study 1A-B. The relative gene expression levels of selected genes measured by qRT-PCR were in good agreement with the RNA-sequencing data (Supp. Figs 6A-B). The strong correlation between the RNA-Seq and qRT-PCR data indicates the reliability of our transcriptomic profiling data. Therefore, the culture condition dependency and reproducibility were analyzed by re-growing the fibroblasts of patients included in study 1A and the expression levels of down and upregulated genes were analyzed by qRT-PCR. The majority of those genes are in line with our previous results by qRT-PCR.

The expression of some genes seemed culture condition dependent (Supp. Figs 6C-D). The same approach was taken to evaluate the outcome of study 2. The Venn diagram of this experiment revealed only 7 genes differentially expressed between adult onset and childhood onset Pompe patients, and 80 in childhood versus adult unaffected controls (Fig. 4A[A]). Of those 7 genes, controls and patients share 3, and all three are downregulated in adult compared to childhood Pompe disease (Fig. 4A[B]). Four genes are upregulated (Fig. 4A[C]). Although, the DEGs in study 2 are not in line with our previous results (study 1), we performed RT-qPCR of genes selected for study 1. Conversely, the majority of these genes showed the same trend of down and upregulation as in study 1. However, they were not statistically significant (data not shown).

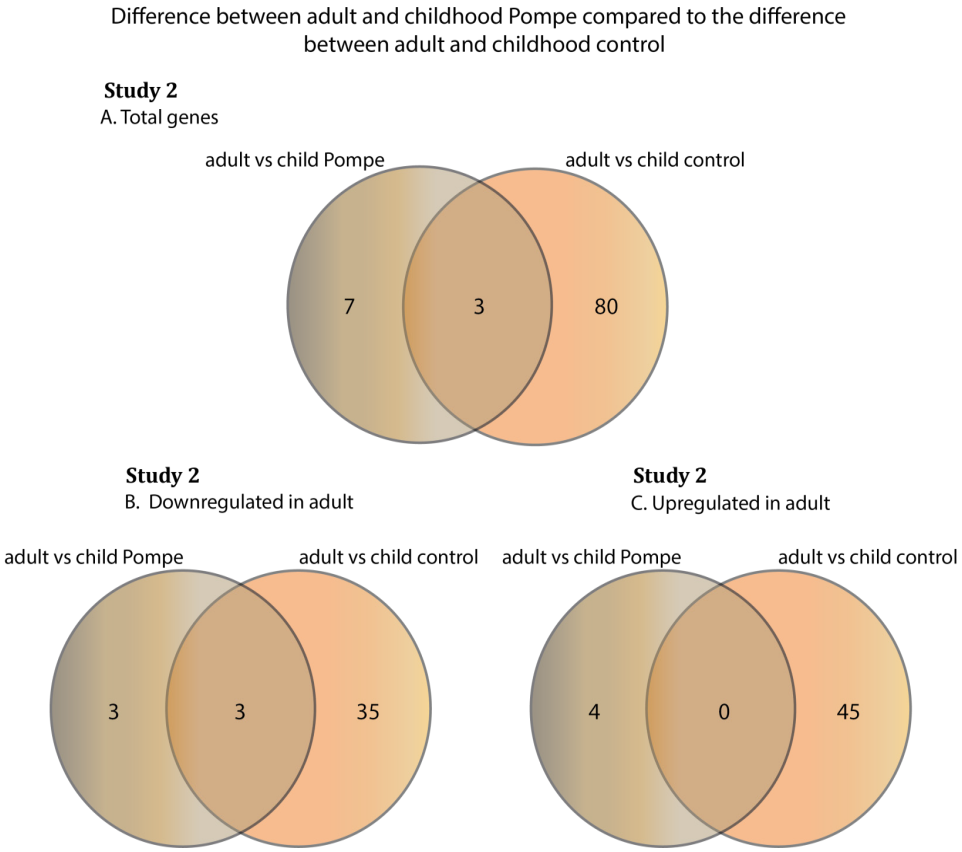


Figure 4A. Venn diagram showing the numerical difference between genes expressed in adult *versus* childhood onset Pompe disease (brown) compared to the difference between adult *versus* and childhood control (orange) of study 1. (A) A total of 7 genes are differentially expressed in childhood *versus* adult onset Pompe disease (corrected for controls), whereas 80 in the comparison of unaffected children *versus* unaffected adults. **(B)** Of the 7 genes, 3 are downregulated genes in adult onset disease and **(C)** 4 upregulated. Foreskin fibroblasts were excluded. The DEGs were filtered using a cut off value of 1.5 fold change (FC) and a false discovery rate (FDR) of <0.05.

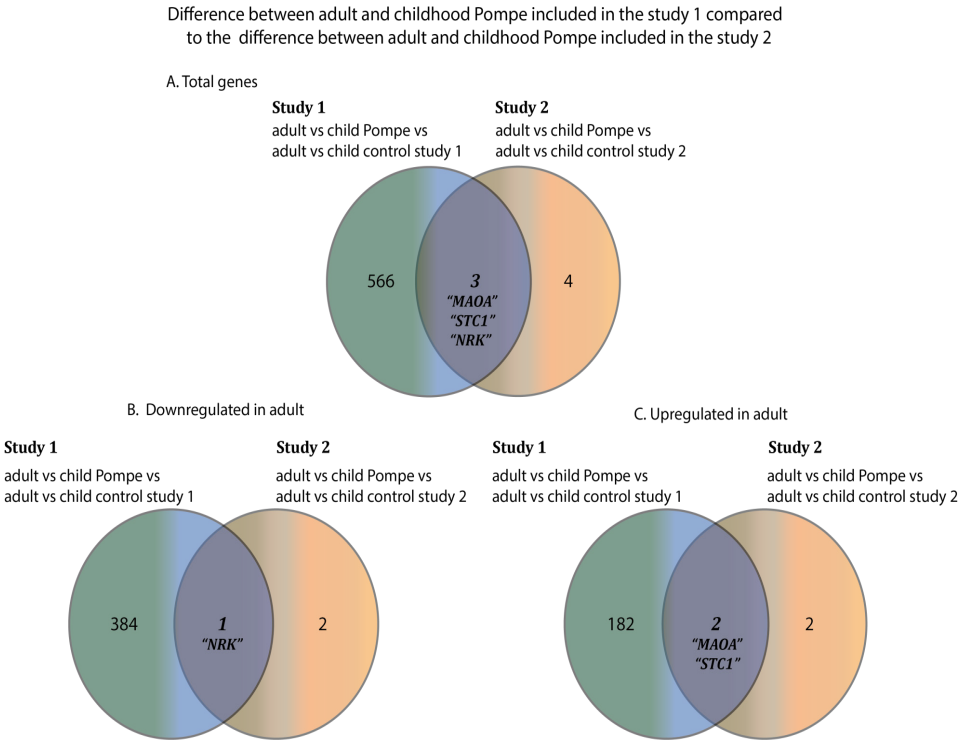


Figure 4B. Venn diagram showing the numerical difference between genes expressed in adult *versus* childhood onset Pompe disease combining study 1 and 2. The venn diagram is corrected for the differences between unaffected children and adults, and shows that **(A)** A total of 3 genes are differentially expressed in adult *versus* childhood onset Pompe disease, of which **(B)** one is downregulated in adult onset disease and **(C)** two upregulated.

Differentially expressed genes in childhood vs adult Pompe disease combining study 1 and 2

Venn diagrams were also made to identify the DEGs in adult vs childhood onset Pompe disease, corrected for childhood vs adult controls, identified in the combined studies 1 + 2. (Fig. 10A). By combining the results of all these studies only 3 DEGs came out distinguishing between childhood and adult onset Pompe disease: one downregulated in adult onset compared to childhood onset disease and two upregulated (Fig. 4B) The nik related kinase (*NRK*) gene showed over 14-fold diminished expression in 13 of the 22 samples from Pompe patients with adult onset disease relative to childhood Pompe (slightly up regulated in 4 samples and slightly down regulated in 4 others. (Fig. 5A). Genes coding for a mitochondrial enzyme *MAOA* and a secreted, homodimeric glycoprotein: *STC1*, exhibited over two-fold increased expression in adult Pompe compared to the childhood Pompe patients (Fig. 5B). Interestingly, those genes have been implicated in various biologic processes including bone and muscle development,

cellular calcium/phosphate homeostasis, cellular metabolism, and cellular biogenic amine metabolic process. (Supp. Tables S3A-B).

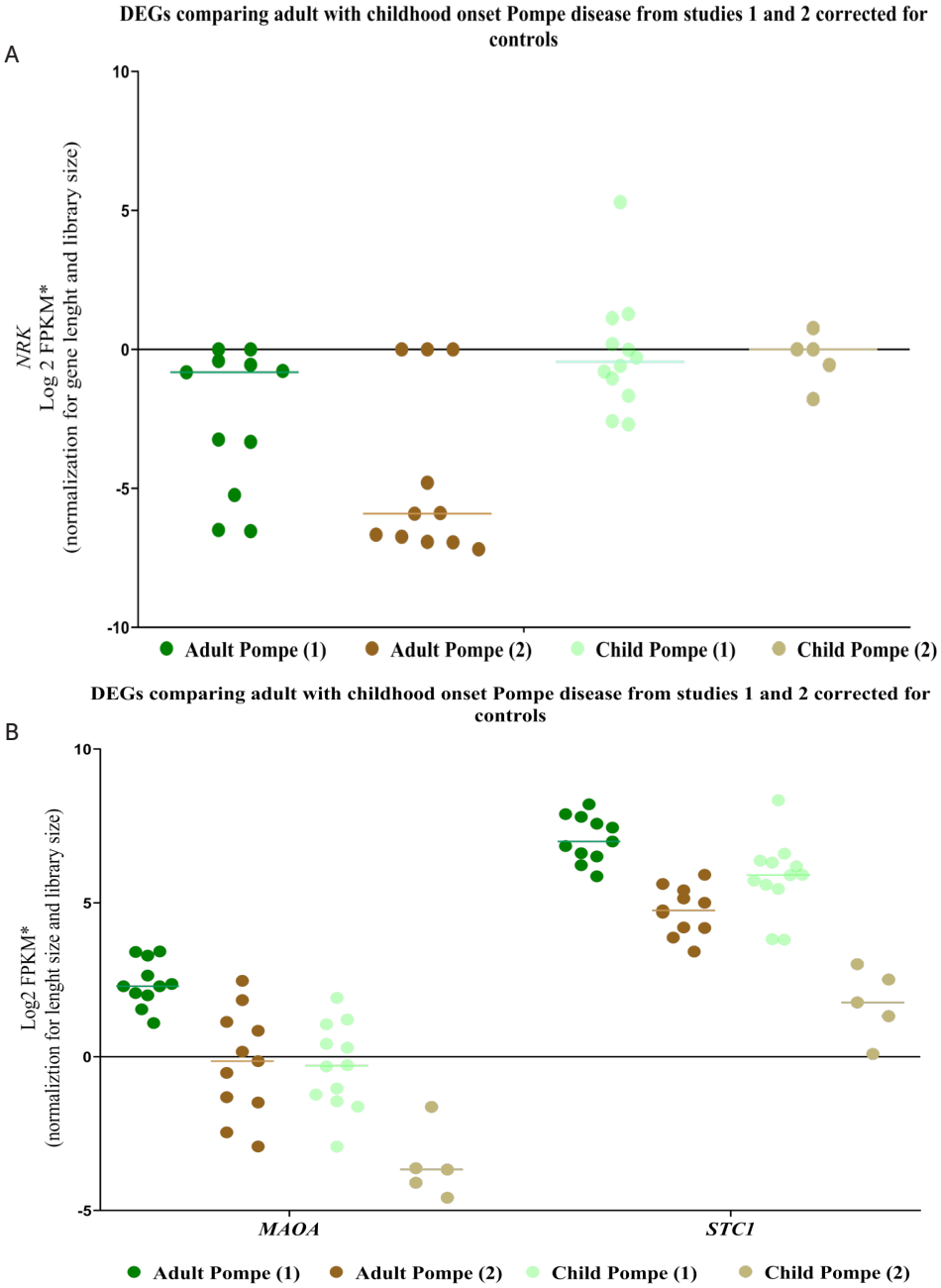


Figure 5. DEGs comparing adult with childhood onset Pompe disease from studies 1 and 2 (corrected for controls). The horizontal line represents the median value; (1/2)= study number. *Fragments per kilobase of exon per millon reads mapped. **Fig. 5A.** Downregulated DEGs in adult versus childhood onset disease. **Fig. 5B**. Same as Fig. 5A but upregulated genes.

Differentially expressed genes in adult Pompe patients vs adult Controls: Study 1
Apart from determining the difference between early and late onset disease among Pompe patients we also investigated whether particular genes would be differentially expressed in patients compared to unaffected controls.

In study 1 (A-B), pairwise comparisons were made between expression profiles of adult Pompe disease patients and adult controls. It turned out that 6607 genes were significantly differentially expressed. Of these 6607 genes, 3292 were significantly downregulated and 3315 significantly upregulated in adult Pompe patients compared to adult healthy controls (Supp. Table S4A).

All those DEGs were subjected to gene ontology and pathway analysis as described above. It appeared that over 600 GO terms were enriched among all DEGs and 390 GO terms among the downregulated DEGs (MF= 30, BP= 265 CC= 95). Enrichment of these GO terms suggests that genes involved in amino acid synthesis, RNA translation, and protein binding play a role. These results were supported by GO pathway analysis terms (Supp. Fig. S4A). Furthermore, we found by pathway analysis (IPA) that most altered pathways are associated with protein synthesis, apoptosis, and cellular stress such as EIF2 signaling, mTOR signaling, regulation of eIF4 and p70S6K signaling and mitochondrial dysfunction (Supp. Fig. S5A). The mTOR signaling pathway was the principal core and key pathway involved. It plays an important role in protein synthesis and apoptosis, but particularly in regulation of the autophagy machinery. Interestingly, dysfunctional autophagy contributes greatly to the skeletal muscle pathology in Pompe disease [28]. In fact, autophagy has a vital role in maintaining the amino acid pool. Among the upregulated DEGs, 265 GO terms were enriched (MF= 26, BP= 190, CC= 49). Immune response genes and related genes were significantly over-represented in adult vs childhood onset Pompe disease relative to adult control (Supp. Fig. S4B), which agrees with the outcome of the most significant pathway analysis by IPA (Supp. Fig. S5B).

Differentially expressed genes in adult Pompe patients vs adult Controls: study 2
In study 2, pairwise comparisons of adulthood Pompe versus adulthood control showed that 53 genes were significantly differentially expressed. Among these, 28 genes were significantly downregulated and 25 genes were significantly upregulated in adult Pompe patients compared to the adult healthy controls (Supp. Table S4A).

Differentially expressed genes in childhood Pompe patients vs childhood Controls: study 1
Using the same filtering parameters and considering FDR significance of only less

than or equal to 0.05, we identified 4426 DEGs in the cohort of childhood Pompe patients, of which 2002 were downregulated and 2424 were upregulated compared to the childhood controls (Supp. Table S4B). Those DGEs were subjected to the analysis of Gene Ontology (GO) terms and pathway analysis. More GO terms were enriched by genes having lower expression level as well as their relationships in the cell cycle process, and cellular division in childhood Pompe relative to the childhood control (Supp. Fig. S4A), which highly correlated with the pathways potentially affected (Supp. Figs. S4A and S5A). The poor expression of genes related to cell division might be caused by a cascade of events induced by the cell culture reaching confluence. The analysis including upregulated genes by Panther pathway term and Ingenuity pathway provided clear indication of over-expressed genes in childhood Pompe relative to childhood Control play a role in immune response (Supp. Figs. S4B and S5B).

To identify the DEGs specific for Pompe disease, we employed Venn diagrams comparing both adult Pompe vs adult controls as well as childhood Pompe disease vs childhood control. They illustrate that 3857 genes were expressed at a significantly different level (Supp. Fig. S5A). Among these, 1748 genes were significantly downregulated (Supp. Fig. S5B) and 2107 genes were significantly upregulated in Pompe disease (Supp. Fig. S7). In addition, this Venn diagram revealed a total of 2750 genes, which were only expressed in adult Pompe patients and 569 genes only in childhood patients (Supp. Fig. S7, top). The number of representative genes in the adult and childhood Pompe patients for the up or downregulated DEGs are shown in (Supp. Figs. S7[-B-C]).

Differentially expressed genes in childhood Pompe patients vs childhood Controls: study 2

We identified 89 DEGs in childhood Pompe, of which 49 were downregulated and 40 were upregulated compared to childhood controls (Supp. Table S4B).

We performed the same approach in study 2 as in study 1 for identifying the DEGs in Pompe patients by comparing the difference between adult Pompe and adult controls compared to the difference between childhood Pompe and childhood control. As shown in Supplementary Figure S8[A], only seven genes were differentially expressed. Among these, five genes were significantly downregulated (Supp. Fig. S8[B]) and two genes were significantly upregulated in Pompe patients (Supp. Fig. S8[C]). In addition, this Venn diagram revealed a total of 46 genes that were differentially expressed in adult Pompe patients and 82 genes in childhood patients (Supp. Figs. S8[B-C]).

With regard to the biological significance of DEGs identified in study 2, the GO and IPA

analysis did not yield any conclusive results as no processes or pathways were found to be significantly enriched, which can be explained by the much lower number of DEGs identified in study 2 compared to study 1.

Differentially expressed genes in Pompe disease compared to unaffected controls combining the results of study 1 + 2.
Apart from analyzing the differences in gene expression profiles between childhood and adult Pompe patients, we also compared the profiles of fibroblasts from affected versus unaffected individuals. The results are depicted in the Venn diagrams of Supplementary Figures S9 (adults), S10 (children) and S11 (adults and children combined). In the comparison adult Pompe/ adult control 32 genes came out as differentially expressed (Supp. Fig. S9[A]), but only 17 genes were identified that had the same behavior (up or down) in both studies. Nine of these 17 genes were downregulated (Supp. Fig. S9[B]) and 8 upregulated in adult Pompe patients compared to adult controls (Supp. Fig. S9[C]). In the comparison childhood Pompe / childhood control 31 genes were identified (Supp. Fig. S10[A]), whereby 9 showing the same behavior came out in both studies. Six of these were downregulated (Supp. Fig. S10[B]) and 3 upregulated in childhood Pompe patients compared to childhood controls (Supp. Fig. S10[C]).

Finally, the DEGs identified in the childhood/childhood and adult/adult comparisons were analyzed in combination (Supp. Fig. S11) and this resulted in the identification of only one single gene that was downregulated in all Pompe disease patients compared to all controls. It turned out to be the *GAA* gene, and its lower expression in fibroblasts of all patients is explained by the disease-associated c.-32-13T>G variant that all patients in our study share and leads to 80-90% diminished GAA mRNA expression. The finding is comforting as it contributes to the validity of our approach and supports the additional findings.

Validation of the 3 common DEGs between adulthood and childhood patients in study 1 and 2.
Again, to validate the accuracy of the RNA-Seq analysis, qRT-PCR was performed in study 3, which included fibroblasts of a new set of six childhood and six adult patients with Pompe disease. This new test confirmed the statistically significant higher expression level of *MAOA* in affected adults compared to affected children (Fig. 6A), also when applied to all patients and controls included in the studies 1-3 ($P<0.01$) (Fig. 6B). Figure 6C is suggestive for an upward trend of the *MAOA* gene expression level with age of onset. In contrast, in healthy controls age-related changes in gene expression of *MAOA* were not observed (Suppl. Figure S12A). Similar results were obtained for

STC1 gene expression in study 3 (Fig. 7A) as well as in studies 1-3 (Figs. 7B-C) albeit that 4 samples in particular had a strong effect on the upward trend. Supplementary Figure 12B does not show age-related gene expression changes of *STC1* in controls. Analysis of *NRK* expression by RT-qPCR revealed that this gene was not expressed in most individuals but that it was highly expressed in selected individuals, without preference for one group. Therefore, we did not further pursue *NRK* expression.

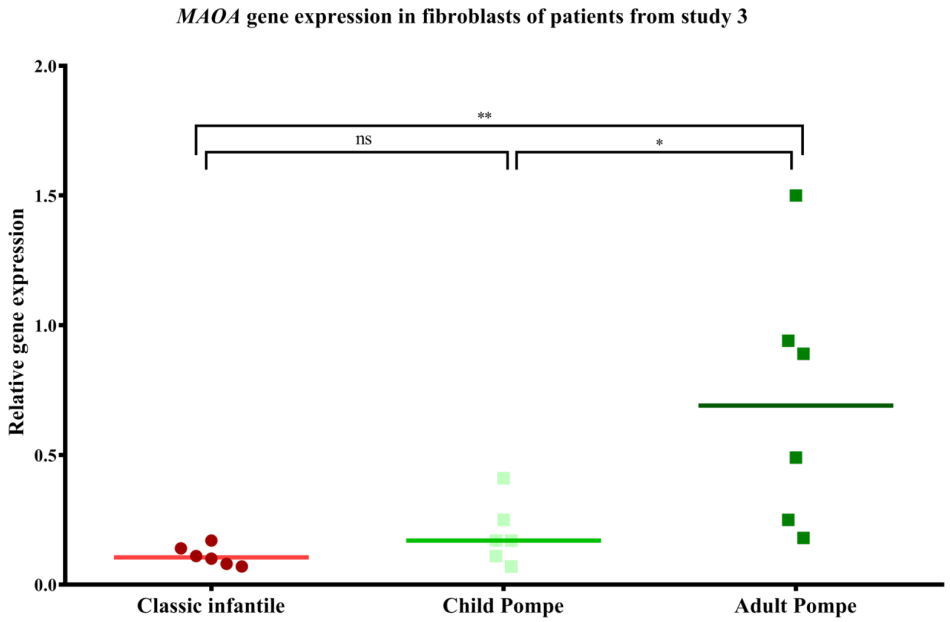


Figure 6A. *MAOA* gene expression in fibroblasts of patients from study 3. The gene expression level was measured by RT-qPCR in a third series of samples to validate the results obtained in study 1-2. The horizontal line presents the median value. ns, not significant (P>0.05), * significant (P<0.05); ** indicate P<0.01. The one-way ANOVA test was used for analysis.

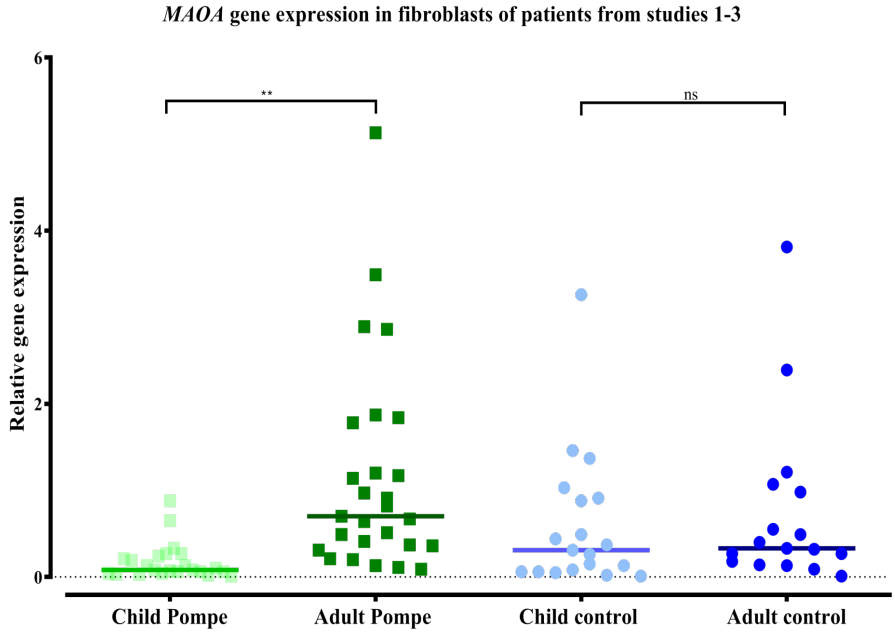


Figure 6B. *MAOA* gene expression in fibroblasts of patients from studies 1-3. The gene expression level was measured by RT-qPCR combining studies 1-3. The horizontal line presents the median value. ns, not significant (P>0.05). ** significant (P<0.01). The one-way ANOVA test was used for the analysis.

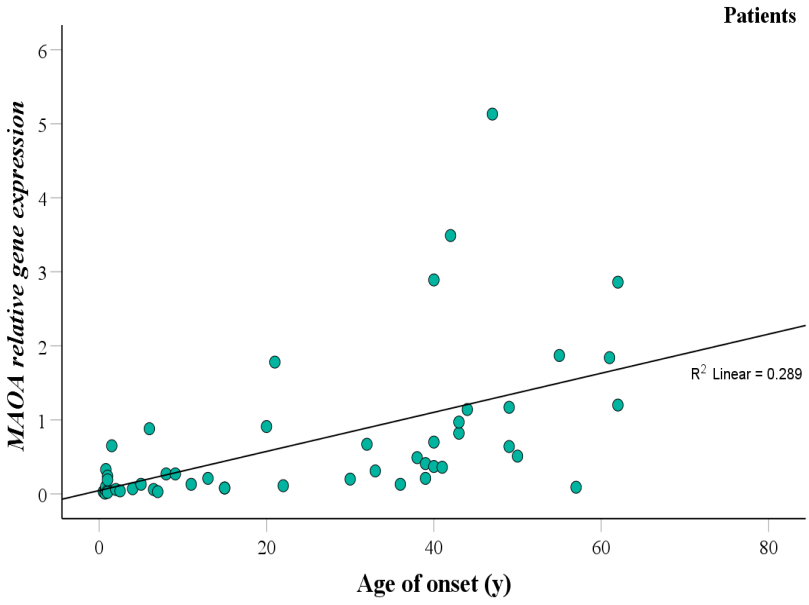


Figure 6C. *MAOA* gene expression by age of onset in studies 1-3. The expression level was measured by RT-qPCR. Linear regression analysis using SPSS was used.

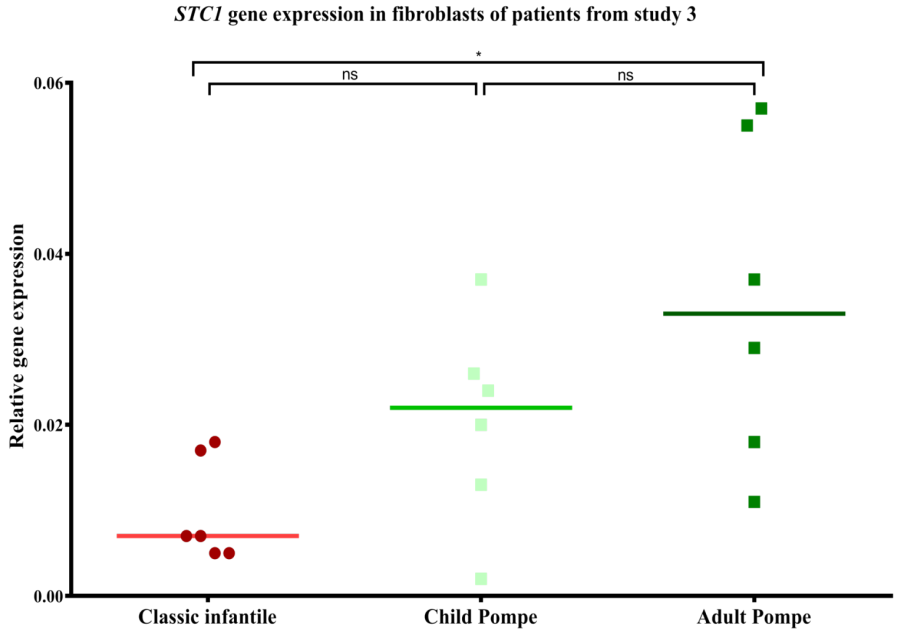


Figure 7A. *STC1* gene expression in fibroblasts of patients from study 3. The gene expression level was measured by RT-qPCR to validate the results obtained in study 1-2. The horizontal line presents the median value. The one-way ANOVA test was used for the analysis. ns, not significant (P>0.05), significant *P<0.05.

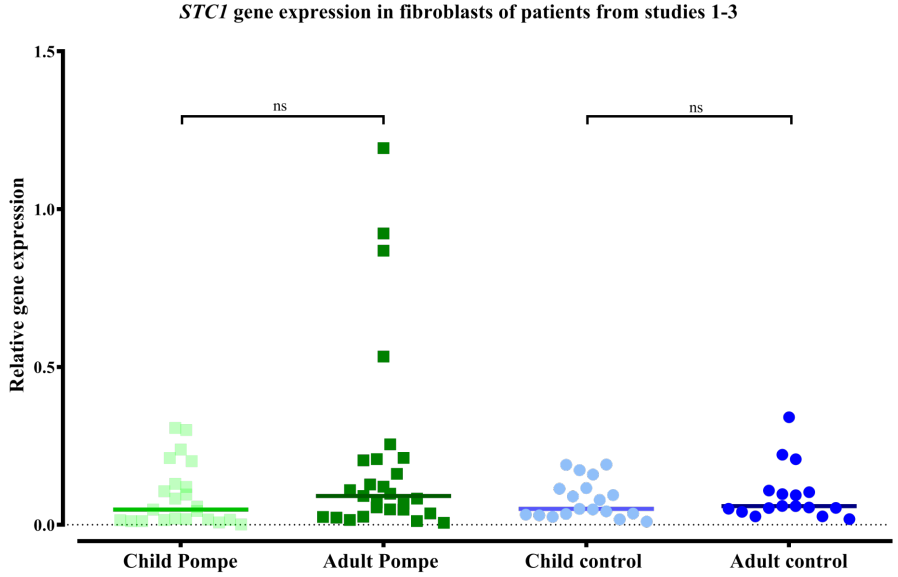


Figure 7B. *STC1* gene expression in fibroblasts of patients from studies 1-3. The gene expression level was studied by RT-qPCR. The horizontal line presents the median value. The one-way ANOVA test was used for the analysis. ns, not significant (P>0.05).

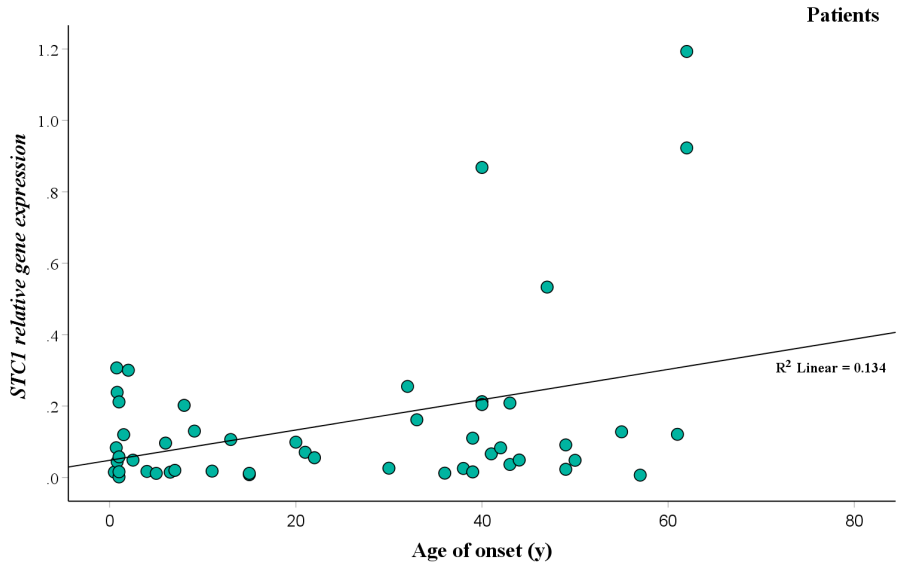


Figure 7C. *STC1* gene expression by age of onset in fibroblast of patients from studies 1-3. The expression level was measured by RT-qPCR. Linear regression analysis was performed using SPSS.

DISCUSSION

The transcriptomes of fibroblast cultures from patients with childhood and adult Pompe disease were investigated to gain insights in the DEGs pattern of various Pompe disease phenotypes. To the best of our knowledge this is a unique study for several reasons. Firstly, there is hardly any literature on markers and/or factors that can be used for monitoring disease progression in Pompe disease. Secondly, transcriptome analysis using Pompe patient fibroblasts has never been thoroughly described, although one paper has analyzed gene expression by using microarray, but this was performed on muscle biopsies of classic infantile Pompe patients that don't have the c.-32-13T>G variant that leads to residual GAA activity like the patients in our study [29]. Thirdly, cultured skin fibroblasts were specifically chosen for this purpose as they are by far the best sample source for accurate measurement of the residual GAA enzyme activity level. In fact, our results of the lysosomal glycogen content confirmed a previous finding that cultured fibroblasts from patients other than those with classic infantile Pompe disease do not store measurable amounts of glycogen [26]. Finally, our study utilizes the more powerful deep sequencing strategy and different approaches to identify DEGs instead of the more commonly used microarray datasets. Lastly, we compared the transcriptome of childhood and adult Pompe patients carrying the

functionally similar c.-32-13T>G / 'null', genotype in 35/39 cases to understand the broad clinical variation observed in those patients.

The significantly downregulated genes related to the amino acid biosynthesis pathway in adult vs childhood Pompe disease compared to childhood versus adult controls discovered in study 1 might suggest a compensatory molecular mechanism in response to cellular stress and cellular metabolic adaptation. The overexpression of genes associated with immune response pathways might point to better cellular regeneration in fibroblasts from patients with adult compared to childhood Pompe disease.

Importantly, comparison between the analyses of study 1 and study 2 revealed only one common gene that showed altered expression between adult and childhood Pompe disease, which was the *GAA* gene. This can be taken as a clear indication of unbiased analysis. Despite the fact that study 2 identified a low number of DEGs, the trend of significantly up- or down-regulated genes in this study was the same but not significant (data not shown), which is likely due to a large variation between individuals. The large variation observed in the second study reflects reality but serves as a highly valuable 'stress test'. For this reason, we argue that the core dataset determined under this approach survived this stress test and may represent a more robust read-out for symptom onset in Pompe disease compared to either dataset from study 1 and 2 separately. The 3 genes *MAOA, Stc-1*, and *NRK-1* have been implicated in various biologic processes including bone and muscle development, cellular calcium/phosphate homeostasis, cellular metabolism, and cellular biogenic amine metabolic process. In addition, from the total of 24 DGEs identified in fibroblasts from Pompe patients compared to controls (Supp. Fig. S11), 16 were differentially expressed only in adult Pompe. The downregulated DEGs (in adulthood Pompe) were mainly involved in cell adhesion and cell-cell adhesion. Of particular interest, *WNT2* plays an important role in the Wnt and mTOR pathways, which are involved in several cellular processes such as proliferation, cell migration, apoptosis, protein biosynthesis, and autophagy [30, 31]. In fact, *WNT2* is an upstream signaling factor of the mTOR pathway that regulates glycogen synthase kinase (GSK-3) activity. GSK-3 has been shown to function in cellular division, proliferation, and survival, but furthermore, it is an important enzyme in glycogen metabolism. In addition the mTOR complex is involved in anabolic activities in the lysosome including autophagy, which has been shown to be dysfunctional in Pompe patients and consequently contributes to the pathological cascade in skeletal muscle [32, 33]. The down-regulation of *WNT2* expression might be explained as part of a compensatory molecular mechanism in response to cellular stress and cellular metabolic adaptation in adult Pompe patients.

On the other hand, patients with childhood Pompe disease showed up-regulation of differentially expressed genes such as *FNIP2* and *SLC9A7* that may be involved in energy and/or nutrient sensing through the AMPK and mTOR signaling pathways, transport of glucose plus other sugars, bile salts and organic acids, metal ions and amine compounds [www.genecards.org/]. Notably, diseases associated with pathogenic *SLC9A7* gene variants include scapuloperoneal myopathy in which muscle fibers are also damaged or dysfunctional [34], as in Pompe disease [35-37]. In fact, scapular winging is a condition that has been reported in Pompe disease, but with low frequency. Also, three childhood Pompe patients (C2, C3 and C6) in this study (1A) presented scapular winging. However, as shown in Supplementary Figure S10, patient (C2) who's symptoms were so severe they interfered with daily life (moderate status) at the time of biopsy had a higher *SLC9A7* gene expression than the two other children (C3 and C6) who's symptoms barely interfered with daily life (mild status) at the time of biopsy. In fact, the gene expression of *SLC9A7* in C3 and C6 is almost the same. (Supp. Fig. S13). This argues against a causative role of *SLC9A7* in scapular winging.

Part of the explanation for the lack of consistency between study 1 and 2 may be technical, variations in control samples, small sample size, and possible batch effects because the patients and controls of the first study were handled in separate studies (1A and 1B). However, we included three identical RNA samples in study 1A in study 1B, and these gave similar results, making a batch effect unlikely. A more likely explanation can be deduced from Figs. 2B and C. In study 1, the patient cohorts are very well separated, while in study 2, the separation is far less defined with respect to age of onset and disease progression. Also in study 3, separation between the two cohorts is not as good as in study 1 (Table 3). This indicates that differences in gene expression profiles can be found at the group level provided that patient cohorts are sharply separated, but that it is difficult to link expression profiles to individual patients.

Table 3. Minimum and maximum ages of onset among all patients in studies 1-3.

Study number	Study 1		Study 2		Study 3	
Phenotype	Childhood	Adult	Childhood	Adult	Childhood	Adult
Patients (n)	12	11	5	11	6	6
Median age at first symptoms (y)	1.25	47	7	33	8.75	39.5
Minimum age at first symptoms (y)	0.5	40	0.75	20	1	36
Maximum at first symptoms (y)	8	62	13	62	15	57
Range (y)	7.5	22	12.25	42	14	21
Mean age at first symptoms (y)	2.7	48.4	6.4	34.9	8.3	43.2
Std. Deviation	2.5	8.0	5.1	12.3	6.4	8.1

The median, mean and std. deviation of age of symptoms onset is indicated per phenotype and study number.

Pompe disease manifests with a broad clinical heterogeneity. Currently, the diagnosis and progression of Pompe disease relies predominately on GAA activity indices. Studies and development of factors that may modulate the underlying pathogenic etiologies of glycogen accumulation should offer the possibility to target those disease-markers that would monitor disease progression and could help in decision making about the timing of therapeutic intervention. Our study may also help to reveal new insight in potential disease-markers for Pompe Disease.

Our findings provide an argument to further investigate the impact of *MAOA* and *STC1* on the clinical course of Pompe disease, as they comprise the core set of DEGs that were consistently upregulated in patients with adulthood onset compared to patients with childhood onset. In light of these results, this core set of DEGs might be considered as potential disease-markers in patients with the IVS1/null genotype.

Funding information
Administrative Department of Science, Technology and Innovation (Colciencias, Colombia), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Brasil).

REFERENCES

1. Pompe, JC (1932). Over idiopathische hypertrofie van het hart. *Ned Tijdsch Geneesk* **76**: 304-311.
2. Schoser, B, Laforet, P, Kruijshaar, ME, Toscano, A, van Doorn, PA, van der Ploeg, AT*, et al.* (2015). 208th ENMC International Workshop: Formation of a European Network to develop a European data sharing model and treatment guidelines for Pompe disease Naarden, The Netherlands, 26-28 September 2014. *Neuromuscul Disord*.
3. Engel, AG, Seybold, ME, Lambert, EH, and Gomez, MR (1970). Acid maltase deficiency: comparison of infantile, childhood, and adult types. *Neurology* **20**: 382.
4. Engel, AG, Gomez, MR, Seybold, ME, and Lambert, EH (1973). The spectrum and diagnosis of acid maltase deficiency. *Neurology* **23**: 95-106.
5. Reuser, AJJ, Hirschhorn, R, and Kroos, MA (2019). Pompe Disease: Glycogen Storage Disease Type II, Acid α-Glucosidase (Acid Maltase) Deficiency. In: Valle, D, S Antonarakis, A Ballabio, A Beaudet and GA Mitchell (eds). *The Online Metabolic and Molecular Bases of Inherited Disease*. McGraw-Hill Education: New York, NY.
6. van der Ploeg, AT, and Reuser, AJ (2008). Pompe's disease. *Lancet* **372**: 1342-1353.
7. Kishnani, PS, Corzo, D, Nicolino, M, Byrne, B, Mandel, H, Hwu, WL*, et al.* (2007). Recombinant human acid [alpha]-glucosidase: major clinical benefits in infantile-onset Pompe disease. *Neurology* **68**: 99-109.
8. van der Ploeg, AT, Clemens, PR, Corzo, D, Escolar, DM, Florence, J, Groeneveld, GJ*, et al.* (2010). A randomized study of alglucosidase alfa in late-onset Pompe's disease. *N Engl J Med* **362**: 1396-1406.
9. Van den Hout, H, Reuser, AJ, Vulto, AG, Loonen, MC, Cromme-Dijkhuis, A, and Van der Ploeg, AT (2000). Recombinant human alpha-glucosidase from rabbit milk in Pompe patients. *Lancet* **356**: 397-398.
10. Winkel, LP, Kamphoven, JH, van den Hout, HJ, Severijnen, LA, van Doorn, PA, Reuser, AJ*, et al.* (2003). Morphological changes in muscle tissue of patients with infantile Pompe's disease receiving enzyme replacement therapy. *Muscle Nerve* **27**: 743-751.
11. de Vries, JM, van der Beek, NA, Hop, WC, Karstens, FP, Wokke, JH, de Visser, M*, et al.* (2012). Effect of enzyme therapy and prognostic factors in 69 adults with Pompe disease: an open-label single-center study. *Orphanet J Rare Dis* **7**: 73.
12. Toscano, A, and Schoser, B (2013). Enzyme replacement therapy in late-onset Pompe disease: a systematic literature review. *J Neurol* **260**: 951-959.
13. Kuperus, E, Kruijshaar, ME, Wens, SCA, de Vries, JM, Favejee, MM, van der Meijden, JC*, et al.* (2017). Long-term benefit of enzyme replacement therapy in Pompe disease: A 5-year prospective study. *Neurology* **89**: 2365-2373.
14. van der Ploeg, AT, Kruijshaar, ME, Toscano, A, Laforet, P, Angelini, C, Lachmann, RH*, et al.* (2017). European consensus for starting and stopping enzyme replacement therapy in adult patients with Pompe disease: a 10-year experience. *Eur J Neurol* **24**: 768-e731.
15. de Vries, JM, Kuperus, E, Hoogeveen-Westerveld, M, Kroos, MA, Wens, SC, Stok, M*, et al.* (2017). Pompe disease in adulthood: effects of antibody formation on enzyme replacement therapy. *Genet Med* **19**:

90-97.

16. van Gelder, CM, Hoogeveen-Westerveld, M, Kroos, MA, Plug, I, van der Ploeg, AT, and Reuser, AJ (2015). Enzyme therapy and immune response in relation to CRIM status: the Dutch experience in classic infantile Pompe disease. *J Inherit Metab Dis* **38**: 305-314.
17. van der Meijden, JC, Kruijshaar, ME, Harlaar, L, Rizopoulos, D, van der Beek, N, and van der Ploeg, AT (2018). Long-term follow-up of 17 patients with childhood Pompe disease treated with enzyme replacement therapy. *J Inherit Metab Dis*. **41**: 1205-1214.
18. Manwaring, V, Prunty, H, Bainbridge, K, Burke, D, Finnegan, N, Franses, R, *et al.* (2012). Urine analysis of glucose tetrasaccharide by HPLC; a useful marker for the investigation of patients with Pompe and other glycogen storage diseases. *J Inherit Metab Dis* **35**: 311-316.
19. Chien, YH, Han, DS, Hwu, WL, Thurberg, BL, and Yang, WS (2013). Myostatin and insulin-like growth factor I: potential therapeutic biomarkers for pompe disease. *PLoS One* **8**: e71900.
20. Carrasco-Rozas, A, Fernandez-Simon, E, Lleixa, MC, Belmonte, I, Pedrosa-Hernandez, I, Montiel-Morillo, E, *et al.* (2019). Identification of serum microRNAs as potential biomarkers in Pompe disease. *Ann Clin Transl Neurol* **6**: 1214-1224.
21. Tarallo, A, Carissimo, A, Gatto, F, Nusco, E, Toscano, A, Musumeci, O, *et al.* (2019). microRNAs as biomarkers in Pompe disease. *Genet Med* **21**: 591-600.
22. Wens, SC, van Gelder, CM, Kruijshaar, ME, de Vries, JM, van der Beek, NA, Reuser, AJ, *et al.* (2013). Phenotypical variation within 22 families with Pompe disease. *Orphanet J Rare Dis* **8**: 182.
23. Kroos, M, Hoogeveen-Westerveld, M, van der Ploeg, A, and Reuser, AJ (2012). The genotype-phenotype correlation in Pompe disease. *Am J Med Genet C Semin Med Genet* **160C**: 59-68.
24. Oba-Shinjo, SM, da Silva, R, Andrade, FG, Palmer, RE, Pomponio, RJ, Ciociola, KM, *et al.* (2009). Pompe disease in a Brazilian series: clinical and molecular analyses with identification of nine new mutations. *J Neurol* **256**: 1881-1890.
25. Huie, ML, Chen, AS, Tsujino, S, Shanske, S, DiMauro, S, Engel, AG, *et al.* (1994). Aberrant splicing in adult onset glycogen storage disease type II (GSDII): molecular identification of an IVS1 (-13T->G) mutation in a majority of patients and a novel IVS10 (+1GT->CT) mutation. *Human Molecular Genetics* **3**: 2231-2236.
26. Umapathysivam, K, Hopwood, JJ, and Meikle, PJ (2005). Correlation of acid alpha-glucosidase and glycogen content in skin fibroblasts with age of onset in Pompe disease. *Clin Chim Acta* **361**: 191-198.
27. Smith, PK, Krohn, RI, Hermanson, GT, Mallia, AK, Gartner, FH, Provenzano, MD, *et al.* (1985). Measurement of protein using bicinchoninic acid. *Anal Biochem* **150**: 76-85.
28. Feeney, EJ, Spampanato, C, Puertollano, R, Ballabio, A, Parenti, G, and Raben, N (2013). What else is in store for autophagy? Exocytosis of autolysosomes as a mechanism of TFEB-mediated cellular clearance in Pompe disease. *Autophagy* **9**: 1117-1118.
29. Palermo, AT, Palmer, RE, So, KS, Oba-Shinjo, SM, Zhang, M, Richards, B, *et al.* (2012). Transcriptional response to GAA deficiency (Pompe disease) in infantile-onset patients. *Mol Genet Metab* **106**: 287-300.
30. Lim, JA, Li, L, Shirihai, OS, Trudeau, KM, Puertollano, R, and Raben, N (2017). Modulation of mTOR signaling as a strategy for the treatment of Pompe disease. *EMBO Mol Med* **9**: 353-370.

31. Lyu, JW, Xu, XB, Ji, KQ, Zhang, N, Sun, Y, Zhao, DD, *et al.* (2019). Activated mTOR signaling pathway in myofibers with inherited metabolic defect might be an evidence for mTOR inhibition therapies. *Chin Med J (Engl)* **132**: 805-810.
32. Lim, JA, Meena, NK, and Raben, N (2019). Pros and cons of different ways to address dysfunctional autophagy in Pompe disease. *Ann Transl Med* **7**: 279.
33. Nascimbeni, AC, Fanin, M, Tasca, E, Angelini, C, and Sandri, M (2015). Impaired autophagy affects acid alpha-glucosidase processing and enzyme replacement therapy efficacy in late-onset glycogen storage disease type II. *Neuropathol Appl Neurobiol* **41**: 672-675.
34. Tzschach, A, Ullmann, R, Ahmed, A, Martin, T, Weber, G, Decker-Schwering, O, *et al.* (2011). Christianson syndrome in a patient with an interstitial Xq26.3 deletion. *Am J Med Genet A* **155A**: 2771-2774.
35. Herbert, M, Case, LE, Rairikar, M, Cope, H, Bailey, L, Austin, SL, *et al.* (2019). Early-onset of symptoms and clinical course of Pompe disease associated with the c.-32-13T>G variant. *Mol Genet Metab* **126**: 106-116.
36. van der Beek, NA, de Vries, JM, Hagemans, ML, Hop, WC, Kroos, MA, Wokke, JH, *et al.* (2012). Clinical features and predictors for disease natural progression in adults with Pompe disease: a nationwide prospective observational study. *Orphanet J Rare Dis* **7**: 88.
37. Oya, Y, Morita, H, Ogawa, M, Nonaka, I, Tsujino, S, and Kawai, M (2001). [Adult form of acid maltase deficiency presenting with pattern of muscle weakness resembling facioscapulohumeral dystrophy]. *Rinsho Shinkeigaku* **41**: 390-396.

SUPPLEMENTARY DATA / SUPPORTING INFORMATION

Supplementary text

Clinical features and disease severity

Table 2 provides demographic characteristics of the patients in cohort 1 and cohort 2. The majority (76.5%) of affected children in cohort 1 is male. Only one of the 17 patients (5.9%) in cohort 1 was partially wheelchair and ventilation dependent at the time of study. All others (94%) were still ambulant. The most frequent first symptom was delayed motor development (35.3%) followed by elevated CK (29.4%). Among the adult patients in cohort 2, 36.4% are males, Five of these patients (22.7%) were partially wheelchair dependent, 2 (9.1%) wheelchair dependent, 1 (4.5%) using a walking aid, while the others (63.6%) were ambulant. Six (27.3%) patients were ventilation dependent. The most frequent first symptom in cohort 2 was fatigue, presenting in 6 (27.3%) patients, followed by difficulty climbing stairs in 4 (18.2%).

One childhood patient (C11) became wheelchair dependent at 4 years of age, but is able to walk a distance of 200 meter. Another child (C1) was wheelchair dependent at the age of 11 years, but regained walking ability during ERT. Two adult patients were fully wheelchair dependent at 51 (A17) and 60 years of age (A21), respectively. Five adult patients had become partially wheelchair dependent from 36 and 63 years onwards. One childhood patient (C2) became ventilator-dependent, using nightly bi-level positive airway pressure (BiPaP) at 12 years, and six adult patients were treated with either invasive or non-invasive ventilation support. The patients using non-invasive ventilation were: A11 at age 57 years, A19 at age 54 years and A20 at age 53 years. The patients that required invasive ventilator support are listed as follow: A17 at age 39 years, A18 at age 33 and A21 at age 67 years (Supp. Tables S1A-B).

As shown in Supplementary Tables S1A-B the currently most severely affected Pompe patients are two adults (patients A17 and A21). Patient A17 had first symptoms at 30 years, started invasive ventilation support at 39 years and became fully wheelchair dependent at the age of 51 years. Patient A21 with an age of onset at 40 years, was fully wheelchair dependent at the age of 60 years and was invasively ventilated at the age of 67 years. Of course, disease severity in Pompe disease is a matter of age of onset, rate of disease progression, and moment of evaluation. In this respect, the childhood patient (C2) that had first symptoms under the age of 1 year and started to use ventilation support at 12 years of age may in time become the most severely affected. The clinical course is hard to predict in most cases. The clinical features of all child and adult patients with Pompe disease and kind of first symptoms are

described in Supplementary Tables S1A-D. The onset of symptoms and severity of disease progression varied substantially even between these patients with the same set of disease-associated variants on both *GAA* alleles. Their *GAA* genotypes and complementary information of the second disease-associated variant are listed in Supplementary Table S2.

Supplementary Table S1A. Clinical features of patients with Pompe disease in study 1.

Clinical phenotype	Childhood Pompe patients												Adult Pompe patients										
Patient id	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8	Ch9	Ch10	Ch11	Ch12	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Age at first symptoms (y)	0.5	1	0.7	2.5	0.8	4	5	6	0.8	1.5	1	8	49	43	40	43	40	61	42	50	55	62	47
Age at time of biopsy (y)	2.8	16.3	2.6	10.9	9.1	16	11	8.6	9.8	13.4	8.3	1.6	58.1	48.4	49.5	50.8	48.7	68	58.6	54	68.9	72.9	60
Clinical status at biopsy	mild	mod	mild	mild	mild	mild	mild	very mild	very mild	mild/ mod?	?	unk-nown	mild/ mod	mild	mod	mod	mild	mild	mod	mild	mild/ mod	mild/ mod	sev
Diagnosis age (y)	2.5	1.3	2.3	3	1.6	4	10.8	7.8	2.5	2	2.2	7.3	50	26	43	41	43	67	57	52	68	72	56
Start ERT (y)	11.9	15.2	2.9	12.9	8.9	16	11	8.5	9.9	no	no	?	58	no	50	53	49	na	58	54	69	73	61
Gender	M	M	M	M	F	M	M	F	M	M	M	F	F	F	F	F	M	M	F	F	M	F	M
Wheelchair age (y)	11										<4												
Mobility (wheelchair)											+^												
Mobility (ambulant)	+*	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+
Ventilation age (y)		12																				57	
Ventilation type		BP																					Non-I
Elevated CK			+			+		+	+			+											
FVC (sitting) % predicted	97	66	86	66	102	67	102	104	82	na	59	71	79	97	53	70	84	94	89	96	98	68	54
FVC (supine) % predicted	87	45	na	54	101	53	97	92	80	na	48	na	52	77	33	38	73	75	75	86	77	49	22
MRC score (%)	78	89	80	83	92	98	97	99	99	na	80	na	84	90	85	84	89	95	84	94	89	93	86

Abbreviations: Ch1-Ch12= childhood onset Pompe disease and A1-A11= adult onset Pompe disease included in study 1. BP = Nightly BiPaP; Non-I= non-invasive; na= not available/applicable; M= male, F= female; mild= symptoms barely interfere with daily life; mod= moderate: symptoms so severe that they interfere with daily life; sev= severe: in the need of a wheelchair and/or ventilator; +*= Ch1 starts walking during ERT; +^= Ch11 is partly wheelchair dependent, can walk 200m.

Supplementary Table S1B. Clinical features of patients with Pompe disease in study 2.

Clinical phenotype	Childhood Pompe patients					Adult Pompe patients										
Patient id	Ch13	Ch14	Ch15	Ch16	Ch17	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22
Age at first symptoms (y)	0.75	9.1	2	7	13	32	39	20	22	41	30	33	21	44	40	62
Age at time of biopsy (y)	1.06	10.4	0.35	10.8	14.3	39	48	52	41	53.9	57.6	40	61.5	62.7	70	72
Clinical status at biopsy	mild	mild	mild	sev	mild	mild	mild	mild/ mod	sev	mod	sev	sev	mod/ sev	sev	sev	mild
Diagnosis age (y)	0	9.4	0.33	9.5	14.1	31	48	52	37	53	41	34	42	44	45	63
Start ERT (y)	1	10.6	13.1	33	14.9	39	48.9	52.3	52	54.7	57.7	59	61.6	62.9	70.4	74
Gender	F	M	M	M	M	F	M	F	F	F	M	M	F	M	F	F
Wheelchair age (y)				22					39		51	36	43	62	60	63
Mobility (wheelchair)				+					+*		+	+*	+*	+*	+	+*
Mobility (ambulant)	+	+	+		+	+	+	+								
Mobility (walking aid)									+	+		-*	-*	-*		-*
Ventilation age (y)											39	33	54	53	67	
Ventilation type											INV	INV	Non-I	Non-I	INV	
FVC (sitting) % predicted	115	84	97	na	90	87	61	103	48	75	na	na	18	33	34	81
FVC (supine) % predicted	109	63	na	na	85	86	50	89	23	41	na	na	na	na	na	76
MRC score (%)	98	97	33	na	97	82	94	85	64	79	57	60	68	72	na	81

Abbreviations: Ch13-Ch17 are childhood onset Pompe patients, and A12-A22 are adult onset Pompe patients included in study 2. INV= invasive; Non-I= non-invasive; +*= partially wheelchair dependent; -*= no walking aid; na= not available; M= male, F= female; mild= symptoms barely interfere with daily life; mod= moderate: symptoms so severe that they interfere with daily life; sev= severe: in need of a wheelchair and/or ventilator; +=yes.

Supplementary Table S1C. Kind of first symptoms of patients with Pompe disease in study 1.

Clinical phenotype	Childhood onset Pompe patients												Adult onset Pompe patients										
patient ID	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8	Ch9	Ch10	Ch11	Ch12	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11
Kind of first symptoms																							
hypotonia		+	+		+				+														
delayed motor development		+			+				+	+	+												
delayed growth								+															
growth stagnation							+																
feeding problems	+									+													
difficulty swallowing																						+	
difficulty chewing																						+	
slipping through									+														
fatigue													+	+						+			+
pulmonary weakness																							+
lower back pain																		+					
muscle ache																		+					
mild muscle weakness													+										
falling				+																			
walking problems				+																			
abnormal gait										+													
difficulty walking																+							
difficulty running																					+		
difficulty with practicing sports															+								
difficulty climbing stairs																	+		+				
difficulty standing from squatting position														+									
difficulty changing posture																			+				
difficulty standing straight																				+			
difficulty stumbling																+							
stiffness															+								
dysarthria																						+	
diarrhea							+	+															

Supplementary Table S1D. Kind of first symptoms of patients with Pompe disease in study 2

Clinical phenotype	Childhood onset Pompe patients							Adult onset Pompe patients								
patient ID	Ch13	Ch14	Ch15	Ch16	Ch17	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22
Kind of first symptoms																
delayed motor development	+															
fatigue		+			+			+								+
Respiratory failure												+				
lower back pain							+									+
muscle ache		+				+				+						
mild muscle weakness of legs															+	
falling		+														
waddling gait														+		+
abnormal gait					+											
difficulty walking								+		+						
difficulty with practicing sports				+												
difficulty climbing stairs						+							+			
difficulty standing from supine position											+					
difficulty changing posture							+									
stumbling									+							
stiffness													+			
reduced force legs			+													
difficulty standing up									+							

Supplementary Table S2. *GAA* genotypes of patients with Pompe disease reported in studies 1 plus 2.

	pt ID	Location	DNA	Protein	Type of variant (DNA)	predicted severity	ACMG classification	phenotype with null allele	CR IM
CHILDHOOD ONSET PATIENT	**Ch1**	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch2	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch3	E15	c.2135T>C	p.(Leu712Pro)	substitution	less severe	likely pathogenic	unknown (disease-associated)	+
	Ch4	I16	c.2331+2T>A	p.?	substitution	very severe	pathogenic	unknown (disease-associated)	-*
	Ch5	E5	c.923A>C	p.(His308Pro)	substitution	potentially less severe	likely pathogenic	classic infantile	+
	Ch6	E10	c.1441T>C	p.(Trp481Arg)	substitution	potentially less severe	likely pathogenic	classic infantile	+
	Ch7	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch8	I16	c.2331+2T>A	p.?	substitution	very severe	pathogenic	unknown (disease-associated)	-*

Supplementary Table S2 continued.

	pt ID	Location	DNA	Protein	Type of variant (DNA)	predicted severity	ACMG classification	phenotype with null allele	CR IM
CHILDHOOD ONSET PATIENT	**Ch9**	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch10	I17	c.2481+102_2646+31del	p.(Gly828_Asn882del)	deletion	very severe	pathogenic	classic infantile	+
	Ch11	E6	c.1051del	p.(Val351Cysfs*41)	deletion	very severe	pathogenic	unknown (disease-associated)	-
	Ch12	E15	c.2066_2070dup	p.(Ala691Serfs*7)	duplication	very severe	pathogenic	unknown (disease-associated)	-
	Ch13	E15	c.2135T>C	p.(Leu712Pro)	substitution	less severe	likely pathogenic	unknown (disease-associated)	+
	Ch14	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch15	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	Ch16	E10	c.1548G>A	p.(Trp516*)	substitution	very severe	pathogenic	classic infantile	-*
	Ch17	E14	c.1933G>A	p.(Asp645Asn)	substitution	potentially less severe	pathogenic	classic infantile	+ #
ADULT ONSET PATIENT	**A1**	E10	c.1548G>A	p.(Trp516*)	substitution	very severe	pathogenic	classic infantile	-*
	A2	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A3	I17	c.2481+102_2646+31del	p.(Gly828_Asn882del)	deletion	very severe	pathogenic	classic infantile	+
	A4	E5	c.925G>A	p.(Gly309Arg)	substitution	potentially less severe	likely pathogenic	classic infantile	+
	A5	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A6	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A7	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A8	I6	c.1076-22T>G	p.?	substitution	potentially mild	pathogenic	childhood	?
	A9	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A10	E10	c.1548G>A	p.(Trp516*)	substitution	very severe	pathogenic	classic infantile	-*
	A11	E5	c.896T>C	p.(Leu299Pro)	substitution	potentially less severe	likely pathogenic	classic infantile	+
	A12	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A13	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A14	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A15	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A16	I6	c.1076-22T>G	p.?	substitution	potentially mild	pathogenic	childhood	?
	A17	I17	c.2481+102_2646+31del	p.(Gly828_Asn882del)	deletion	very severe	pathogenic	classic infantile	+
	A18	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	*
	A19	E5	c.925G>A	p.(Gly309Arg)	substitution	potentially less severe	likely pathogenic	classic infantile	+
	A20	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A21	E2	c.525del	p.(Glu176Argfs*45)	deletion	very severe	pathogenic	classic infantile	-*
	A22	E5	c.925G>A	p.(Gly309Arg)	substitution	potentially less severe	likely pathogenic	classic infantile	+

All patients included in this study share the same variant on the first GAA allele, c.-32-13T>G, and a second variant on allele 2 as listed. Ch1-C1h7: patients with childhood onset Pompe disease; A1-A22: patients with adult onset Pompe disease. E= exon, I= intron, += positive; -= negative; ?= unknown. *CRIM, as published by Bali et al. 2012. *Predicting cross-reactive immunological material (CRIM) status in Pompe disease using GAA mutations: lessons learned from 10 years of clinical laboratory testing experience. Am J Med Genet C Semin Med Genet. 160C(1):40-9*; #= CRIM as published by Kishnani et al. 2006. *Chinese hamster ovary cell-derived recombinant human acid alpha-glucosidase in infantile-onset Pompe disease J Pediatr. 149(1):89-97.*

Supplementary Table S3. Description of genes significantly differentially expressed in adult onset compared to childhood onset Pompe disease. These genes were revealed by pairwise comparisons of differentially expressed genes in Pompe disease affected individuals from studies 1 plus 2 (corrected for controls).

Supp. Table S3A.

Gene symbol	Chromosomal location	Subcellular location	Molecular function	Tissue expression	Signaling pathway	Biomarker application(s)	Additional information
NRK	Xq22.3	cytosol	-receptor signaling protein -serine/threonine kinase activity -ATP binding	overexpressed in: -ovary -adrenal gland	-TNF receptor signaling pathway		The encoded protein may be involved in the induction of actin polymerization in late embryogenesis

Description of genes significantly downregulated in adult compared with childhood onset Pompe disease.

Supp. Table S3B.

Gene symbol	Chromosomal location	Subcellular location	Molecular function	Tissue expression	Signaling pathway	Biomarker application (s)	Additional information
STC1	8p21.2	-extracellular -nucleus -plasma membrane	-hormone activity	overexpressed in: -thyroid -pituitary	ND	Diagnosis (ovarian cancer)	The protein may play a role in the regulation of renal and intestinal calcium and phosphate transport, cell metabolism, or cellular calcium/phosphate homeostasis. Overexpression of human stanniocalcin 1 in mice produces high serum phosphate levels, dwarfism, and increased metabolic rate
MAOA	Xp11.3	-mitochondrion -cytosol	-primary amine oxidase activity -flavin adenine dinucleotide binding -serotonin binding	ubiquitous	-Enzymatic degradation of dopamine by monoamine oxidase -Tryptophan metabolism -Tyrosine metabolism Glycine, serine and threonine metabolism -Arginine and proline metabolism -Histidine metabolism Phenylalanine metabolism -Metabolic pathways		Catalyzes the oxidative deamination of biogenic and xenobiotic amines and has important functions in the metabolism of neuroactive and vasoactive amines in the central nervous system and peripheral tissues

Description of genes significantly upregulated in adult compared with childhood onset Pompe disease.

Supplementary Table S4. Number of genes differentially expressed in adult Pompe disease compared to adult controls and in childhood Pompe compared with childhood controls in the studies 1 plus 2.

Supp. Table S4A.

Comparison	Study	Down	Up	Total of Genes
adult patient vs adult control	1 (A-B)	3292	3315	6607
adult patient vs adult control	2	28	25	53

Number of genes differentially expressed in adult onset Pompe disease compare with adult controls in studies 1 plus 2.

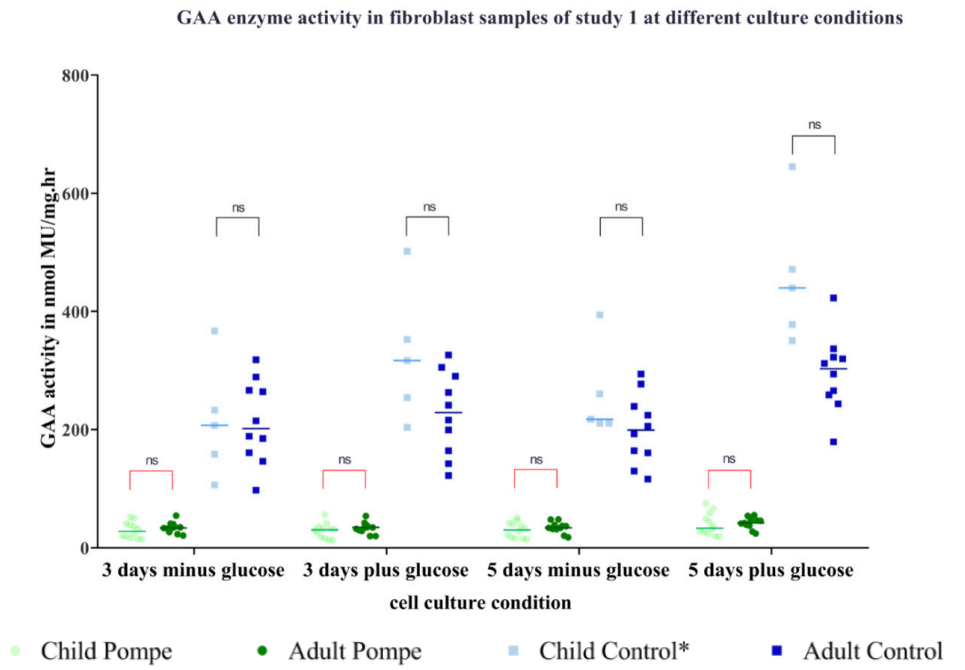
Suppl. Table S4B.

Comparison	Study	Down	Up	Total of Genes
childhood patient vs childhood control	1 (A-B)	2002	2424	4426
childhood patient vs childhood control	2	49	40	89

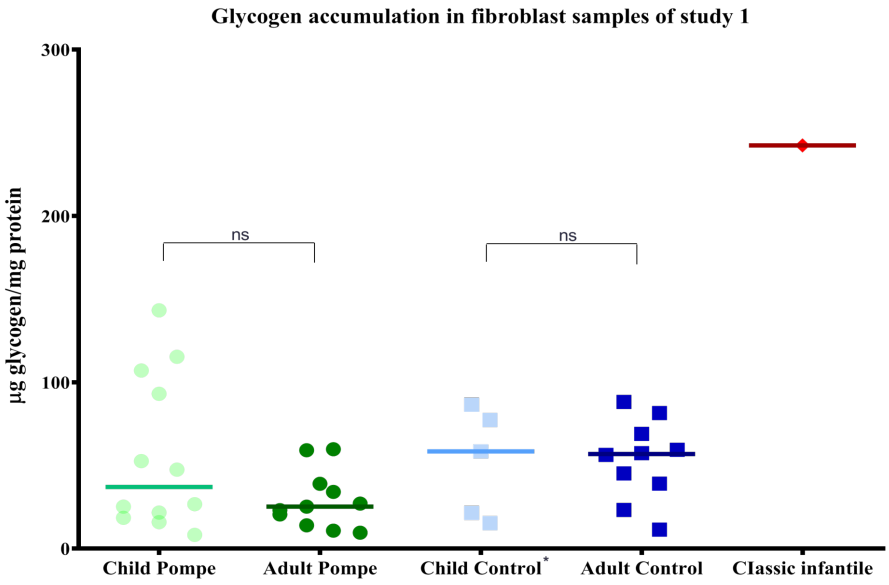
Number of genes differentially expressed in childhood onset Pompe disease compared with childhood controls in studies 1 plus 2.

Suppl. Table S5. The corresponding number identification of childhood and adult onset Pompe patients in the tables compared to the multidimensional scaling plot (MDS, Figure 3A)

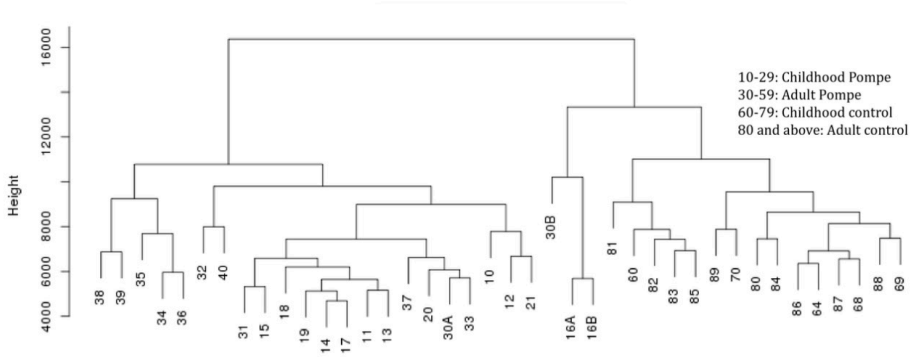
Clinical phenotype	Tables Patient ID	MDS patient ID	Clinical phenotype	Tables Patient ID	MDS patient ID
Childhood onset Pompe patients	CH1	12	Adult onset Pompe patients	O1	30A, 30B
	CH2	10		O2	37
	CH3	21		O3	40
	CH4	11		O4	38
	CH5	14		O5	39
	CH6	15		O6	34
	CH7	16A, 16B		O7	36
	CH8	17		O8	31
	CH9	13		O9	32
	CH10	20		O10	33
	CH11	18		O11	35
	CH12	19			



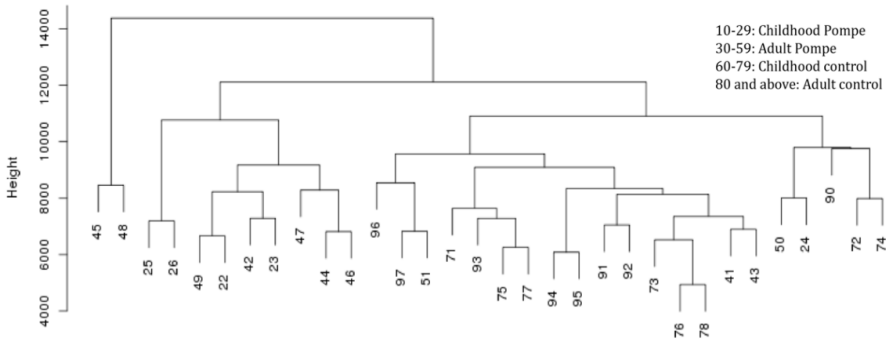
Supplementary Figure S1. GAA enzyme activity in fibroblast samples of study 1 at different culture conditions. The cells were routinely grown in DMEM as described in Material and Methods. For measurement of the lysosomal glycogen content the medium was replaced by glucose free medium for either 3 or 5 days where after the cells were harvested and the GAA activity measured. The horizontal line presents the median value. As in Fig.4A, ns = non-significant (P>0.05). The differences between the GAA activities in patients *versus* controls were in all instances significant (P<0.05). The same methodology was used as in Fig 4A. *Fibroblasts from foreskin were not included. The 2way ANOVA test was used for the analysis



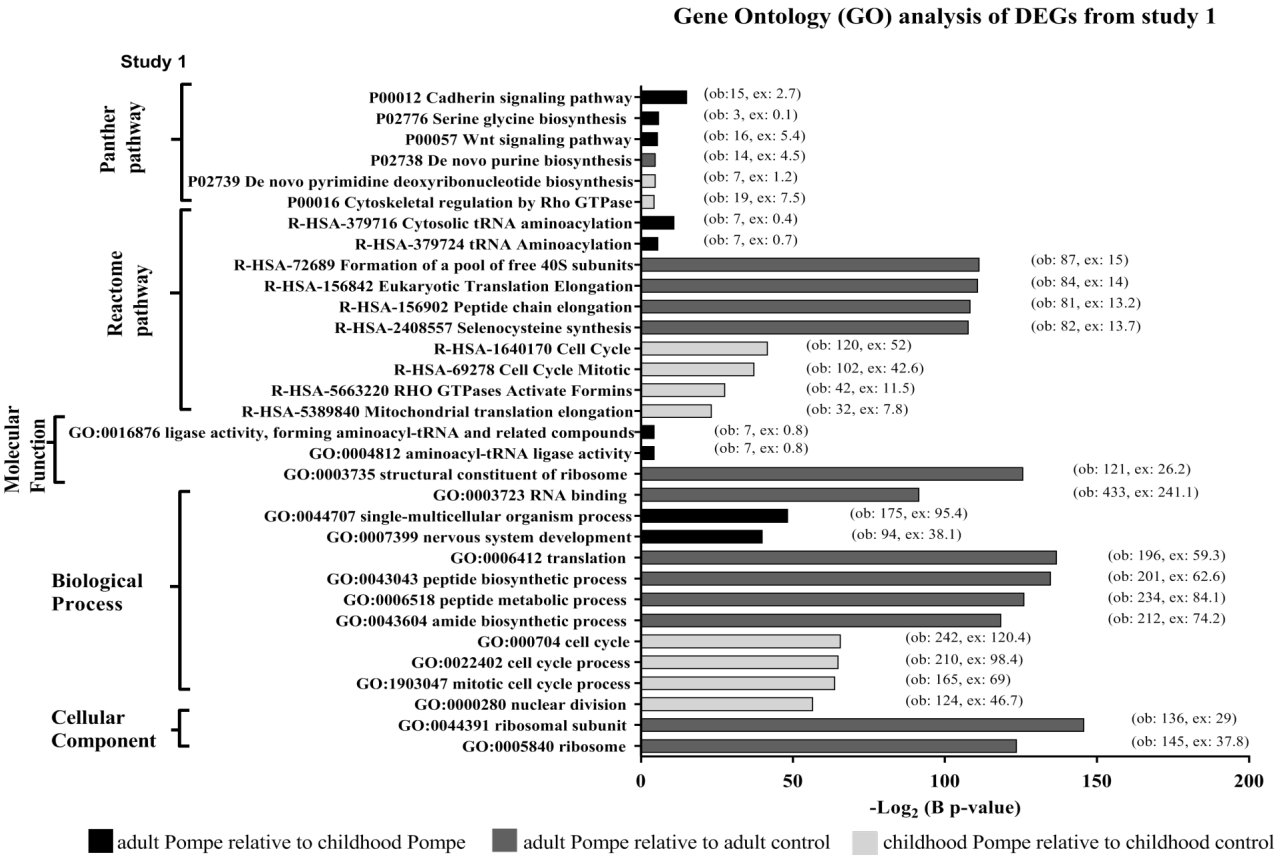
Supplementary Figure S2. Glycogen accumulation in fibroblast samples of study 1. The horizontal line presents the median value. There were no significant differences between the groups (ns; P>0.05). Only the single cell line from a patient with classic infantile Pompe disease that was included as a control in this study showed significant accumulation of glycogen when compared with all other samples. The same methodology and statistical analysis was used as in Fig 4A. *Fibroblasts from foreskin were not included.



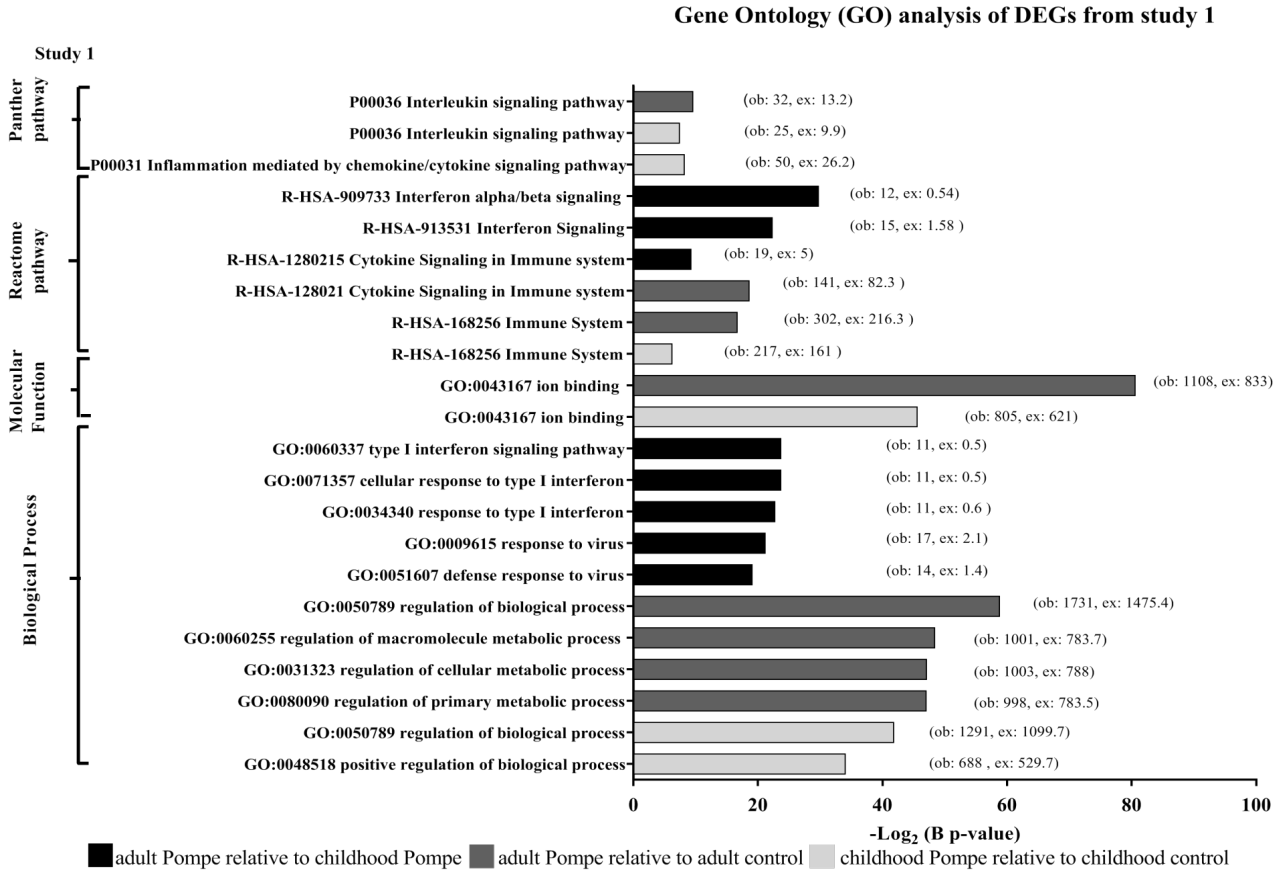
Supplementary Figure S3A. Hierarchical cluster analysis by RNA sequencing comparing Pompe patients and healthy controls from studies 1A plus B. The patients cluster together in the left branch, and the controls in the right sub-branch of the top-right branch. The left sub-branch of the top-right branch includes one sample of a child and one of an adult onset of Pompe disease from study 1B and one sample of a child from study 1A. These three samples are quite different from the controls however, as expected. Within the patients' tree, the left branch is all patients, whereas the right branch is all childhood patients, together with 4 adult patients. Within the controls tree, there is a clear separation between a branch with only child controls, and a branch with a mixture of child and adult.



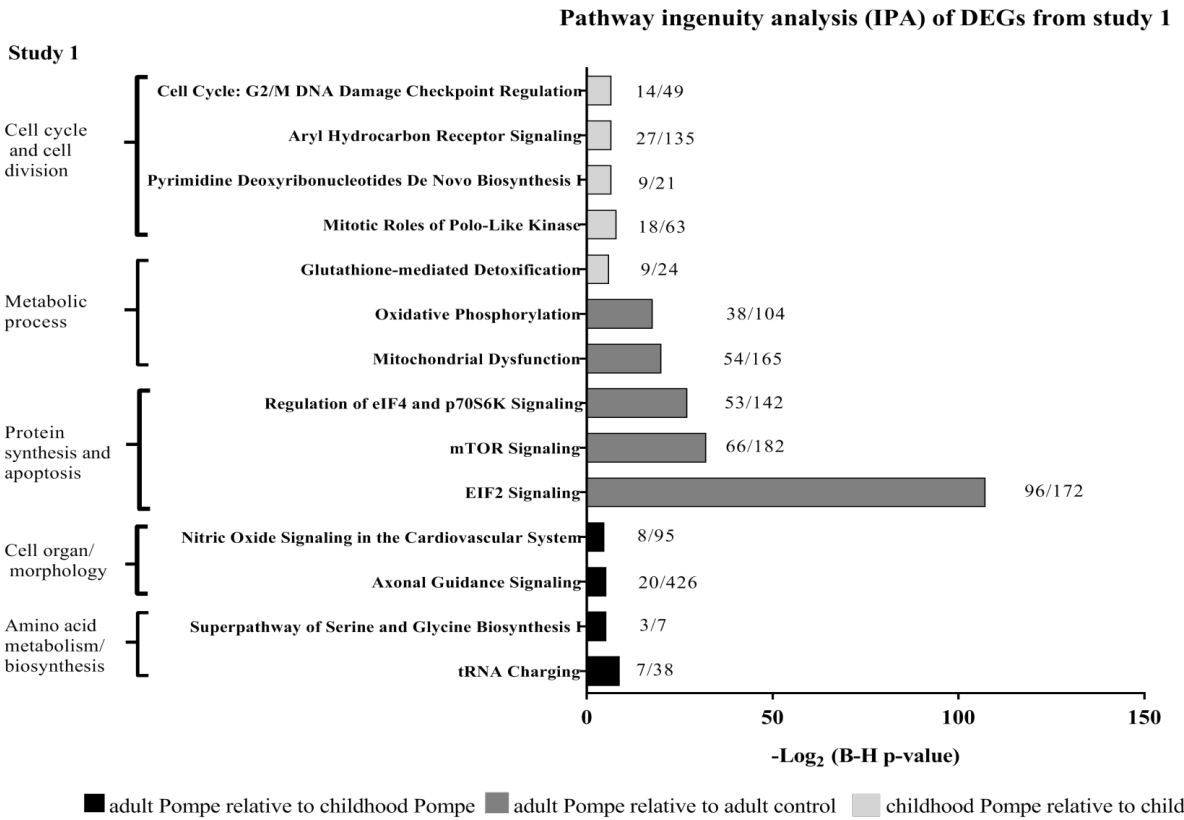
Supplementary Figure S3B. Hierarchical cluster analysis by RNA sequencing comparing Pompe patients and healthy controls from study 2. Most patient samples cluster apart from the controls except for five (24, 41, 43, 50, 51) that cluster with the controls. Two samples from (adult) patients (45 and 48) cluster together apart from all other.



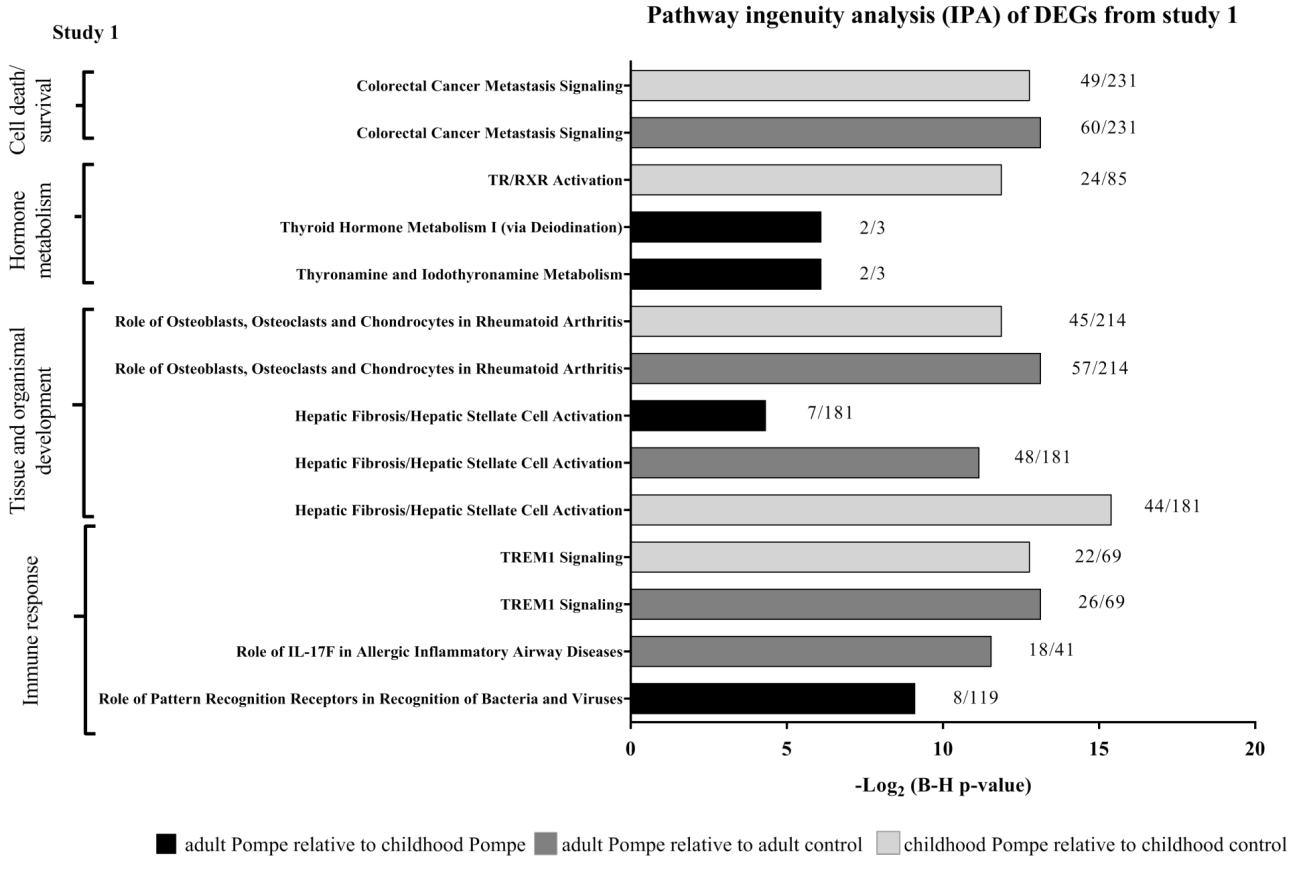
Supplementary Figure S4. Gene Ontology (GO) analysis of DEGs from study 1. The-DEGs detected by RNA-Seq were investigated for significantly enriched GO terms (B p-value <0.05) using the Panther classification system. All significant GO terms were ranked according to the Bonferroni correction p-values. Bars represent p-values adjusted by Bonferroni (B) correction multiple test (-Log_2 B p-value) of the most significant GO terms. Colors represent the GO enriched terms in adult *versus* childhood onset Pompe disease, relative to healthy controls. The GO terms are listed by category: Panther pathways, Reactome pathways, molecular function (MF), biological process (BP) and cellular component (CC). GOID and the term are included. Number of observed genes (ob), and number of expected genes (ex) are given at the end of the bars. **Supp. Fig. S4A**, depicts the set of downregulated genes, **Supp. Fig. S4B** the set of upregulated genes.



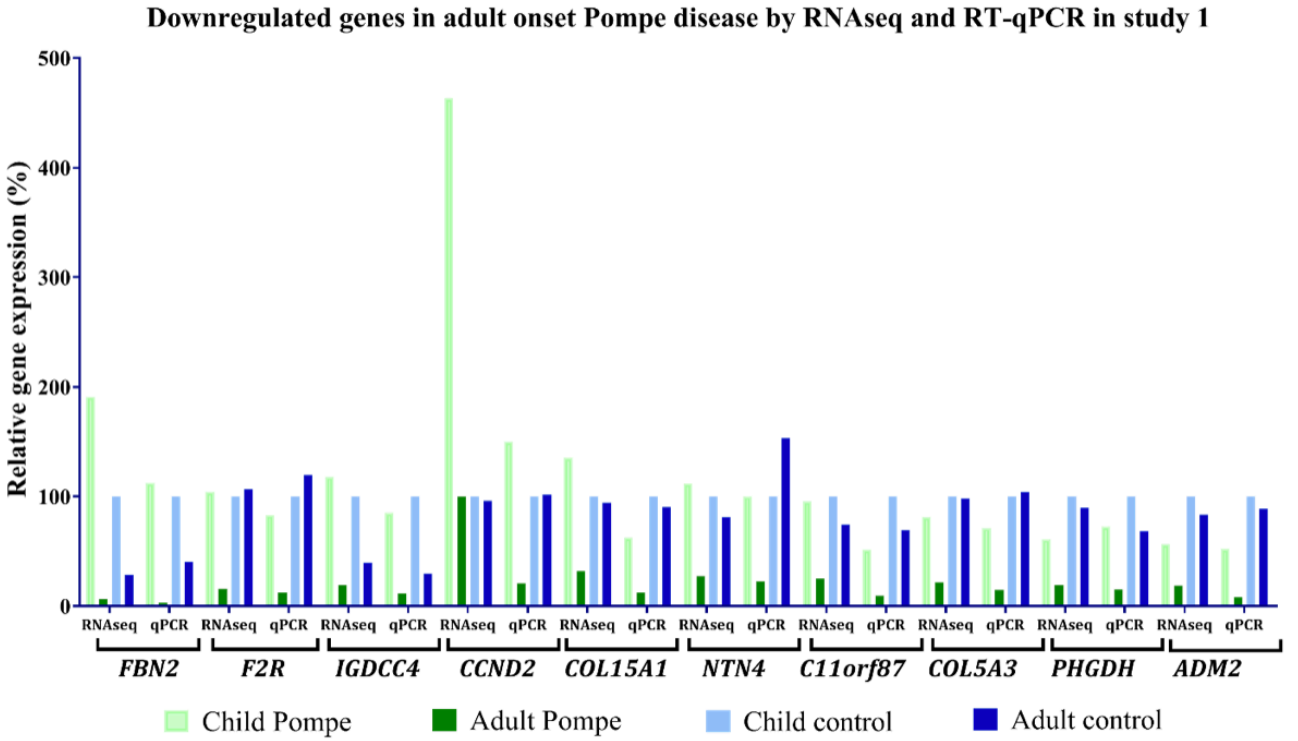
Supplementary Figure S4B



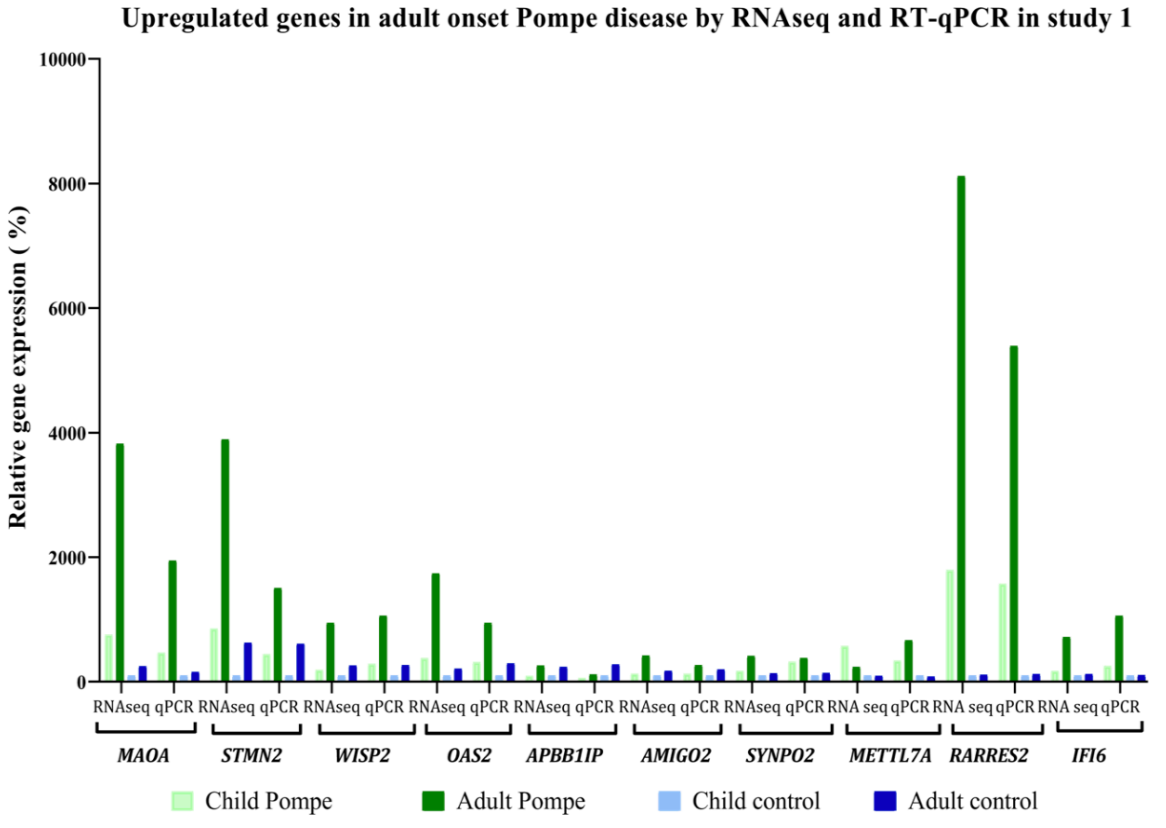
Supplementary Figure S5. Pathway ingenuity analysis (IPA) of DEGs from study 1. The DEGs detected by RNA-Seq were investigated for identification of pathways (B-H p-value <0.05) using Ingenuity software and ranked accordingly. Bars represent adjusted p-values (-Log_2 B-H p-value) of the most significant pathways. Colors represent the pathways identified in adult *versus* childhood onset disease relative to healthy controls. Top functions of the pathways are listed at the right. The figures at the end of the bars represent the number of DEGs (in patients) per pathway relative to the total number of genes assigned to that pathway. **Supp. Fig. S5A**, depicts the set of downregulated genes by pathway, **Supp. Fig. S5B** the set of upregulated genes by pathway.



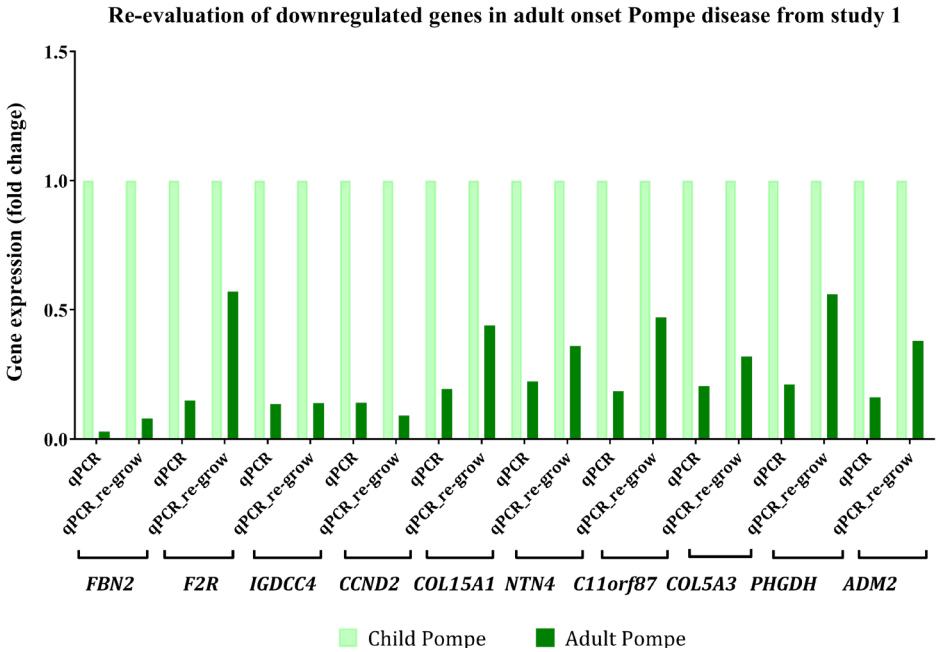
Supplementary Figure S5B



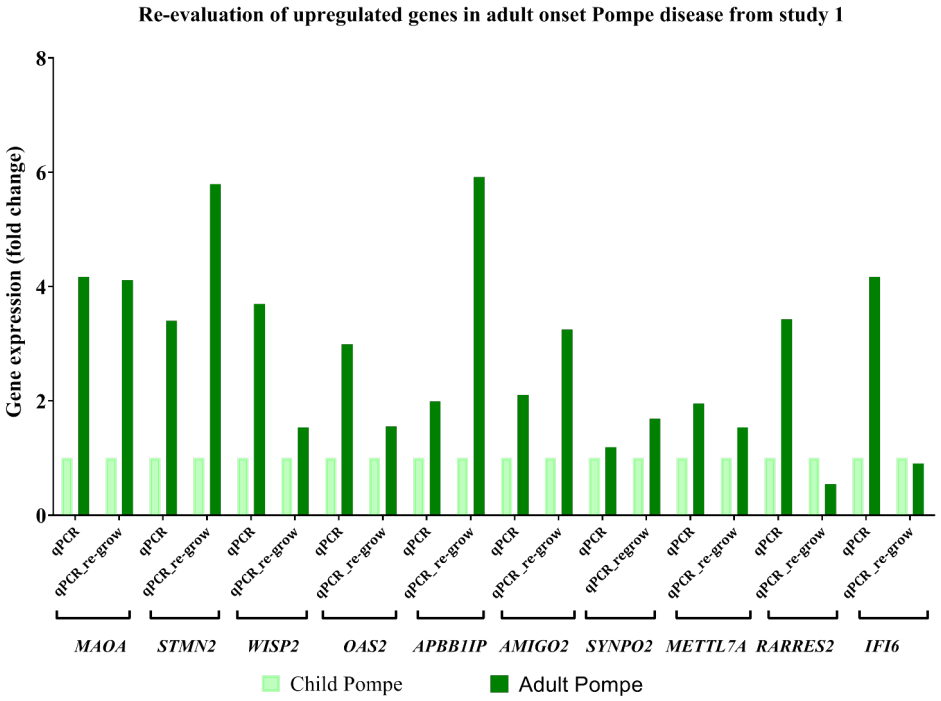
Supplementary Figure S6A. Downregulated genes in adult onset Pompe disease by RNAseq and RT-qPCR (study 1). Ten downregulated genes identified by RNAseq were analyzed by qPCR to validate the quality of the RNAseq data. The relative gene expression levels of selected genes measured by RT-qPCR were in good agreement with the RNA-sequencing data



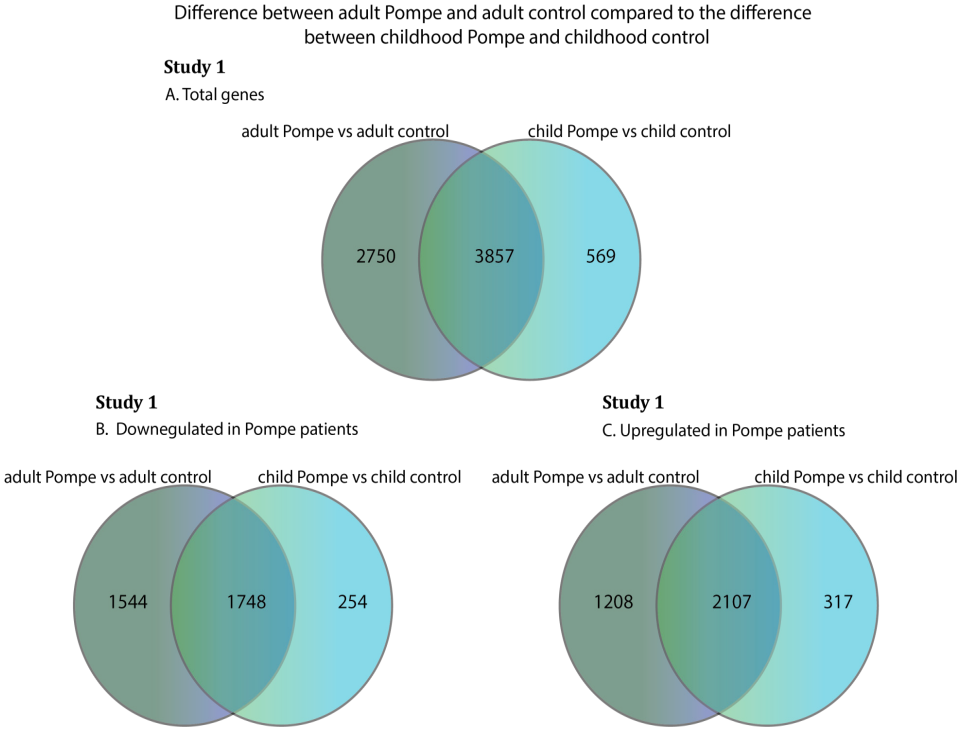
Supplementary Figure S6B. Upregulated genes in adult onset Pompe disease by RNAseq and RT-qPCR (study 1). Ten upregulated genes identified by RNAseq were analyzed by qPCR to validate the quality of the RNAseq data. The relative gene expression levels of most selected genes measured by RT-qPCR were in good agreement with the RNA-sequencing data.



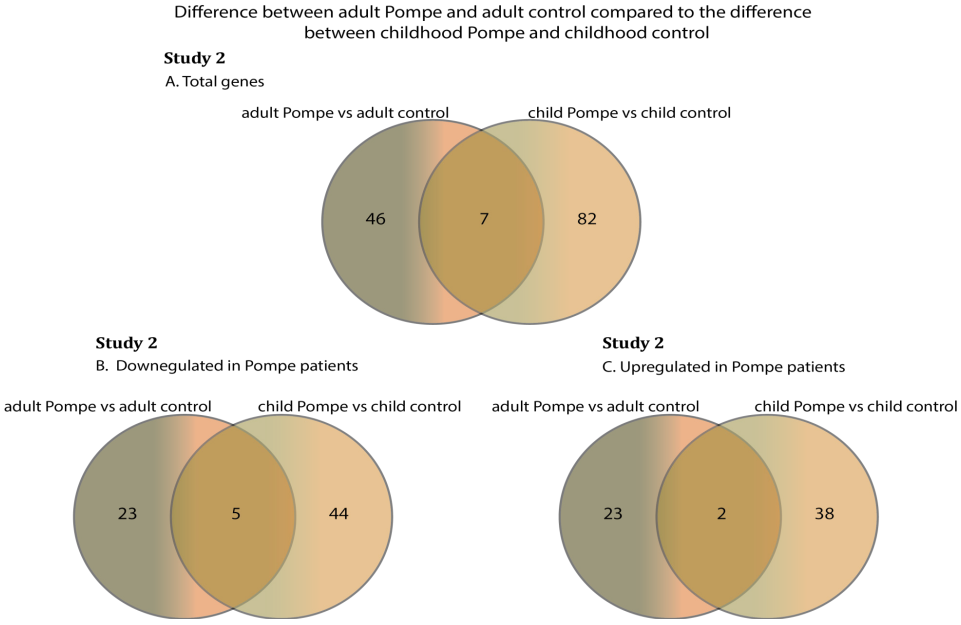
Supplementary Figure S6C. Re-evaluation of downregulated genes in adult onset Pompe disease from study 1. The fibroblasts of patients included in study 1 were re-grown and gene expression levels of downregulated genes were re-evaluated by RT-qPCR in order to analyze dependency on culture conditions and reproducibility.



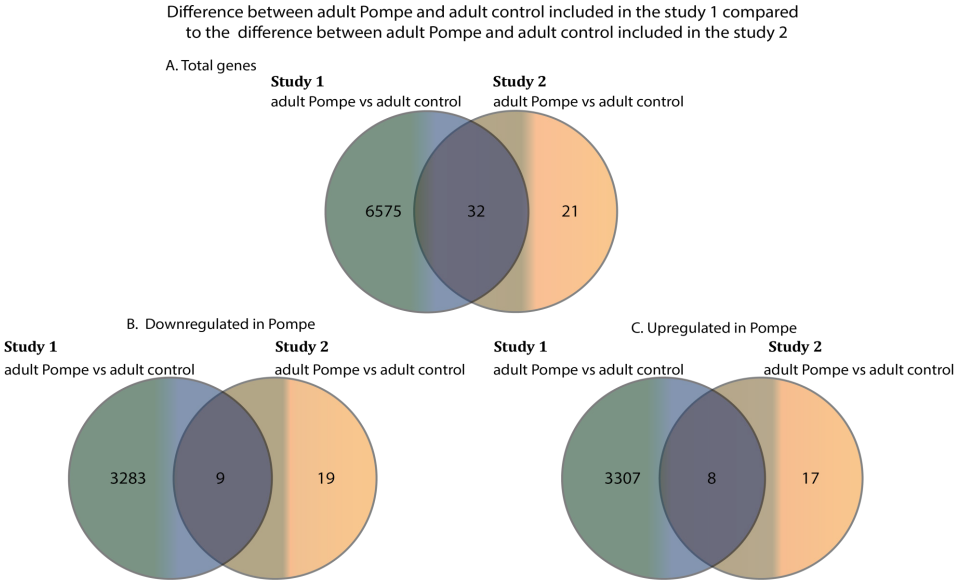
Supplementary Figure S6D. Re-evaluation of upregulated genes in adult onset Pompe disease from study 1. The fibroblasts of patients included in study 1 were re-grown and gene expression levels of upregulated genes were re-evaluated by RT-qPCR in order to analyze dependency on culture conditions and reproducibility.



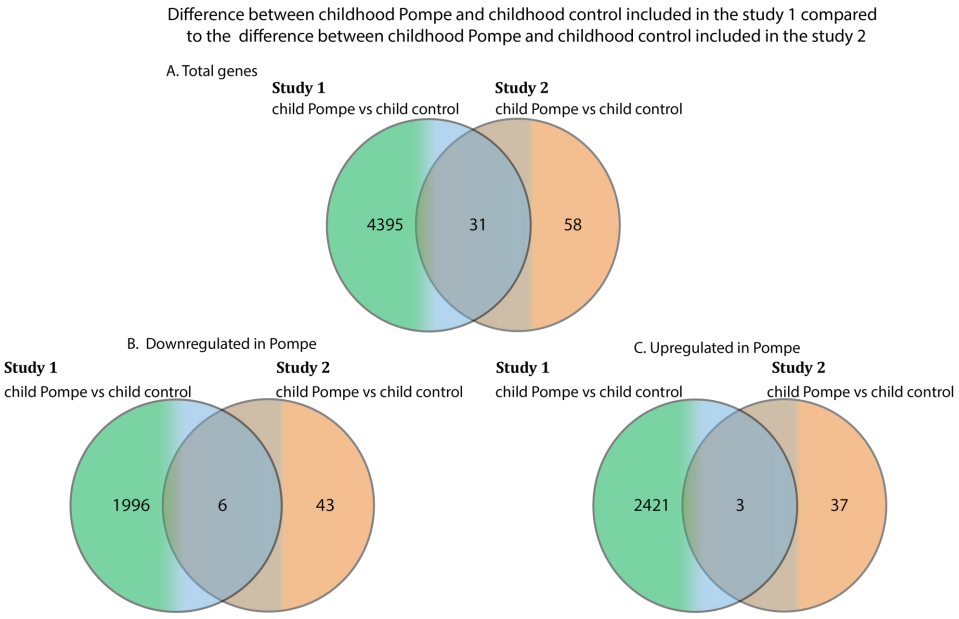
Supplementary Figure S7. Venn diagram showing the difference between affected adults and children compared to unaffected controls from study 1. (A) A total of 3857 genes are differentially expressed in fibroblasts of the patients. **(B)** 1748 genes are downregulated and **(C)** 2107 are up-regulated in the patients. The DEGs were filtered using a cut off value of fold change (FC) < -1.5/>1.5, and false discovery rate (FDR) <0.05.



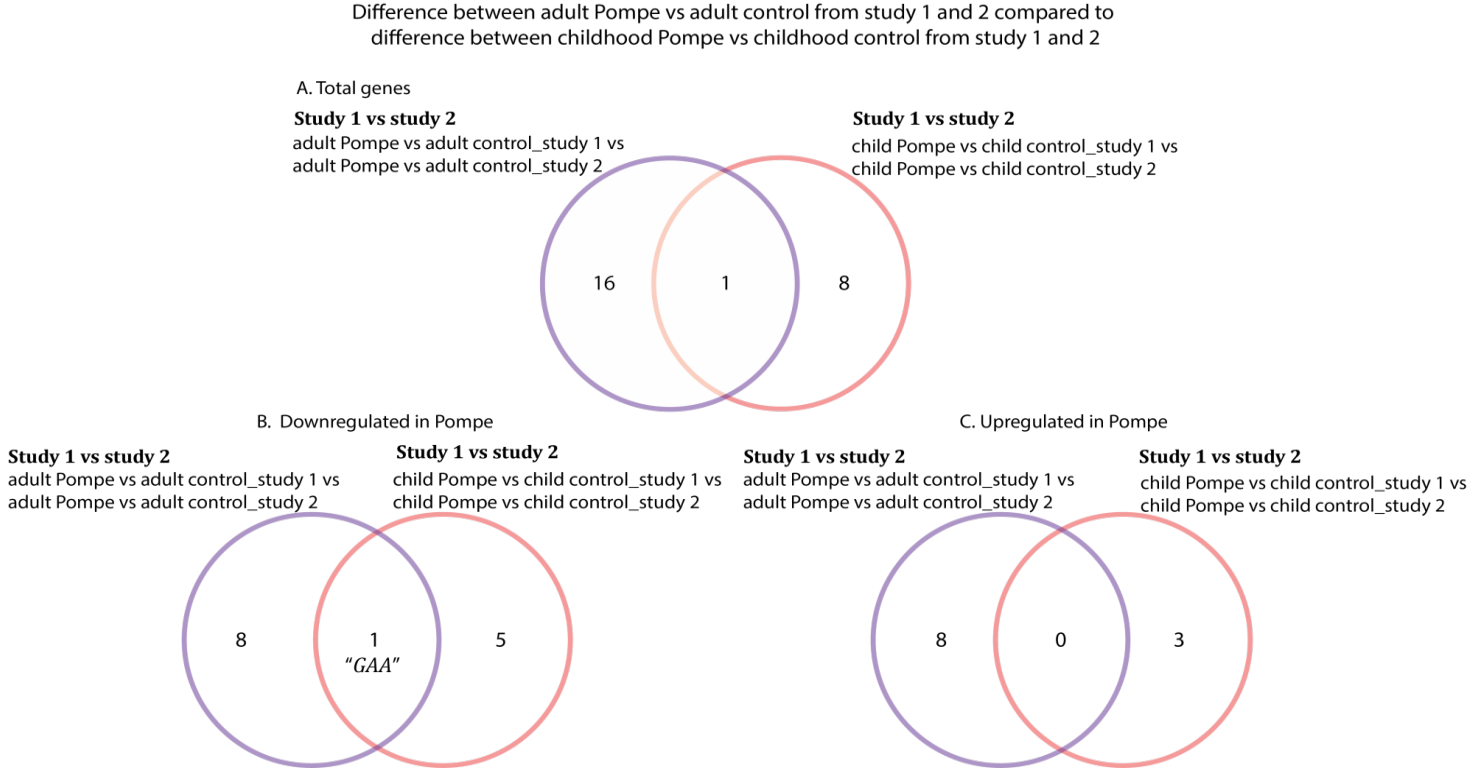
Supplementary Figure S8. Venn diagram showing the difference between affected adults and children compared to unaffected controls from study 2. (A) A total of 7 genes are differentially expressed in Pompe patients. **(B)** 5 genes are downregulated and **(C)** 2 are upregulated genes in the patients. The DEGs were filtered using a cut off value of fold change (FC) < -1.5/>1.5, and false discovery rate (FDR) <0.05.



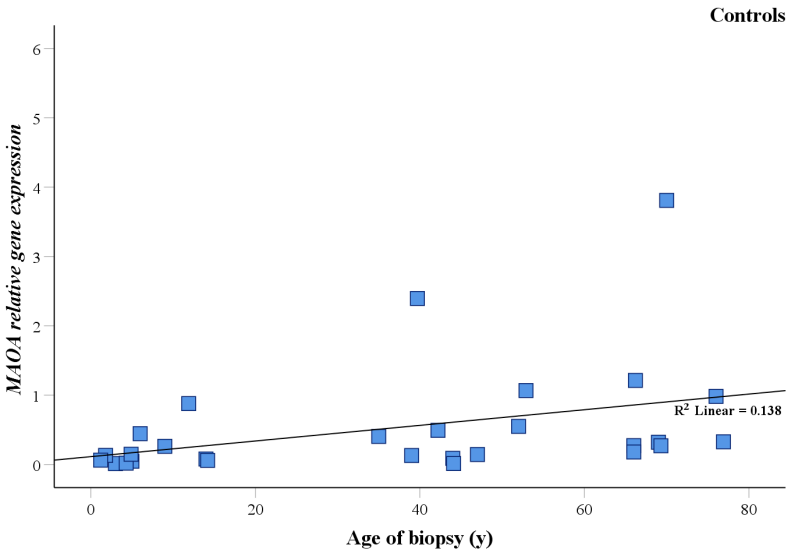
Supplementary Figure S9. Venn diagram showing the difference between affected adults and unaffected adult controls combining studies 1 and 2. (A) A total of 32 genes are differentially expressed in adult patients compared to adult controls **(B)** 9 genes are downregulated in the patients and **(C)** 8 are upregulated.



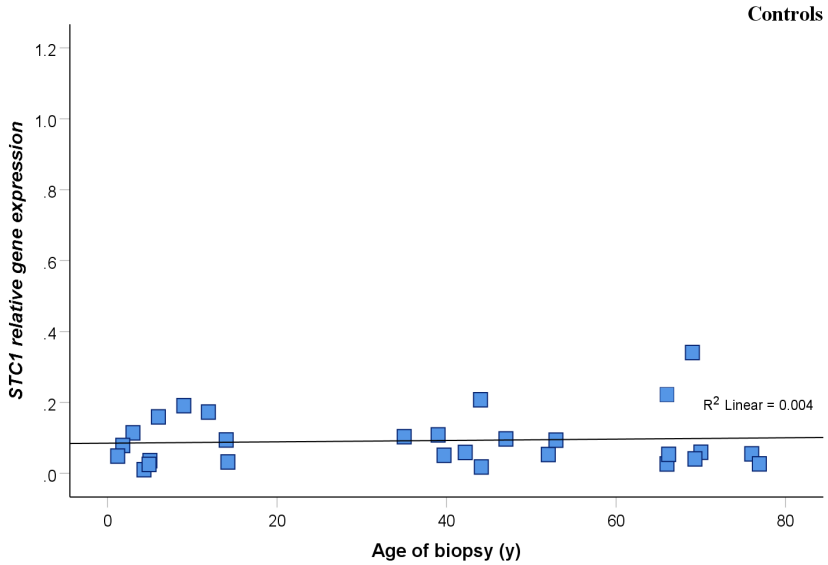
Supplementary Figure S10. Venn diagram showing the difference between affected children and unaffected childhood controls combining studies 1 and 2. (A) A total of 31 genes are differentially expressed in childhood patients compared to controls. **(B)** 6 genes are downregulated in the patients and **(C)** 3 are upregulated.



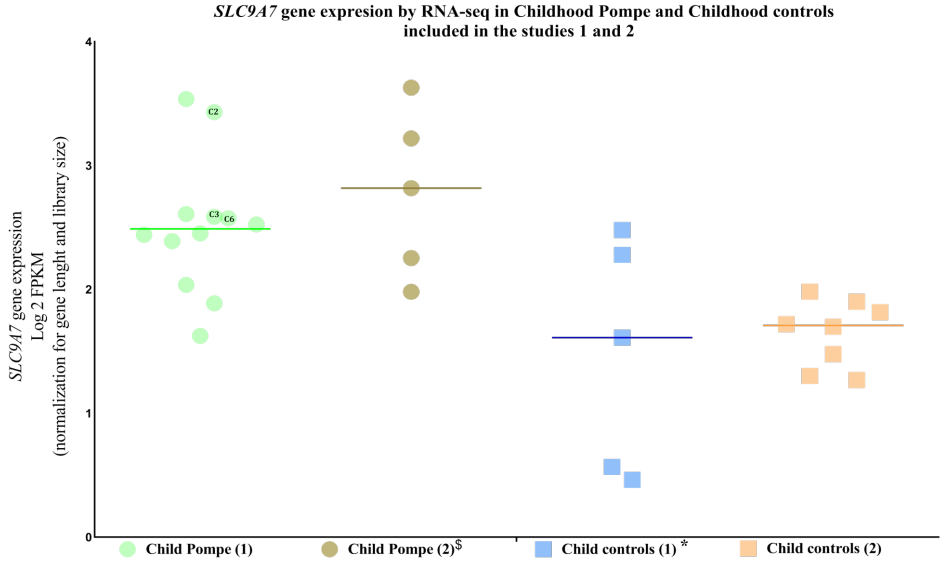
Supplementary Figure S11. Venn diagram showing the number of DEGs comparing both adult patients with adult controls as well and affected children with childhood controls combining the data sets from studies 1 and 2 (**A)** A total of 17 genes are differentially expressed in adult patients compared to controls, and 9 are differentially expressed in childhood patients compared to controls. **(B)** 9 are downregulated in adult onset Pompe disease and 6 in childhood onset disease. Notably, the single one gene downregulated in both adult as well as childhood patients, compared with age matched controls, is the *GAA* gene, which showed over two-fold decreased in Pompe patients. **(C)** 8 genes are upregulated in affected adults and 3 in affected children compared to age matched controls.



Supplementary Figure S12A. *MAOA* gene expression by age in controls of studies 1- 3. The expression level was measured by RT-qPCR. Samples from foreskin were not included. Linear regression analysis was performed using SPSS.



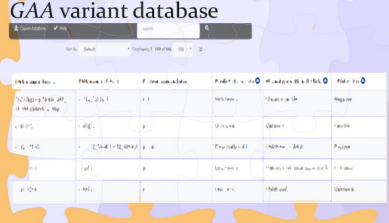
Supplementary Figure S12B. *STC1* gene expression by age in controls of studies 1- 3. The expression level was measured by RT-qPCR. Samples from foreskin were not included. Linear regression analysis was performed using SPSS.



Supplementary Figure S13. *SLC9A7* gene expression by RNA-seq in the patients included in studies 1 plus 2. C2= childhood Pompe disease with scapular winging and moderate (symptoms so severe they interfere with daily life) status at time of biopsy; C3 and C6= childhood Pompe disease with scapular winging and mild phenotypes (symptoms barely interfere with daily life) at time of biopsy**.** The horizontal line represents the median value. *Fragments per kilobase of exon per millon reads mapped. (1/2) = study number: $scapular winging unknown, * foreskin fibroblasts were not included.

Pompe disease

GAA variant database



CHAPTER 7

GENERAL DISCUSSION AND PERSPECTIVES

7

General discussion and perspectives

As briefly mentioned in the Scope of this thesis Pompe disease can be addressed as a glycogen storage disorder (GSD), as a lysosomal storage disorder (LSD), and also as a neuromuscular metabolic disorder (NMD). For understanding the cellular pathology, the pathologic cascade of events, and the clinical phenotype of Pompe disease information can be gathered by searching for features that Pompe disease has in common with other diseases within these three disease categories.

With regard to glycogen storage, Pompe disease is different from all other GSDs in that the major cellular pathways of glycogen metabolism -synthesis and turnover- are not seriously affected. For instance, energy deficit has not been documented in Pompe disease. The crux lays in the fact that the lysosomal contribution to glycogen metabolism is far smaller than the cytoplasmic. What primarily counts in Pompe disease is the lysosomal pathology and dysfunction that is caused by physical overload of the lysosomal system. This affects a cascade of lysosome related processes and pathways, most prominently autophagy, endocytosis, and cellular renewal. In this respect, the LSDs share many aspects of cellular pathology, but they differ in clinical signs largely depending on which compound stores in which cell type based on a given enzyme deficiency or other lysosomal defect. The grouping of Pompe disease as NMD follows from the sharing of clinical symptoms related to skeletal-muscle dysfunction. In Pompe disease the dysfunction derives from the muscle itself while in other NMDs it may be caused by neural defects.

Above all Pompe disease is a genetic disorder caused by disease-associated sequence variants in the gene that encodes an enzyme called acid α-glucosidase (GAA). GAA is needed for the stepwise degradation of glycogen to glucose, but only for the portion of total cellular glycogen that ends up in lysosomes by autophagy. The genetic information provided by the *GAA* gene and the surrounding genetic network are together responsible for GAA activity, the metabolism of glycogen, the process of autophagy and the synthesis of GAA. In concert these determine how much lysosomal glycogen is degraded.

The concepts 'lysosome', 'glycogen storage', 'acid α-glucosidase (GAA)', and 'muscle cell damage resulting in clinical signs' were introduced in **Chapter 1** in their simplest ways. Figure 1 shows how 2 *GAA* alleles -Pompe disease is autosomal recessive- primarily determine how much GAA is synthesized and in which forms (both fully functional, one partially functional the other fully dysfunctional, one partially present the other absent, and in any other thinkable combination). More and functionally better GAA provides faster and more efficient lysosomal glycogen degradation and slower development of lysosomal pathology. The figure illustrates in addition how genes other than *GAA* can influence the lysosomal glycogen turnover (surrounding genetic network). At the right side of this figure are genetic factors depicted involved in lysosomal enzyme synthesis and transport. A positive influence may upgrade GAA synthesis. A negative influence may lead to reduction of GAA synthesis. At the left side of the figure are genes depicted that are involved in cytoplasmic glycogen metabolism (extra-lysosomal) and genes involved in autophagosome formation and functioning. If the lysosomal import of glycogen is low, relatively little GAA activity may suffice to prevent lysosomal storage, but if the import is high the demand for GAA may be higher too. The figure also shows the contribution of the 'cellular organization' (for instance, a muscle fiber is not a liver cell. One cell type needs more GAA than the other, one stores more glycogen than the other, one metabolizes more glycogen than the other outside the lysosome, one contracts -causing mechanical damage- the other does not, and so on). Figure 1 is not more than a hypothetical model, but it does show how the clinical phenotype of Pompe disease relates to the amount and quality of the GAA, which in turn is primarily dictated by the two *GAA* alleles and the type of disease-associated sequence variants that they incorporate. The figure illustrates in addition how the *GAA* genotype-phenotype correlation in Pompe disease is potentially influenced by the expression level and quality of numerous other genes (surrounding genetic network) that are addressed as genetic modifiers.

The concept of a lysosomal storage disorder (LSD) was proposed around 50 years ago by HG Hers [1], following the discovery that 'Pompe disease' was caused by lysosomal enzyme deficiency. Soon after that, the ground-breaking concept of "cross correction" introduced by Elizabeth Neufeld [2] lead to the notion of enzyme replacement therapy (ERT). Since then, there has been a huge progress and expansion in the field with regard to the level of understanding of both the molecular and clinical basis of LSDs. Treatments for these disorders have advanced rapidly and now provide several safe and effective enzyme replacement therapies. However, of the more than 70 LSDs identified so far, approved enzyme treatment is currently available for some including Gaucher disease, (type I and III), Fabry disease, Mucopolysaccharidosis (type I, II, IV, VI, and VII),

Mannosidosis, and Pompe disease, which is the first inherited muscle disorder for which therapy has become available. However, multidisciplinary research remains a crucial factor for understanding the whole impact of the disease and the clinical outcome.

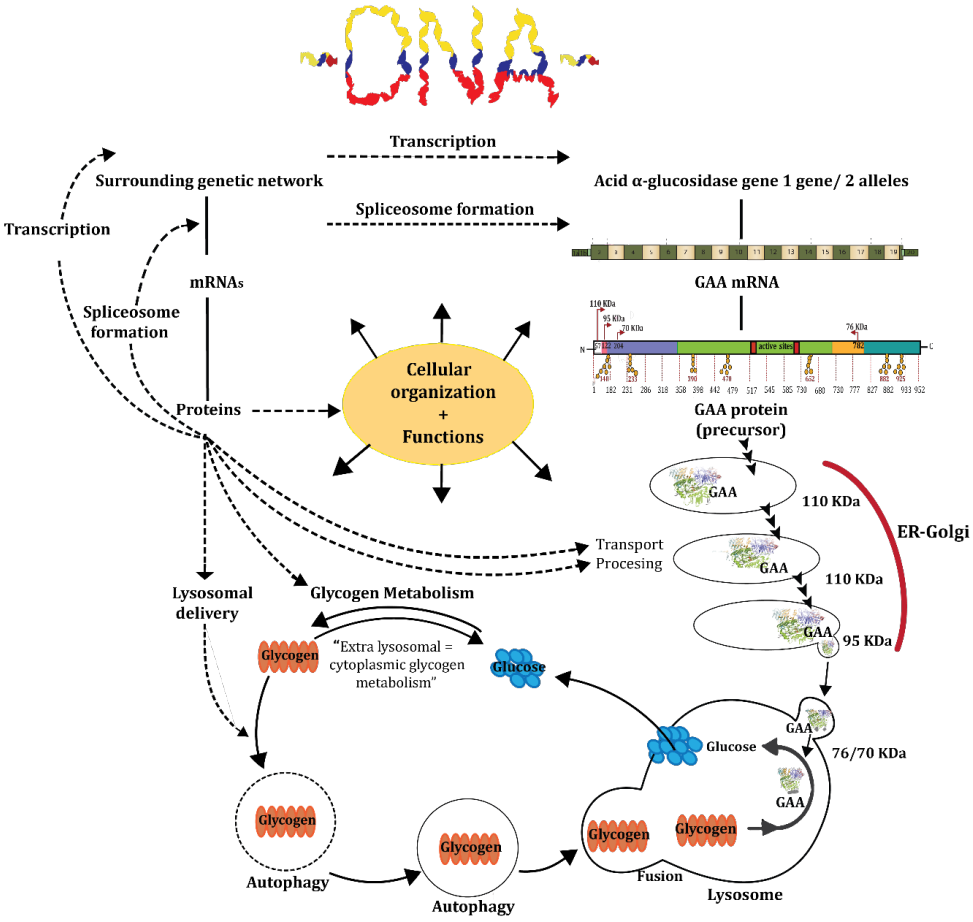


Figure 1. How expression of *GAA* dictates and genetic background modifies lysosomal degradation. This model illustrates how the amount of functionally active GAA is primarily determined by the 2 *GAA* alleles and how a substantial number of other genes (surrounding genetic network) exert their impact on the lysosomal glycogen degradation. The latter are involved in processes related to the synthesis, processing and transport of GAA, the formation and functioning of cell organelles (including autophagy), and the extra-lysosomal glycogen metabolism.

The research described in this thesis was aimed to determine which enzymatic diagnostic assay to use for which purpose. For this first aim a considerable number of diagnostic cases from over the past 28 years was reviewed and the information

used for determining the disease threshold of GAA enzyme activity in different assay systems **(Chapter 2)**. Our findings suggest that Pompe disease can be excluded when leukocytes serve as sample source using either glycogen or 4MUG as substrate. However, if the GAA activity is below the normal range using glycogen, deficiency should be confirmed using 4MUG, which excludes cases of *GAA* pseudodeficiency. Clearly, based on decades of experience with large numbers of diagnostic samples of various kinds and assay procedures we conclude that cultured skin fibroblasts are by far the best sample source for enzymatic diagnosis. The method provides the best discrimination between classic infantile and non-classic phenotypes; mixed leukocytes do not. However, using fibroblasts as sample source has it's drawbacks, such as relative invasiveness and the time needed to establish the fibroblast cultures. This limits the use of fibroblasts as first assay, especially in baby's that need a fast diagnosis in order to allow early start of treatment with ERT, which can be lifesaving in these patients. With respect to measuring the activity in blood our findings indicate that using glycogen as natural substrate provides better distinction between affected and unaffected than 4-MUG. In practice very few laboratories use glycogen because it makes the assay more complicated. It would be valuable to be able to establish an enzymatic assay than can rapidly be performed and that has the same qualitative outcomes as the fibroblast assay. Future research is required for this and might include downscaling sample size to allow a direct measurement of primary cells obtained from different sources such as skin or cheek mucosa.

Furthermore, this thesis aimed to define conceptual disease models by way of assigning a functional severity scale to each *GAA* variant not only based on predictive values but more so on clinical outcome. This second aim was meant to facilitate diagnosis, support genetic counseling in prediction of genetic risk and clinical course and guide in decision-making with regard to treatment options **(Chapter 3).** Such models allow the extrapolation of outcomes based on clinical information described in the literature via PubMed (previously: www.pompecenter.nl; now: www.pompevariantdatabase.nl) and brings knowledge of the genotype-phenotype correlation in Pompe disease a step forward **(Chapter 4)**. For instance, it has led to a third aim of this thesis i.e.-the finding of genes that can possibly explain differences in disease severity and progression between patients with the same or very similar *GAA* genotypes, be used as predictor of disease progression, and be a potential target for the development of alternative treatments **(Chapters 5 and 6).** Salient aspects of the various studies are discussed in the following paragraphs followed by a short final paragraph on perspectives for future research. The main findings of this thesis are summarized in Table 1.

Table 1. Main Findings and Perspectives

	Main findings	Perspectives
CHAPTER 2 **Enzymatic Diagnosis of Pompe Disease: lessons from 28 years of experience**	-The GAA activity measured in fibroblasts is the gold standard for diagnosing Pompe disease, which discriminates between clinical subtypes. -Assays in leukocytes using glycogen as substrate partly distinguish between clinical phenotypes but can result in false-positive diagnosis due to *GAA2* pseudodeficiency. - DBS assays are ideally suited for screening purposes but the relatively high incidence of false-positive and false-negative results calls for confirmation by other methods.	-International standardization of nomenclature and diagnostic protocols is needed, also in terms of which method to apply in which situation. -Newborn screening technical advances are recommended in order to distinguishing between Pompe disease and *GAA* pseudodeficiency in Asian populations.
Chapter 3 **Extension of the Pompe mutation database by linking** **disease-associated variants to clinical severity**	-The extended Pompe variant database not only reports the nature of *GAA* variants but also links *GAA* genotypes to phenotypes, which means a step forward in understanding the disease. -The database is foreseen to support diagnostic, prognostic, and genetic counseling.	-The Pompe variant database gains power if periodically updated and supplemented with sophisticated functional analysis of variants. - The database may also contribute to the development of personalized therapies.
CHAPTER 4 **The extended Pompe mutation database reveals distinct phenotypic spectra of Pompe disease-associated *GAA* variants**	- Homozygous and compound heterozygous *GAA* genotypes manifest genotype specific phenotypic spectra. - The broad clinical spectrum of patients with c.-32-13T>G / 'null' genotypes suggest that the impact of modifying factors is particularly large in this genotype.	- A search for genetic modifiers is strongly recommended as this may lead to new therapeutic opportunities.
CHAPTER 5 **The *ACE* I/D polymorphism does not explain heterogeneity of natural course and response to enzyme replacement therapy in Pompe disease**	-The broad clinical spectrum of patients with -32 13C>T / 'null' *GAA* genotypes is not explained in totality by genetic diversity of the ACE locus.	-Searching for alternative modifiers is warranted.
CHAPTER 6 **Clinical diversity of Pompe patients with c.-32-13T>G genotypes investigated by gene expression profiling**	A search for differential gene expression in fibroblast of patients with c.-32-13T>G / 'null' genotypes (in 35/39 cases) has led to the identification of genes and their associated metabolic and cellular processes that could potentially serve as markers for disease severity.	The preliminary results have to be validated using larger sample size. -Functional studies can be added to consolidate the findings.

MAIN FINDINGS

Enzymatic Diagnosis of Pompe Disease: lessons from 28 years of experience

Since 2006, enzyme replacement therapy (ERT) with alglucosidase alfa (Genzyme, Cambridge, MA, USA) has been available, patients with Pompe disease have experienced significant medical benefits. The most convincing effects of ERT are the reduction of cardiac hypertrophy, the survival for several years beyond natural life expectancy, and the reaching of important developmental milestones in classic infantile Pompe patients, in which the median age at death is 8.7 month without ERT [3, 4]. The short-term fatal prognosis in absence of treatment highly indicates there is more than ever a need for early and reliable diagnosis to reduce mortality and morbidity as well as to increase the chance of maximal therapeutic benefit. These considerations have led to the implementation of newborn screening (NBS) for Pompe disease in some countries. Before briefly discussing the pro's and con's of that approach I will first draw conclusions about the present day methodologies of routine enzymatic diagnosis as based on the studies described in Chapter 2.

In that study the diagnostic outcome data were analyzed that had been obtained over a period of 28 years by the diagnostic laboratory of the department of Clinical Genetics of the Erasmus University Medical Center, Rotterdam. In short, the assay in which cultured fibroblasts had been used as sample source and 4MUG as artificial substrate had provided the most reliable results. With this method a distinction could even be made between classic infantile and childhood or adult forms of Pompe disease on the basis of discriminating levels of residual GAA activity in almost all cases. The method has, however, also it's drawbacks in that the taking of a skin biopsy and the follow-up culturing of the fibroblasts is not a routine procedure at every institution, is time consuming, and culturing the cells under strictly standardized conditions is required. The second-best method turned out to be the one whereby mixed leukocytes were used as sample source. The drawing of blood is a routine procedure and so is the isolation of leukocytes at most sites. Based on our experience we recommend against the isolation of lymphocytes as a means to eliminate the activity of glucoamylase, which is a disturbing factor in the assay of GAA. Especially when the blood sample is a few days old the lymphocyte preparation will easily be contaminated with granulocytes producing the glucoamylase activity [5, 6]. With less effort and better result, the glucoamyslase activity can be inhibited by the inclusion of 3 uM acarbose in the assay with glycogen as substrate, and 8 uM acarbose with 4MUG as substrate.

The second experience with comparing the various assay procedures was

confirmation of an earlier observation that the use of glycogen as natural substrate provided a better distinction between affected and unaffected than the use of 4MUG as artificial substrate [7]. However, the latter procedure has also a draw back in that the so called *GAA2* allele which has a carrier frequency of 1:16 in Caucasian populations causes *GAA* pseudodeficiency that can mistakenly be interpreted as pathologic deficiency. Therefore, one of our conclusive advices in Chapter 2 is to measure the GAA activity in mixed leukocytes using glycogen as substrate and acarbose as inhibitor of glucoamylase, and to be aware that *GAA2* homozygosity and compound heterozygosity reduces the affinity of GAA for glycogen without causing a state of disease. Deficiency needs to be confirmed using a parallel assay with 4MUG substrate. We based our advice on the diagnostic outcome in cases that were included in our study with either 1 or 2 *GAA2* alleles. These samples fell in either the patient range or in a grey zone between affected and unaffected. For individuals with Asian ancestry counts that they may have *GAA* pseudodeficiency as carriers or homozygotes of 1 or 2 c.[1726A;2065A] pseudodeficiency alleles leading to moderately or markedly reduced activity in leukocytes and fibroblasts for both the natural substrate glycogen as well as the artificial 4MUG substrate. A limitation of all leukocyte assays is that they seem not sensitive / specific enough to provide information about the level of residual GAA activity and do not discriminate between classic infantile Pompe disease and less severe phenotypes with later onset. This distinction can be reached in fibroblasts using 4MUG as substrate.

Figure 3 in Chapter 2 clearly demonstrates that measuring GAA activity in cultured fibroblasts is the gold standard for diagnosing Pompe disease, since close to 1% of residual GAA activity is detectable using 4MUG substrate and distinctive clinical subtypes can be diagnosed on the basis of residual activity, while none of the other methods does. The flow diagram illustrates what we believe to be the best approach to diagnose Pompe disease based on our studies described in Chapter 2. Start with measuring the GAA activity in mixed leukocytes using both glycogen as well as 4MUG substrates. If the activity falls within the control range, Pompe disease is excluded and DNA analysis and/or fibroblast assays are not required (Figure 2A, top). However, if the GAA activity in mixed leukocytes using glycogen (+3 µmol/L acarbose) is below the control range but not clearly deficient using 4MUG (+8 µmol/L acarbose) as artificial substrate, the best approach is to perform DNA analysis to exclude *GAA2* pseudodeficiency and to search for other disease-associated *GAA* variants (Figure 2A, middle). In case of questionably low activity for both substrates perform DNA analysis to exclude heterozygosity or homozygosity for the Asian pseudodeficiency allele and to search for disease-associated variants as well (Figure 2A, bottom). In case none

of the former procedures leads to a definitive diagnosis and the outcome remains inconclusive, a skin biopsy can be taken to perform an additional GAA activity assay with 4MUG and any other assay either using DNA or RNA, or protein extracted from these cells or search for glycogen storage with histochemical procedures.

In addition, Figure 2B (text box) proposes on how to act in familial cases of Pompe disease. In this situation, a thorough clinical investigation including ancillary studies should be the first action to take. If the person is clinically affected the diagnosis can be confirmed by GAA enzyme activity assay or by DNA analysis if an index case is available. In case the clinical investigation is inconclusive a DNA analysis is indicated for identifying -the familial- *GAA* disease-associated variants, taking into consideration that the GAA enzyme activity assay may not provide an answer. In a situation like this it will be difficult to distinguish a pre-symptomatic patient with rather high GAA activity from an unaffected carrier. Diagnosis of Pompe disease based on *GAA* sequence analysis is not included in Chapter 2. However, in practice we use and recommend this procedure in most cases, either to confirm the diagnosis or to collect as much information as possible for counseling purposes and future (pre-natal) diagnosis, as well as for deciphering the GAA genotype-phenotype correlation and to learn about the contribution of modifying factors.

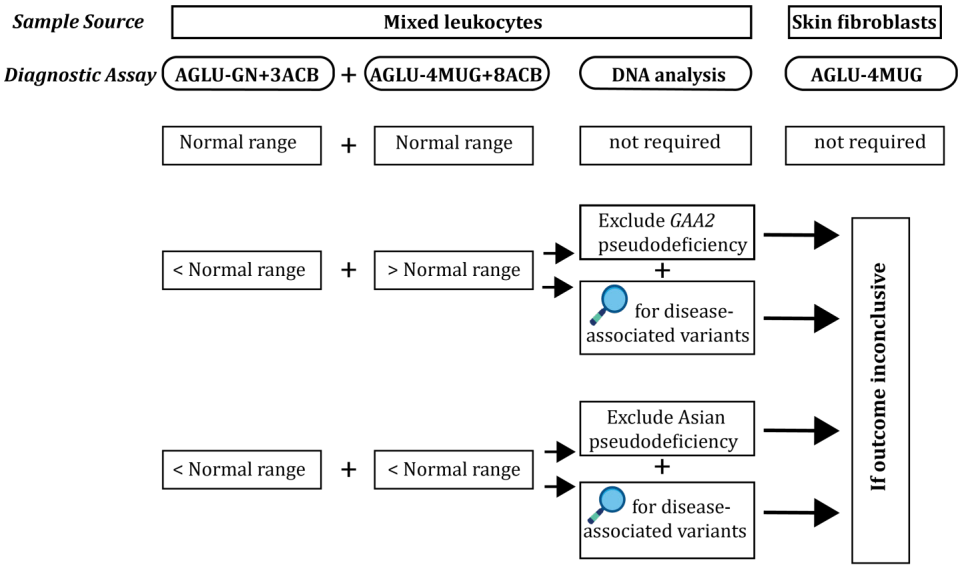


Figure 2A. How to diagnose Pompe disease? Flow diagram showing methods and materials for the enzymatic diagnosis of Pompe disease with special attention for GAA pseudodeficiency *(Niño et al 2020)*.

Familial cases of any ethnicity

1. Start with thorough clinical investigation including ancillary studies.

2. If the person is clinically affected the diagnosis can be confirmed by GAA enzyme assay or by DNA analysis.

3. In case the clinical investigations are inconclusive a DNA analysis is indicated for identifying -the familial- *GAA* disease-associated variants.

4. The GAA enzyme activity assay may not provide a definitive answer, since it will be difficult to distinguish a pre-symptomatic patient with rather high GAA activity from an unaffected carrier.

Figure 2B. How to diagnose Pompe disease? Recommendations for how to act in 'familial cases'.

Exceptional cases

Exceptional cases were found in this study, in which neither the GAA activity in leukocytes nor the GAA activity in fibroblast nor the genotype provided a decisive answer. There are two possible reasons. First, the patient might have a more complex genetic cause of Pompe disease. For example, SNP arrays have proven to be instrumental in demonstrating that some patients with Pompe disease have uniparental isodisomy (UPD), in which part of a chromosome is present in homozygous state and derived from only one parent [8]. This may lead to inclusion of extended analysis such as SNP arrays in obscure diagnostic cases. Second, we suggest that modulators, or modifying factors, having an impact on the clinical course of disease severity may exist. For instance, some studies have identified secondary modulators, as per example loss-of-function (LoF) mutations in zinc transporter 8 that protect against diabetes in obese individuals [9], another example is provided by mutations in globin genes that modify the severity of sickle cell disease by buffering primary mutations in β-globin genes [10]. Moreover, a study including around 600,000 genomes from healthy individuals identified 13 adults with no signs of clinical manifestations harboring mutations for eight severe Mendelian childhood diseases, suggesting that modifying factors can play a role in disease outcome [11]. The identification of modifier genes that can modulate disease progression is expected to have serious diagnostic and prognostic value. It can be envisioned that a panel of known modifiers could be included as a standard diagnostic tool to better predict disease progression. This might help to decide when treatment with ERT should be started.

Newborn screening (NBS)

Newborn screening has been implemented in several countries and US states to establish rapid diagnosis and early therapeutic intervention if needed [12-14]. For instance, a newborn screening program for Pompe disease was initiated in Taiwan

in 2005 and has led to early diagnosis of the disease. Indeed, interest in newborn screening for additional lysosomal storage disorders (LSDs) including Fabry Disease, Gaucher disease, and mucopolysaccharidosis type I (MPSI) continues to expand as new technologies and treatments become available. Pilot studies in LSDs provide insight in the effectiveness of, for instance, the digital microfluidic screening method as well as the refinement of the 'cutoff values' for detection of these LSDs [15, 16]. Technical advances on different methods based on the analysis of GAA activity have recently been reported, and new artificial substrates that are used for detection of GAA activity in dried blood spot (DBS), either by fluorometry, tandem mass spectrometry (MS/MS) [17], or microfluidics combined with fluorometry [18, 19]. These advances have improved throughput and performance of the assays, but they have not improved the intrinsic properties of the assays. It still remains important to validate the results of NBS in second tier biochemical analyses, and to exclude pseudodeficiencies using DNA analysis.

Drawbacks of NBS

The screening methods do serve their function in that affected newborns are identified in their first weeks of life, but they have their drawbacks that need to be considered too and are future challenges to be addressed in this field. For instance, NBS in Pompe disease is complicated by occurrence of false positive and false negative screening results. In **Chapter 2**, we presented a clear example of a false negative case of classic infantile Pompe disease in a patient who presented hypotonia since birth and hypertrophic cardiomyopathy (HCM) at 3.5 months of age. The misdiagnosis of this patient by DBS was due to the clinical picture of viral lung infection with bacterial superinfection. GAA pseudodeficiency due to homozygosity or heterozygosity for the c.[1726A;2065A] allele is a far more common complication of newborn screening in Asian populations with a high frequency of this allele [20, 21], Solutions include DNA analysis [22]. DBSs assays are also ideal for screening purposes, however, it is generally accepted that a second diagnostic procedure in samples other than bloodspots is needed to finally establish the diagnosis. Most problematic is the fact that the current screening test for Pompe disease not only detects cases of classic infantile Pompe disease, but also the childhood and adult forms, for which the age of onset varies considerably from infancy to adulthood. The unpredictable age of onset in those latter cases has a huge impact on emotional stress including psychological burden for the patients-in-waiting and their families. Moreover, patients might be lost to follow up [23, 24]. This has become an important point of discussion among patient representatives and treating physicians. It is evident that NBS provides an important advantage for the early identification of classic infantile Pompe disease. For later onset forms of Pompe

disease, it is more difficult to decide whether NBS provides an advantage. On one hand, patients would like to be diagnosed as early as possible to allow timely treatment and to prevent irreversible damage. On the other hand, there is an ethical concern in reporting late onset Pompe disease in newborns without being able to predict when symptoms will present, as these might occur only very late in life at 60-70 years of age. It will be important to continue these discussions among stakeholders and to closely evaluate the results of NBS for Pompe disease.

Extension of the Pompe mutation database by linking disease-associated variants to clinical severity

The rapidly expanding number of sequence variations that were discovered over the past 15 years has led to the start of the Pompe disease mutation database at www.pompecenter.nl to archive the continuously growing number of variants that were identified. Initially, the contribution by Erasmus MC was relatively large, but gradually more and more entries were retrieved from the literature. By the time that I started the study described in Chapter 3 the database (version May 2016) listed 558 different variants that were classified mainly based on the predicted and measured effect of the individual variants, but not linked to the clinical phenotypes of the patients in whom these variants were found. Thus, conclusions about genotype-phenotype correlations could not be drawn. Therefore, in order to make current knowledge of *GAA* variants useful to physicians and geneticists, it was decided to extend the database by additionally linking variants to clinical outcome of patients with associated genotypes.

In **Chapter 3** we extended the database with clinical information of reported phenotypes by performing a systematic review of published studies and report clinical phenotypes that are associated with *GAA* variants. We also added additional *in silico* predictions for effects on splicing and protein function and for cross reactive immunologic material (CRIM) status, minor allele frequencies (MAFs) and molecular analyses. We classified published patient information based mostly on the criteria stated in Güngör and Reuser [25]. We grossly divided the patients in three clinical subgroups: those with classic infantile Pompe disease, as originally described by Dr J.C. Pompe; patients with a disease onset during childhood; and those with an onset during adulthood. The currently often used terminology of infantile-onset Pompe disease (IOPD) and late-onset Pompe disease (LOPD) was intentionally avoided. Both terms are impractical for investigating genotype-phenotype correlations. In the literature, IOPD is variably used for onset of symptoms at infantile age (without very specific definition of age ranges) or as substitute for classic infantile Pompe disease characterized by onset of symptoms shortly after birth typically including hypertrophic cardiomyopathy

/ cardiomegaly with a mean survival age of 6-8 months, but is also used for onset of symptoms under the age of 1 year with cardiomyopathy. The problem with using the term LOPD is that it does not provide any information whatsoever about the current age of the patient and the age of onset which can be any age from birth to over 70 years. Differentiation between classic infantile, childhood and adult Pompe disease is important since the life expectancy of these sub-groups differs substantially and the projected outcome of enzyme replacement therapy as well.

In our study patients were classified with classic infantile Pompe disease if they presented symptoms at or under 12 months of age, and had evident signs of a hypertrophic cardiomyopathy. Patients were classified with childhood Pompe disease if the age of symptom onset was before 18 years of age and evident hypertrophic cardiomyopathy was absent. Patients were classified with adult Pompe disease if the first symptoms presented at the age of 18 years or later and evident hypertrophic cardiomyopathy was absent. In our study, 18 years was chosen as the cut-off age between childhood and adult Pompe disease. Others suggested around 16 years [25] and 12 years was chosen as cut-off age in previous Pompe Registry publications presenting an overview of the diagnostic methods [26, 27] and the genotype-phenotype correlation, respectively [28].

The first step taken to link *GAA* variants to phenotypes was to identify all patients with an evident form of classic infantile Pompe disease. From earlier studies and from our own review of enzymatic diagnostic procedures it is known that these patients do not have any GAA activity due to any combination of two so called *GAA* 'null alleles' [13, 29-32]. Thus, all *GAA* variants associated with the classic infantile phenotype were marked as "classic-infantile" variants. Subsequently, we analyzed the genotypes of affected children and adults. If a classic infantile variant was found in one allele the variant in the other allele had to be responsible for the milder phenotype (almost always associated with a certain level of residual GAA activity) [33]. This way variants were labeled "childhood" and "adult" or "childhood or adult" when associated with both phenotypes. Moreover, the pathogenicity of each variant was determined based on its MAF, molecular nature, *in vitro* functional studies, and *in silico* predictions. Another important feature added to the database was the CRIM status, which is a relevant criterium for treatment decision making in cases of classic infantile Pompe disease because CRIM-negative patients tend to develop higher antibody titers to the therapeutic enzyme than CRIM-positive patients [13].

As a result of our efforts 49% of the 422 disease-associated variants could be linked to patients with either classic infantile, childhood, or adult phenotypes. Thereby,

genetic counsellors and treating physicians are supported with predicting the patients' prognosis not only based on earlier identified genotypes, but also by combination of variant alleles. A total of 132 CRIM-negative and 216 CRIM-positive variants were identified based on predictions and immunoblot analyses. While disease-associated missense variants were found throughout the GAA protein, they were enriched up to 7-fold in the catalytic site.

The current updated version of the Pompe disease database, from now on named as the "Pompe disease *GAA* variant database" under the link (www.pompevariantdatabase.nl) offers access to a collection of information and data from studies developed in research laboratories (*in vitro* and *in vivo* functional studies), from bioinformatics (*in silico* analysis), and from scientific publications thereby providing meaningful connections between clinical, biochemical, and molecular genetic data. It will promote scientific research and deeper insight in several aspects of Pompe disease including patients' and families' support, diagnosis, newborn screening, genetic counselling, treatment, clinical trials planning, all together leading to improved clinical care and hopefully new modes of treatment.

It has become clear that improved treatment options are desirable for Pompe disease. In particular, ERT with myozyme has a variable clinical response, with both good, moderate, and no responders. In addition, it has become clear that the CNS shows progressive white matter abnormalities in long term surviving classic infantile Pompe patients, and that ERT is not able to cross the blood-brain barrier. Novel treatment options are being developed in several laboratories including our own, and the database will be instrumental in defining patient populations that may be amenable to specific treatment options. Future work is required for a further update beyond 2016, and to increase the number of patients in order to obtain a higher (>49%) percentage of variants that can be classified according to clinical severity.

The extended Pompe mutation database reveals distinct phenotypic spectra of Pompe disease-associated GAA variants

Understanding the clinical heterogeneity in Pompe disease and the variable response to ERT is a challenge on its own. Knowledge of the genotype-phenotype correlation and confounding factors is essential and heavily leans on high quality data coverage by the Pompe disease *GAA* variant database. The availability of information is necessary for answering and understanding several questions that still remain such as the broad phenotypical variation within *GAA* genotypes, common disease-associated variants in relation to geographic origin and founder effect, and as mentioned the different

response to ERT seen in patients with the same genotypes. Also, understanding and analyzing patient-specific variants and phenotypes could potentially improve patient selection for clinical trials, as well as lead to personalized therapies.

A first investigation of the genotype-phenotype correlation based on data in the Pompe disease database clearly supports the notion that modifying factors co-determine the clinical course. As early shown (Kroos et al 2007 and several other publications) [34-37] patients with a c.-32-13T>G / 'null' *GAA* genotype demonstrate a very broad clinical spectrum with regard to age of onset, severity and rate of disease progression. Our analyses do not only confirm this finding, but reveal clinical variation also within other *GAA* genotypes. In case of c.-32-13T>G / 'null' the spectrum encompasses onset from the first year of life till around 70 years. Within families and within other genotypes the variation is considerably less but cannot be ignored. Findings like these have led to the search for genes in studies such as described in Chapters 5 and 6, but also to the development of potentially therapeutic interventions based on the use of antisense oligonucleotides (AONs) to reduce the percentage of aberrantly spliced *GAA* pre-mRNA species and thereby higher the GAA activity to a non-pathogenic level [38-40]. The approach is of great relevance since c.-32-13T>G is the most common disease-causing variant in affected children and adults (86% of cases).

Shortly before completion of this study, my colleagues at Erasmus MC reported the discovery of a *GAA* silent variant in the c.-32-13T>G allele influencing the degree of aberrant splicing. As far as c.-32-13T>G / 'null' compound heterozygotes are concerned, this synonymous variant (c.510C>T located in exon 2 and uniquely co-occurring with c.-32-13T>G) was found in 27% of cases with childhood onset but in none with adult onset. Compared to the C at position c.510, the T appeared to increase the number of aberrant splicing events leading to a lower level of residual GAA activity, which could explain the earlier onset of symptoms in these cases. Similar, as yet to be identified, *GAA* sequence variants co-segregating with c.-32-13T>G might explain the early onset in the remaining 63% of childhood cases, but genetic modifiers located in cis or trans and outside the *GAA* gene must play a role too. Also the fact that -"Homozygous IVS1 (c.-32-13T>G) patients who carry c.510C>T, either at heterozygous or at homozygous state, have the prognosis to develop symptoms at any age, while homozygous IVS1 patients without c.510C>T may remain asymptomatic or may develop symptoms at any age"- points to the existence of genetic modifiers [41]. In the future, information on the modifier genes would present a major improvement. Furthermore, the database is meaningful source for diagnostics, newborn screening implementation, trial planning that can be used as platform tool to supporting patients and treating clinicians.

The ACE I/D polymorphism does not explain heterogeneity of natural course and response to enzyme replacement therapy in Pompe disease.

Chapter 5 is devoted to the potential contribution of the *ACE* I/D polymorphism to the clinical diversity in -32-13T>G / 'null' genotypes. ACE stands for angiotensin converting enzyme, which is produced by lung endothelial cells and catalyzes among others the conversion of angiotensin I to angiotensin II instrumental in the up-regulation of somatic functions such as water and salt retention via vasoconstriction. The Insertion / deletion (I/D) polymorphism in the *ACE* gene has been suggested to modify the age of onset and the response to ERT in Pompe disease [35-37, 42]. In Chapter 5 we have investigated the effect of the *ACE* I/D polymorphism in a larger cohort of children and adults (131) than earlier reported. One hundred and twelve of these patient had received ERT for several years and the clinical course of each individual has been accurately monitored over the 5 years' study duration. At first visit, we did not find significant differences in clinical outcome measures of patients with the *ACE* I/I , I/D or D/D genotypes. Also the response to ERT after a 5 years period did not reveal significant differences in any outcome measure between the ACE groups using linear mixed model analyses. The same analysis also failed to explain intra familial differences in age of onset and/or clinical course. Our conservative conclusion from this study was that the broad clinical spectrum of patients with -32-13C>T / 'null' *GAA* genotypes was not explained by genetic diversity of the *ACE* locus. Its contribution could only be minor and the search for alternative modifiers is justified. Consequently, it was decided to embark on a further search for transcriptional differences contributing to the clinical outcome that can be used as disease-marker as described in Chapter 6.

Identification of genes related to clinical heterogeneity in Pompe disease by gene expression profiling analysis.

Variable age of onset, clinical severity, and treatment efficacy among patients with *GAA* c.-32–13T>G / 'null' genotypes and lack of clear genotype-phenotype correlations also within and between families suggests that genetic modifiers exist that affect the phenotypic expression of Pompe disease [33, 36, 43]. Importantly, the identification of transcriptional differences contributing to the clinical outcome might lead to biomarker development, improvement of the therapeutic intervention, and prediction of disease phenotype. In **Chapter 6,** we applied RNA sequencing (RNA-seq) to examine gene expression profiles in cultured skin fibroblasts from two Pompe patients cohorts with similar *GAA* genotypes (c.-32-13T>G/ 'null', in 35/39 cases) but different ages of onset. In the two studies that we conducted, we compared expression differences in patients with those in age-matched controls. The two studies were evaluated separately and the results of each comparison were subsequently analyzed in categories to

identify expression patterns characteristic for Pompe disease phenotypes and those distinguishing between Pompe patients and controls. By doing this we specifically aimed to identify factors favoring either the adult or the childhood phenotype and/or factors that distinguish it from other LSDs and myopathies, which can be used as potential disease-marker in Pompe disease.

Expression differences between Pompe patients relative to controls

Analysis of the multidimensional scaling plot (MDS) revealed a separation of 23 (12 childhood onset, 11 adult onset) Pompe and 21 (11 childhood, 10 adult) control samples in the gene expression space indicating that Pompe disease induces gene expression differences. Additionally, the clear separation of adult onset and childhood onset patients indicates a distinctive transcriptional signature between these two Pompe phenotypes. In one study (study 1B in Chapter 6) six control samples emerged as outliers when compared to the other control samples. These samples were derived from foreskin in contrast to the other samples that were derived from skin biopsies. Apparently but unexpectedly the two cell types have a distinctive transcriptional signature.

Across the two studies that we performed, our analyses showed a number of common gene sets that were altered in disease, although the exact identity and pattern of gene expression varied between the studies. Part of the explanation for this variation could relate to inappropriate control selection criteria, to sample size, and possibly to batch effects, even though it is unlikely since two patient samples were included in study 1B, which included the controls, but they cluster instead of with the patients. Despite some variation in outcome data, it became clear from these two studies that common sets of differentially expressed genes (DEGs) are altered in disease compared to health and in affected adults compared with affected children.

Differentially expressed genes in common between studies

The common expression differences between the two studies were examined at the level of individual genes. To accomplish this, probe sets were identified in each dataset which were significantly different between adult and childhood onset Pompe disease (corrected for controls), and between patients with Pompe disease and controls at a false discovery rate (FDR) of <0.05 and Log fold change of 1.5. The overlap between studies was assessed using Venn diagram comparison. Between the two studies in adult Pompe compared to childhood Pompe, 3 DEGs were altered, and between Pompe compared to controls, 24 genes showed significant expression differences.

The 3 gene sets differentiating between adult and childhood onset Pompe disease are linked to known metabolic and cellular processes including bone and muscle development, cellular calcium / phosphate homeostasis, and cellular biogenic amine metabolic process. Among these DEGs, one gene was significantly downregulated and two genes were significantly upregulated in adult Pompe disease relative to childhood onset Pompe disease. Notably, the nik related kinase (*NRK*) gene showed over 14-fold decreased expression in adult Pompe relative to childhood Pompe. Its encoded protein is tentatively involved in the induction of actin polymerization through phosphorylation of cofilin-1 in late embryogenesis and plays a rol in actin cytoskeleton re-organization. Among its related pathways are NAD metabolism and TNF signaling. Gene Ontology (GO) annotations related to this gene include transferase activity, transferring phosphorus-containing groups and protein tyrosine kinase activity.

Genes coding for a secreted, homodimeric glycoprotein: *STC1*, and a mitochondrial enzyme *MAOA* exhibited over two-fold increased expression in adult Pompe compared to the childhood Pompe patients. The protein encode by *STC1* may play a role in the regulation of renal and intestinal calcium and phosphate transport, cell metabolism, or cellular calcium/phosphate homeostasis. *MAOA* is bound to the outer membrane of mitochondria in most cell types including skeletal muscle. It catalyzes the oxidative deamination of biogenic and xenobiotic amines and has important functions in the metabolism of neuroactive and vasoactive amines in the CNS and peripheral tissues. Among its related pathways is Cytokine Signaling in the immune system, which is in agreement with our results of GO terms by Reactome pathway analysis in our study 1. Indeed, genes with higher expression levels in adult Pompe compare to child Pompe tend to be expressed in pathways and biological processes that play a role in immune system functions as shown by GO analysis.

The identified genes are involved in a broad range of biological processes crucial for mitochondrial and muscle metabolism including oxidative phosphorylation, TCA cycle, fatty acid metabolism, amino acid degradation, cell survival, and calcium homeostasis, and relevant processes of the immune system. Since mitochondria play a central role in energy metabolism, while cytokines serve to mediate and regulate immune and inflammatory responses, the DEGs identified in our study could be partially responsible for mitochondrial dysfunction and inflammatory condition, which have been described in Pompe disease. Therefore, these genes could be attractive candidates that could influence the clinical course of Pompe disease.

The expression of this set of 3 genes might be associated with disease clinical status

and progression, and may therefore be useful as a putative signature in predicting symptom onset, and prognosis, and instrumental for genetic counseling in relation to neonatal screening programs for Pompe disease and the timing of treatment initiation. In parallel, it may provide a new insight of putative modifying factors, and consequently provide a better understanding of molecular pathogenesis of Pompe disease.

A retrospective study has evaluated gene expression changes in muscle from classic infantile Pompe patients and identified genes involved in muscle-specific pathways and genes involved in inflammation and apoptosis pathways and networks [44]. The pathologic process in muscle will have induced these changes, but interestingly, we identified pathways and networks playing a role in muscle metabolism and inflammatory responses as well using fibroblasts as experimental cell type. The similarity of findings is intriguing and deserves further attention. Nevertheless, the gene expression found in our studies in fibroblast presents good read out and reflects the disease. Important questions to be addressed remain, for instance, *are these candidate genes the cause or the consequence of disease progression, and are gene-specific patterns of expression variation per individual?*

Unfortunately, at present, there are not yet answers to these questions that could help with decision making about disease management and therapeutic intervention. However, each little step forward into the direction of biomarker development or cost-effective method to predict and diagnose the onset and progression of the disease, may ultimately be useful for therapeutic monitoring or intervention. The genes described in our study that were associated with the course of Pompe disease can potentially serve this purpose. However, the variability in expression among individuals (controls and patients) will make these genes unsuitable for predicting the clinical course of individual patients. Different approaches migt be needed to identify additional modifying factors. For example, recent work from our laboratory using allele-specific PCR identified the cis-acting silent *GAA* variant c.510C>T as a strong modifier of symptom onset [41]. This modifier was statistically highly significant, but explains only part of the phenotypic variation seen. Therefore, we cannot exclude that other not yet identified modifiers exist.

PERSPECTIVES AND CHALLENGES TO OVERCOME

Diagnosis

The studies described in this thesis incorporate several relevant aspects of Pompe disease that need be addressed for making steps forward in the understanding of Pompe disease and the development of therapeutic interventions. While most of our studies were already performed on relatively large numbers of cases and patient derived biological samples there remains the desire to further increase the sample size and to minimize 'local effects' by preferential inclusion of cases managed by our own Center for Lysosomal and Metabolic Diseases. One way to achieve this goal is to enhance international collaboration of clinical and fundamental research centers and vividly interact with Pompe patient associations and establish data repositories. Such actions are prone to validate the outcome of our studies and will benefit the accuracy and efficacy of future projects.

The introduction of new and the updating of existing laboratory procedures including ancillary studies, NBS, prenatal screening, enzyme analytical procedures and NGS technologies have facilitated the diagnosis of Pompe disease. However, deep knowledge, intelligent application of available procedures, and correct interpretation of outcome data remains the key to optimal disease management. In some cases it is difficult to decide which method to use for which purpose. As illustrated in Chapter 2 on the enzymatic diagnosis of Pompe disease, the outcome depends on sample source, substrates used, and awareness of pseudo-deficiency alleles. If DNA sequence analysis is used as diagnostic procedure other pitfalls arise in that *GAA* sequence variants may be missed or their effect be misinterpreted with currently available technologies. Despite the enormous advances in diagnostic procedures in recent years, but actually because of these advances some real challenges are faced. A salient example is the NBS assay, which is based on GAA activity measurement in dried blood spots. The assay is hampered by a relatively large percentage of false positives, troubled by the occurrence of GAA pseudodeficiency in Asian populations and does not discriminate between classic infantile Pompe disease and cases of childhood and/or adult onset disease.

The latter causes ethical dilemmas and devastating uncertainty in affected families. Notably, the diagnosis 'infantile onset Pompe disease' (IOPD) is a true ordeal that ignites a reaction of immediate start of ERT whereas the long-term clinical outcome of ERT remains uncertain with present day treatment protocol and clinical outcome. The verdict "late-onset Pompe disease" (LOPD) is accepted as the better outcome but

anxiousness about the first moment of symptom presentation lays a huge burden on the families' psychological wellbeing [23, 45]. On the one hand one may conclude that the current implementation of NBS in countries like Taiwan, Austria, Italy, Hungary, Japan, and several states in the United States [24] is to the benefit of patients, on the other hand it causes unwanted anxiety in absence of totally curative therapeutic modes of intervention. Discussion among all stakeholders involved is required to evaluate and possibly adapt the practice of NBS for Pompe disease.

Other than but certainly next to the introduction of NBS, there is an urgent need for international standardization of diagnostic protocols or at least the introduction of generally accepted diagnostic protocols in basic terms of when to apply what method. For instance, should one still measure GAA activity in isolated lymphocytes or use mixed leukocytes instead with either glycogen or 4MUG as substrate and acarbose as inhibitor of glucoamylase and in what concentration**?** Or alternatively start with DNA sequence analysis in all instances and rely on the features of the variants either predicted by their nature or by *in silico* analysis? Decision about which method to use should be made in expert diagnostic laboratories. This is crucial, because if the wrong test is chosen, particularly in cases of pseudodeficiency, or if outcome data are incorrectly interpreted, it may lead to misdiagnosis, for instance the unnecessary prescription of ERT to carriers.

With regard to newborn screening technical advances have been made [19, 24, 46] and contributed to improved diagnostic methods for distinguishing between Pompe disease and GAA pseudodeficiency in Asian populations [47]. With regard to *GAA* sequence analysis and the interpretation of *GAA* variants and their effects, the introduction of methods for functional analysis and methods to assay the efficacy of splicing events has substantially helped with diagnostic decision making [30, 48]. Most of all we believe that the systematic collection and analysis of all available information on *GAA* variants including both their functional as well as their clinical effects, as reported in Chapter 3 and 4, lends a tremendous support to diagnostic practice.

Future developments, such as the introduction of NGS facilities, will further contribute to easier, quicker and broader means of Pompe disease diagnosis including the detection of genetically determined modifying factors as well as the identification of variants affecting the promotor, untranslated regions (3' UTR and 5' UTR) and splice sites.

Genetic counseling and prognosis

Prognosis and genetic counseling should be intimately related to the diagnostic tract and 'loss to follow-up' considered inacceptable. Currently, the diagnosis and prognosis of Pompe disease relies predominately on clinical observation, GAA activity indices and/or common *GAA* genotypes. The same parameters are used as guidelines for the start of enzyme replacement therapy while, unfortunately, the very long-term outcome of the currently available enzyme replacement therapy is still under investigation. The diagnosis 'classic infantile Pompe disease' at birth is at present interpreted as an indication to start enzyme replacement therapy as soon as possible while it has become gradually known that grey matter changes and residual muscle weakness occur despite of therapy. This warrants the development of alternative treatment options that can target both muscle and the CNS. The diagnosis 'childhood' or 'adult' Pompe disease at birth is problematic too because none of the currently available assay procedures is able to predict with sufficient accuracy when the first symptoms will manifest and what the life expectancy of the patient will be given the broad clinical diversity within this group of patients. The first to be done is to develop methods and institute guidelines to discriminate at least between childhood and adult forms of Pompe disease and to abandon the terminology 'late-onset Pompe disease' (LOPD), as well as establishing an uniform terminology of Pompe disease classification described in the scientific literature. Advanced insights in the role of *GAA* variants to phenotypes, and disease modifiers could assist the counseling process. Attempts to identify these correlations and factors were described in Chapters 3-6.

What to expect from current developments?

With respect to diagnosis and the identification of clinical sub-types we believe that the worldwide collection and sharing of clinical, enzymatic, and molecular data is the most powerful tool for becoming optimally informed about the incidence of Pompe disease, the clinical presentation and the genotype-phenotype correlation; all supportive to optimal management. Twenty years ago, the Pompe disease mutation database at www.pompecenter.nl was set up to achieve this goal and will serve it even better in its present format (www.pompevariantdatabase.nl) including a variant severity rating not only based on molecular but also on clinical parameters (Chapters 3 and 4). Therefore, updating the database periodically will be established. In fact, our research group is currently working on a new update, which includes the addition of new scientific publications and variants previously not reported. Future updates should consider the addition of more extensive functional studies. It may be possible to create an algorithm that provides an estimation about the possible clinical phenotype and disease progression of a particular patient based on the *GAA* genotype

and clinical symptoms. This prediction of prognosis might help physicians to making timely treatment-decisions. However, much future research on larger patient cohorts is required to establish such algorithm. The establishment of the European Pompe Consortium is expected to be instrumental in these efforts as it will provide a much larger patient cohort that will be useful to validate putative modifying factors.

With respect to disease modifiers, it is evident that, on top of *GAA* genotypes, disease modifying factors can steer into the direction of either childhood or adult onset. In addition, the observed activation or depression of certain functional networks might lead to the identification of useful genes involved in predicting the disease course (Chapter 6). While exploring this topic, the key issue is determining whether the differences in gene expression profiles between phenotypes are the cause or the consequence of the phenotypic diversity. Our findings reveal that the c.-32-13T>G variant stands out for its exceptionally broad phenotypic variation. The current challenge is to identify the factor(s) involved in this variation, which may result in the design of a personalized treatment strategy. Additional plans include evaluating the association of genetic variants and gene expression levels by performing eQTL (expression quantitative trait loci) analysis. This will allow identification of the relationship between genomic variants and regulation of gene expression. Consequently, establish if the clinical variation with respect to disease onset and progression might be associated to individual variants per patient. Our recent work demonstrates that it is possible to predict an early presentation of symptoms based on the presence of the c.510C>T *GAA* modifier variant. This modifier was identified in a cohort of ~130 patients, demonstrating the feasibility of identifying disease modifiers even in relatively small cohorts [41]. However, collaborative efforts are needed to bring together large cohorts, which will facilate the finding and validation of also other moderate risk factors.

The glucose tetrasaccharide (Glc4) in urine is a well-known and useful biomarker in Pompe disease, but is not specific for the disease since increased levels are also found in other GSDs [49]. Myostatin and insulin-like growth factor 1 (IGF-1) have been proposed as potential therapeutic biomarkers in Pompe disease as well but are not regularly reported [50]. The insertion/deletion (I/D) polymorphism of the angiotensin I-converting enzyme (*ACE*), was reported as potential disease modifier and affect the response to ERT, but this was not supported by our own findings in a larger number of cases [36, 37, 42]. A demand for informative biomarkers in Pompe disease remains.

The studies described in this thesis do not provide a direct answer on how to **approach therapy,** but signifies that therapy starts with proper and rapid diagnosis at the deepest level of under-standing. In the era of personalized medicine, the aim

will be to design the appropriate and most effective treatment based on molecular, enzymatic, and clinical findings. Affected children and adults are expected to benefit from new generation enzyme replacement therapies employing improved enzyme formulas, higher doses and possibly co-administration of chaperones. Patients with classic infantile Pompe disease might benefit the same from novel modes of enzyme replacement therapy, but more so from the development of gene therapy protocols that address the white matter changes that have recently been reported. Therapies based on gene manipulation, for instance the employment of AONs for correcting splicing defects, the application of *ex vivo* or *in vivo* gene therapy compensating the defective *GAA* gene, or interference with genetic modifiers, are among the future options.

May the studies described in this thesis add their share to the very complex challenges that lay ahead.

REFERENCES

1. Hers, H.G., *alpha-Glucosidase deficiency in generalized glycogen storage disease (Pompe's disease).* Biochem. J., 1963. **86**: p. 11-16.
2. Fratantoni, J.C., C.W. Hall, and E.F. Neufeld, *The defect in Hurler and Hunter syndromes. II. Deficiency of specific factors involved in mucopolysaccharide degradation.* Proc Natl Acad Sci U S A, 1969. **64**(1): p. 360-6.
3. Amalfitano, A., et al., *Recombinant human acid alpha-glucosidase enzyme therapy for infantile glycogen storage disease type II: results of a phase I/II clinical trial.* Genet Med, 2001. **3**(2): p. 132-8.
4. Kishnani, P.S., et al., *A retrospective, multinational, multicenter study on the natural history of infantile-onset Pompe disease.* J Pediatr, 2006. **148**(5): p. 671-676.
5. Zhang, H., et al., *Comparison of maltose and acarbose as inhibitors of maltase-glucoamylase activity in assaying acid alpha-glucosidase activity in dried blood spots for the diagnosis of infantile Pompe disease.* Genet Med, 2006. **8**(5): p. 302-6.
6. van Diggelen, O.P., et al., *Enzyme analysis for Pompe disease in leukocytes; superior results with natural substrate compared with artificial substrates.* J Inherit Metab Dis, 2009. **32**(3): p. 416-23.
7. Okumiya, T., et al., *A new diagnostic assay for glycogen storage disease type II in mixed leukocytes.* Mol Genet Metab, 2006. **88**(1): p. 22-8.
8. Labrijn-Marks, I., et al., *Segmental and total uniparental isodisomy (UPiD) as a disease mechanism in autosomal recessive lysosomal disorders: evidence from SNP arrays.* Eur J Hum Genet, 2019. **27**(6): p. 919-927.
9. Flannick, J., et al., *Loss-of-function mutations in SLC30A8 protect against type 2 diabetes.* Nat Genet, 2014. **46**(4): p. 357-63.
10. Galarneau, G., et al., *Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation.* Nat Genet, 2010. **42**(12): p. 1049-51.
11. Chen, R., et al., *Analysis of 589,306 genomes identifies individuals resilient to severe Mendelian childhood diseases.* Nat Biotechnol, 2016. **34**(5): p. 531-8.
12. Reuser, et al., *Enzymatic and molecular strategies to diagnose Pompe disease.* Expert Opin Med Diagn, 2010. **4**(1): p. 79-89.
13. van Gelder, C.M., et al., *Enzyme therapy and immune response in relation to CRIM status: the Dutch experience in classic infantile Pompe disease.* J Inherit Metab Dis, 2015. **38**(2): p. 305-14.
14. Kishnani, P.S., et al., *Recombinant human acid [alpha]-glucosidase: major clinical benefits in infantile-onset Pompe disease.* Neurology, 2007. **68**(2): p. 99-109.
15. Chiang, S.C., et al., *Algorithm for Pompe disease newborn screening: Results from the Taiwan screening program.* Mol Genet Metab, 2012.
16. Hopkins, P.V., et al., *Lysosomal storage disorder screening implementation: findings from the first six months of full population pilot testing in Missouri.* J Pediatr, 2015. **166**(1): p. 172-7.
17. Gelb, M.H., C.R. Scott, and F. Turecek, *Newborn screening for lysosomal storage diseases.* Clin Chem,

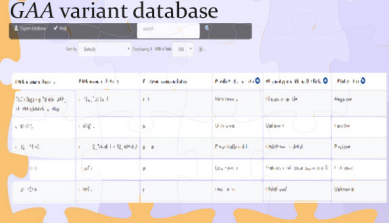
2015. **61**(2): p. 335-46.

18. Sista, R.S., et al., *Digital microfluidic platform for multiplexing enzyme assays: implications for lysosomal storage disease screening in newborns.* Clin Chem, 2011. **57**(10): p. 1444-51.
19. Sista, R.S., et al., *Multiplex newborn screening for Pompe, Fabry, Hunter, Gaucher, and Hurler diseases using a digital microfluidic platform.* Clin Chim Acta, 2013. **424**: p. 12-8.
20. Tajima, Y., et al., *Structural and biochemical studies on Pompe disease and a "pseudodeficiency of acid alpha-glucosidase".* J Hum Genet, 2007. **52**: p. 898-906.
21. Labrousse, P., et al., *Genetic heterozygosity and pseudodeficiency in the Pompe disease newborn screening pilot program.* Mol Genet Metab, 2009. **99**(4): p. 379-83.
22. Liao, H.C., et al., *Mass Spectrometry but Not Fluorimetry Distinguishes Affected and Pseudodeficiency Patients in Newborn Screening for Pompe Disease.* Clin Chem, 2017. **63**(7): p. 1271-1277.
23. van El, C.G., et al., *Newborn screening for pompe disease? a qualitative study exploring professional views.* BMC Pediatr, 2014. **14**(1): p. 203.
24. Bodamer, O.A., et al., *Newborn Screening for Pompe Disease.* Pediatrics, 2017. **140**(Suppl 1): p. S4-S13.
25. Gungor, D. and A.J. Reuser, *How to describe the clinical spectrum in Pompe disease?* Am J Med Genet A, 2013. **161A**(2): p. 399-400.
26. Kishnani, P.S., et al., *Timing of diagnosis of patients with Pompe disease: data from the Pompe registry.* Am J Med Genet A, 2013. **161A**(10): p. 2431-43.
27. Kishnani, P.S., et al., *Methods of diagnosis of patients with Pompe disease: Data from the Pompe Registry.* Mol Genet Metab, 2014. **113**(1-2): p. 84-91.
28. Reuser, A.J.J., et al., *GAA Variants and Phenotypes Among 1079 Patients with Pompe Disease: Data from the Pompe Registry.* Hum Mutat, 2019 **40**: 2146-2164.
29. Kroos, M.A., et al., *Glycogen storage disease type II: frequency of three common mutant alleles and their associated clinical phenotypes studied in 121 patients.* J Med Genet, 1995. **32**(10): p. 836-7.
30. Kroos, M., et al., *Update of the Pompe disease mutation database with 107 sequence variants and a format for severity rating.* Hum Mutat, 2008. **29**(6): p. E13-26.
31. Kroos, M., et al., *Update of the pompe disease mutation database with 60 novel GAA sequence variants and additional studies on the functional effect of 34 previously reported variants.* Hum Mutat, 2012. **33**(8): p. 1161-5.
32. Bali, D.S., et al., *Predicting cross-reactive immunological material (CRIM) status in Pompe disease using GAA mutations: lessons learned from 10 years of clinical laboratory testing experience.* Am J Med Genet C Semin Med Genet, 2012. **160C**(1): p. 40-9.
33. Kroos, M., et al., *The genotype-phenotype correlation in Pompe disease.* Am J Med Genet C Semin Med Genet, 2012. **160C**(1): p. 59-68.
34. Kroos, M.A., et al., *Broad spectrum of Pompe disease in patients with the same c.-32-13T->G haplotype.* Neurology, 2007. **68**(2): p. 110-5.
35. de Filippi, P., et al., *The angiotensin-converting enzyme insertion/deletion polymorphism modifies the clinical outcome in patients with Pompe disease.* Genet Med, 2010. **12**(4): p. 206-11.

36. De Filippi, P., et al., *Genotype-phenotype correlation in Pompe disease, a step forward.* Orphanet J Rare Dis, 2014. **9**(1): p. 102.
37. Ravaglia, S., et al., *Can genes influencing muscle function affect the therapeutic response to enzyme replacement therapy (ERT) in late-onset type II glycogenosis?* Mol Genet Metab, 2012. **107**(1-2): p. 104-10.
38. Bergsma, A.J., et al., *From Cryptic Toward Canonical Pre-mRNA Splicing in Pompe Disease: a Pipeline for the Development of Antisense Oligonucleotides.* Mol Ther Nucleic Acids, 2016. **5**(9): p. e361.
39. Goina, E., et al., *Glycogen Reduction in Myotubes of Late-Onset Pompe Disease Patients Using Antisense Technology.* Mol Ther, 2017. **25**(9): p. 2117-2128.
40. Goina, E., et al., *Assessment of the functional impact on the pre-mRNA splicing process of 28 nucleotide variants associated with Pompe disease in GAA exon 2 and their recovery using antisense technology.* Hum Mutat, 2019 **40**: 2121-2130.
41. Bergsma, A.J., et al., *A genetic modifier of symptom onset in Pompe disease.* EBioMedicine, 2019. **43**: p. 553-561.
42. Baek, R.C., et al., *The influence of a polymorphism in the gene encoding angiotensin converting enzyme (ACE) on treatment outcomes in late-onset Pompe patients receiving alglucosidase alfa.* Mol Genet Metab Rep, 2016. **8**: p. 48-50.
43. Wens, S.C., et al., *Phenotypical variation within 22 families with Pompe disease.* Orphanet J Rare Dis, 2013. **8**(1): p. 182.
44. Palermo, A.T., et al., *Transcriptional response to GAA deficiency (Pompe disease) in infantile-onset patients.* Mol Genet Metab, 2012. **106**(3): p. 287-300.
45. Pruniski, B., E. Lisi, and N. Ali, *Newborn screening for Pompe disease: impact on families.* J Inherit Metab Dis, 2018. **41**(6): p. 1189-1203.
46. Gelb, M.H., et al., *Direct multiplex assay of enzymes in dried blood spots by tandem mass spectrometry for the newborn screening of lysosomal storage disorders.* J Inherit Metab Dis, 2006. **29**(2-3): p. 397-404.
47. Shigeto, S., et al., *Improved assay for differential diagnosis between Pompe disease and acid alpha-glucosidase pseudodeficiency on dried blood spots.* Mol Genet Metab, 2011. **103**(1): p. 12-7.
48. Bergsma, A.J., et al., *Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach.* Hum Mutat, 2015. **36**(1): p. 57-68.
49. Manwaring, V., et al., *Urine analysis of glucose tetrasaccharide by HPLC; a useful marker for the investigation of patients with Pompe and other glycogen storage diseases.* J Inherit Metab Dis, 2012. **35**(2): p. 311-6.
50. Chien, Y.H., et al., *Myostatin and insulin-like growth factor I: potential therapeutic biomarkers for pompe disease.* PLoS One, 2013. **8**(8): p. e71900.

Pompe disease

GAA variant database



APPENDIX

Summary

The first chapter of this thesis **(chapter 1)** provides a historic overview of the developments that have led to the naming Pompe disease or glycogen storage disease type II. It further describes acid α-glucosidase (GAA), its biosynthesis, structure, function, and its deficiency that causes lysosomal glycogen storage, and how lysosomal glycogen storage leads to muscle cell damage and clinical symptoms through a cascade of pathologic events. The chapters 2 to 6 describe the research that was done to advance insight in Pompe disease whereby attention is focused on diagnostic methods (**chapter 2**), *GAA* sequence variant collection and analysis 'The Pompe disease *GAA* variant database' (**chapters 3 and 4**), the proposed role of *ACE* as genetic modifier contributing to clinical variation and response to ERT **(chapter 5), and** genome-wide differential gene expression in Pompe disease with emphasis on differences between early- and late-onset phenotypes as potential disease-markers **(chapters 6). Chapter 7** is a general discussion. What follows here is a brief summary of the most essential content of each chapter.

Pompe disease or glycogen storage disease type II (GSDII) **(chapter 1)** is one of more than 10 different Glycogen Storage Diseases (GSDs) and also one of over 70 different lysosomal disorders (LSDs). It is an autosomal recessive disorder caused by acid α-glucosidase (GAA) deficiency. GAA enzyme normally degrades lysosomal glycogen stepwise to glucose. Pathogenic sequence variations in the *GAA* gene are responsible for partial or complete GAA deficiency causing glycogen to accumulate inside lysosomes, causing muscle pathology. Expanded lysosomes not only interfere with muscle contraction but also with other kinds of cellular processes including autophagy. In the final stage of pathology, the cells are damaged and progressive muscle degeneration occurs.

Pompe disease presents as a spectrum of phenotypes, ranging from a rapidly fatal phenotype in infants, usually succumbing in their first year of life due to cardiorespiratory failure, to more slowly progressive disease in older children and adults whereby skeletal and respiratory muscles are affected, but the heart is spared. The clinical heterogeneity with regard to age of onset observed among children and adults with the most common 'c.-32-13 T>G / 'null' genotype has led to the hypothesis that modifying factors play a role in disease severity and progression (addressed in **chapters 4-6**).

In the very early days (1932), Pompe's disease (glycogenosis type II) was known as a disease of infancy. Muscular variants were described much later, but still before acid α-glucosidase deficiency was discovered as the primary cause of lysosomal glycogen storage. By 1973, the full clinical spectrum of glycogenosis type II had come to light, and was first described by Dr. Andrew Engel and his colleagues. At that time, there was no clue to what factors were causing the clinical diversity. The difference between the classic-infantile phenotype as described by Dr Pompe, and the later-onset phenotypes (childhood and adult) is so striking that they could have been caused by mutations in different genes; for instance, a gene encoding GAA structure and a gene regulating the height of *GAA* expression. The latter hypothesis was abandoned since complementation analysis did not lead to correction of GAA deficiency. Later, it was generally accepted that a variety of allelic mutations in the acid α-glucosidase gene (*GAA*) must be causing the clinical heterogeneity, but this could not be proven at that time by lack of technology. However, genetic heterogeneity could be demonstrated indirectly by following the biosynthesis of GAA in cultured skin cells of clinical variants and the finding of informative differences. The full extent of genetic heterogeneity has come to light after the cloning of the *GAA* gene allowing the search for pathogenic sequence variations.

The rapidly expanding number of sequence variations that were discovered over the past 15 years has led to initiate the Pompe disease mutation database at www.pompecenter.nl, which had grown to 558 *GAA* variants in May 2016. These variants had obtained designations from 'very severe' to 'non-pathogenic' either based on their location and/or nature, their *in silico* predicted effect, or their effect on GAA synthesis and enzymatic function as studied *in vitro*. The accuracy of the predictions was never challenged, but not investigated either. Upgrading the Pompe disease mutation database by linking individual *GAA* variants to *GAA* genotypes and clinical phenotypes became one of the major study subjects resulting in the extended *GAA* variant database that replaced the previous database and can be found at www.pompevariantdatabase.nl (**chapters 3** and **4**).

Undoubtedly, one of the most exciting events in the history of Pompe disease has been the introduction of enzyme replacement therapy in 2006. The benefits and still remaining challenges are also addressed in this first chapter, but haven't been subject of this study.

Other background information provided in **Chapter 1** relates to the frequency of Pompe disease, the geographic distribution, the symptoms and the clinical spectrum.

The aims / questions of the studies performed in the context of this thesis were formulated at the end of **Chapter 1** and were as follows:

1. How to diagnose Pompe disease enzymatically?
2. Upgrading of the Pompe disease mutation database and facilitation of the analysis of the genotype-phenotype correlation.
3. Does the common I/D *ACE* polymorphism contribute to the heterogeneity of Pompe disease?
4. Can disease-markers be identified that can distuinguish severe from mild disease**?**. The results of these studies are discussed in broader context, overall conclusions are drawn, and future perspective are formulated in **Chapter 7**.

Chapter 2 is devoted to the enzymatic diagnosis of Pompe disease and reviews the outcome of a series of 2591 assays in 1709 diagnostic cases from various sources using different procedures over the past 28 years in the department of Clinical Genetics of the Erasmus MC University Medical Center in Rotterdam. The GAA activity was measured in white blood cells (leukocytes), dried blood spots, and skin fibroblast with the aim to establish the best assay for particular purposes. It came out that cultured skin fibroblasts provide by far the best results with regard to differentiating between classic infantile, childhood and adult forms of Pompe disease. Mixed leukocytes provide the fastest diagnostic method, whereby the reaction mixture needs to contain acarbose to inhibit glucoamylase. Glycogen as well as 4-methylumbelliferyl-α-D-glucopyranoside are suitable substrates, but glycogen discriminates best between affected and unaffected. The disadvantage of using glycogen is that carriers of *GAA2* pseudodeficiency may be scored as patients. In such cases, additional use of 4-methylumbelliferyl-α-D-glucopyranoside (4MUG) solves the problem. Dried bloodspots are the best sample source for screening purposes, but are in our hands not suitable for final diagnosis, which is in line with the general opinion that a second diagnostic test is required to confirm the diagnosis.

Chapter 3 and **4** describe how the Pompe disease mutation data base with the previous link www.pompecenter.nl and the new link www.pompevariantdatabase.nl was upgraded and logistically altered to provide more and better information on the functional effects of *GAA* sequence variants by linking each variant (one allele) to pathogenic genotypes (a combination of 2 alleles) and their associated clinical phenotypes. The direct link deepens insight in the genotype-phenotype correlation and facilitates diagnosis, prediction of disease severity and genetic counseling. The upgraded version of the database **(chapter 3)** provides quick access to all clinical phenotypes associated with one particular *GAA* genotype and easily delineates the

clinical diversity within genotypes **(chapter 4**). Similar types of analysis have been described in other disorders, but are mostly restricted to single patients per genotype. Our genotype-phenotype analysis illustrates that 49% of the 422 disease-associated *GAA* variants are linked to classic infantile, childhood and adult phenotypes. In addition, 38% of disease-associated variants listed in the database are shown to be or predicted to be Cross Reactive Immunological Material (CRIM) negative (i.e. without any expression of GAA protein), while the majority of missense variants is predictably CRIM positive (i.e. with expression of enzymatically inactive GAA protein) (**chapter 3**). The new features that are added to the Pompe disease mutation database provide an invaluable tool for diagnosis, prognosis of disease progression, and treatment decision-making in Pompe disease. **Chapter 4** is a first demonstration of how the upgraded version of the database can be used for research purposes. In this study we analyzed and compared the phenotypic variation among patients with the most common c-32-13T>G genotypes and other rather common *GAA* genotypes. The conclusion is reached that there must be genetic or environmental factors that co-determine the course of Pompe disease in case patients are not fully GAA deficient.

In **Chapter 5** the effect of one of the few modifying factors previously reported was explored. It concerns the *ACE* I/D polymorphism. Based on rather small patient numbers, patients with the ACE DD genotype compared to the ACE II genotype seemed to have more rapid disease progression and reduced response to enzyme replacement therapy. However, some of the reported findings are conflicting, and studies in a larger patient population were recommended. Therefore, we investigated the role of the *ACE* polymorphism in a relatively large cohort of 131 children and adults covering the full phenotypic spectrum of patients with the c.-32-13T>G / 'null' *GAA* genotype with regard to age of disease onset, disease severity and outcome when treated with ERT over a follow up period of 5 years. In our study we could not confirm earlier results: differences in *ACE* polymorphism could neither explain inter-familial nor intra-familial phenotypic diversity.

In **Chapter 6**, we performed a search for disease-markers in Pompe disease that could be used as read out of disease severity. The chosen method was that of measuring gene expression at the mRNA level in two groups of patients with similar *GAA* genotypes (c.-32-13T>G / 'null' in 35/39 cases); one with onset of first symptoms under the age of 16 years and one with onset of symptoms over the age of 35 years. These studies were hampered by inter-experimental variation in outcome, but three genes stood out as tentative disease-markers: *MAOA* (Monoamine Oxidase Type A), which catalyzes the oxidative deamination of biogenic and xenobiotic amines and has important

functions in the metabolism of neuroactive and vasoactive amines in the central nervous system and peripheral tissues; *STC1* (Stanniocalcin-1) that plays a role in the regulation of renal and intestinal calcium and phosphate transport, cell metabolism, or cellular calcium/phosphate homeostasis; and *NRK* (Nik Related Kinase) that has been associated with the induction of actin polymerization in late embryogenesis. In general, these genes have been implicated in various biologic processes including bone and muscle development, cellular calcium / phosphate homeostasis, cellular metabolism, and cellular biogenic amine metabolic process. In addition, a multitude of genes could be earmarked as potential disease-markers since they were differentially expressed in cultured skin fibroblasts from affected Pompe patients compared to unaffected controls.

Collectively, the findings described in this thesis are a step forward in the diagnosis and understanding of Pompe disease with emphasis on the genotype-phenotype correlation, as discussed in **Chapter 7**.

Samenvatting

Het eerste hoofdstuk van dit proefschrift (**Hoofdstuk 1**) geeft een historisch overzicht over de ziekte van Pompe, ook wel glycogeen stapelingsziekte type II genoemd. In dit hoofdstuk wordt het enzym zure alfa-glucosidase (GAA) besproken, alsmede de biosynthese, structuur, functie en de deficiëntie ervan die uiteindelijk leidt tot stapeling van glycogeen en de ziekte van Pompe. Ook wordt besproken hoe de stapeling van glycogeen in de lysosomen kan leiden tot beschadiging van spiercellen en de bijbehorende klinische symptomen.

De hoofdstukken 2 tot 6 beschrijven het verrichte onderzoek in dit proefschrift om meer inzicht te krijgen in de ziekte van Pompe. De focus van het proefschrift ligt met name op de diagnostiek van de ziekte van Pompe (**Hoofdstuk 2**), het in kaart brengen van *GAA* sequentie varianten en het uitbreiden van de Pompe database waarin alle bekende mutaties zijn opgenomen (**Hoofdstuk 3** en **4**). **Hoofdstuk 5** en **6** focussen op de rol van genen die wellicht een voorspellende factor zouden kunnen zijn voor het klinisch verloop van de ziekte van Pompe. Het laatste hoofdstuk (**Hoofdstuk 7**) is een algemene discussie. Een korte samenvatting van elk hoofdstuk is hieronder weergegeven.

De ziekte van Pompe (**Hoofdstuk 1**) is één van de bijna 70 verschillende lysosomale stapelingsziekten (LSDs), waaronder meer dan 10 verschillende glycogeen stapelingsziekten (GSDs) vallen. Het is een autosomaal recessieve ziekte die veroorzaakt wordt door een mutatie in het *GAA* gen dat zure alfa-glucosidase (GAA) produceert. Het GAA enzym breekt normaal gesproken glycogeen af tot glucose. Pathogene variaties in de sequentie van het *GAA* gen zorgen voor een gedeeltelijk of geheel tekort aan GAA en dit leidt tot stapeling van glycogeen in de lysosomen. Deze stapeling van glycogeen in de lysosomen zorgt ervoor dat de lysosomen opzwellen en daarbij het samentrekken van spieren bemoeilijkt. De glycogeen stapeling kan ook interfereren met andere processen die zich in de cel afspelen, zoals autofagie. In het laatste stadium van de ziekte van Pompe zijn spiercellen zo beschadigd dat er progressieve spierdegeneratie plaatsvindt.

De ziekte van Pompe heeft een breed fenotypisch spectrum. Het snel progressief en fatale fenotype in baby's leidt tot overlijden aan cardiorespiratoir falen. Bij het minder progressieve fenotype in oudere kinderen en volwassenen zijn voornamelijk de skelet- en ademhalingsspieren aangedaan, maar het hart niet. Grote klinische

heterogeniteit wordt gezien bij de genetische afwijking die het meest voorkomt bij kinderen en volwassenen: c.-32-13 T>G / 'null'. Dit heeft geleid tot de hypothese dat naast de genetische mutatie ook modificerende factoren een rol spelen in de mate van progressie en ernst van de ziekte (**Hoofdstuk 4-6**).

In 1932 werd de ziekte van Pompe voor het eerst beschreven door de Nederlandse patholoog J.C. Pompe en was alleen bekend als een ziekte die bij baby's voorkwam. Later werd de kinder- en volwassen variant van de ziekte beschreven waarbij alleen de spieren waren aangedaan. In 1973 werd het brede spectrum van de ziekte van Pompe beschreven door Dr. Andrew Engel en collega's, maar het was nog steeds onduidelijk wat de onderliggende oorzaak hiervan was. Opvallend genoeg wordt het grote verschil in klinische manifestatie tussen het klassiek infantiele fenotype en het niet-klassieke fenotype niet veroorzaakt door mutaties in verschillende genen. Later ontstond het idee dat verschillende mutaties in het *GAA* gen zouden kunnen leiden tot de klinische heterogeniteit van de ziekte van Pompe, maar dit kon in die tijd nog niet bewezen worden. Wel kon de genetische heterogeniteit indirect gedemonstreerd worden door de biosynthese van GAA in gekweekte huidcellen van patiënten met de verschillende varianten van de ziekte te bestuderen welke leidde tot duidelijke verschillen. Pas nadat het *GAA* gen gekloneerd kon worden kon er onderzoek gedaan worden naar de pathogene sequentie varianten en werd de grote genetische heterogeniteit duidelijk.

Het groeiende aantal sequentie varianten die werden ontdekt in de afgelopen 15 jaar leidde uiteindelijk tot het opzetten van de Pompe database (www.pompecenter.nl). In mei 2016 waren er al meer dan 558 genetische varianten van de ziekte van Pompe in opgenomen. De variaties werden geclassificeerd als 'niet-pathogeen' tot 'zwaar aangedaan' en was gebaseerd op de locatie waar in het gen de mutatie zit, een *in silico* voorspelling en/of het effect van de mutatie op GAA synthese en enzym activiteit welke bepaald zijn met *in vitro* testen. De nauwkeurigheid van de *in silico* voorspelling werd nooit in twijfel getrokken, maar werd verder ook nooit onderzocht. Dit proefschrift beschrijft de optimalisatie van de Pompe database en is gedaan door verbanden te leggen tussen de individuele *GAA* varianten, de GAA genotypes en klinische fenotypes. De resultaten hiervan leidden tot een nieuwe Pompe database. De nieuwe database kan hier gevonden worden: www.pompevariantdatabase.nl (**Hoofdstuk 3** en **4**).

De introductie van de enzym vervangingstherapie in 2006 is zonder twijfel de belangrijkste gebeurtenis in de geschiedenis van de ziekte van Pompe. De voordelen van deze enzym vervangingstherapie en uitdagingen die nog openstaan worden ook besproken in het eerste hoofdstuk. Verder wordt de frequentie van de ziekte van

Pompe besproken, de geografische verdeling, symptomen en het klinisch spectrum. De doelstellingen van de studies in dit proefschrift zijn:

1. Hoe kan de ziekte van Pompe enzymatisch het beste worden gediagnosticeerd?
2. Het leveren van een verbeterde Pompe database die kan bijdragen aan het analyseren van een eventuele genotype-fenotype correlatie
3. Draagt het *ACE* polymorfisme bij aan klinische heterogeniteit van de ziekte van Pompe?
4. Kunnen er genen geïdentificeerd worden met behulp van genoom-wijde mRNA expressie analyse die een rol spelen bij de ernst van het fenotype in kinderen en volwassenen met hetzelfde *GAA* genotype? Zouden deze genen gebruikt kunnen worden als voorspellers van het verloop van de ziekte?

Hoofdstuk 2 is geheel gewijd aan de enzymatische diagnose van de ziekte van Pompe en bespreekt de resultaten van 2591 enzym assay's die gedaan zijn met materiaal van 1709 personen in de afgelopen 28 jaar in het Erasmus Medisch Centrum in Rotterdam. De GAA enzym activiteit werd gemeten in witte bloedcellen (leukocyten), gedroogde bloed spots en huid fibroblasten. De gekweekte huid fibroblasten gaven de beste resultaten om te kunnen differentiëren tussen de klassiek infantiele en de niet-klassieke vorm van de ziekte van Pompe. De enzym assay met leukocyten is de snelste methode om de ziekte vast te stellen, waarbij wel acarbose nodig is om neutrale hydrolase glycoamylase te inhiberen, dit interfereert namelijk met de meting van het GAA enzym. Als substraat voor de enzym assay kan zowel glycogeen als 4-methylumbelliferyl-α-D-glucopyranoside gebruikt worden, waarbij glycogeen meer geschikt is om het verschil aan te tonen tussen aangedaan of niet-aangedaan. Het nadeel van het gebruik van glycogeen als substraat is dat dragers van het *GAA2* gen geclassificeerd worden als patiënten met de ziekte van Pompe terwijl ze de ziekte niet hebben. In dat geval moet 4-methylumbelliferyl-α-D-glucopyranoside gebruikt worden. Gedroogde bloedspots zijn de best toegankelijke bron van materiaal om patiënten te screenen, maar zijn niet geschikt om de finale diagnose mee te stellen en vereisen dan ook een tweede diagnostische test om de diagnose te bevestigen.

Hoofdstuk 3 en **4** beschrijven de overgang van de oude Pompe mutatie database naar de nieuwe Pompe mutatie database (www.pompevariantdatabase.nl). De nieuwe Pompe database bevat meer en gedetailleerdere informatie over de effecten van *GAA* sequentie varianten doordat daarin elke variant (op één allel) gelinkt is aan het bijbehorende pathogene genotype (combinatie van 2 allelen) en klinisch fenotype. De gelinkte data geven meer inzicht in de genotype-fenotype correlatie en faciliteren diagnose, voorspelling van het ziekteverloop en kunnen ondersteunend zijn bij

genetische counseling. De nieuwe database (**Hoofdstuk 3**) geeft snel toegang tot alle klinische fenotypes die geassocieerd zijn met een bepaald *GAA* genotype en kan zo klinische verscheidenheid tussen de genotypes gemakkelijk weergeven (**Hoofdstuk 4**).

De genotype-fenotype analyse uit de database laat zien dat 49% van de 422 geanalyseerde *GAA* varianten van de ziekte van Pompe gelinkt zijn aan klassiek infantiele, juveniele en adulte fenotypes. Ook zijn 38% van de pathogene varianten bewezen of voorspeld 'cross reactive immunologic material' (CRIM) negatief, terwijl de meerderheid van de missense varianten een voorspelde CRIM positieve status hebben (**Hoofdstuk 3**). De CRIM-status is belangrijk omdat dit een grote rol speelt bij de prognose van de ziekte en bij het opstellen van een behandelplan: CRIM negatieve patiënten hebben in het algemeen een minder goede prognose.

De uitbreiding van de Pompe database hebben ervoor gezorgd dat het een onmisbare database is geworden voor de diagnose, prognose van het verloop van de ziekte en keuze van behandeling van de ziekte van Pompe. In **hoofdstuk 4** wordt voor het eerst getoond hoe de nieuwe versie van de Pompe database gebruikt kan worden voor onderzoeksdoeleinden. We analyseerden en vergeleken de fenotypische variaties van patiënten met de meest voorkomende c-32-13C>T genotypes (en andere voorkomende *GAA* genotypes) en concludeerden dat er genetische en/of omgevingsfactoren een rol spelen bij het ziekteverloop van de ziekte van Pompe.

In **hoofdstuk 5** is het effect van het ACE I/D polymorfisme, een in de literatuur voorgestelde modificerende factor, onderzocht. In een kleine patiëntengroep werd gerapporteerd dat patiënten met het ACE DD genotype, in vergelijking met het ACE II genotype, een sneller ziekteverloop hebben en een verminderde respons op de enzym vervangingstherapie. Vanwege het kleine aantal patiënten in deze studie waren de resultaten niet eenduidig en nieuwe analyses met een grotere groep patiënten werden aanbevelen. Daarom werd hier de rol van het ACE polymorfisme in een cohort van 131 kinderen en volwassenen bestudeerd. Deze patiënten hadden allen het c.-32-13T>G / 'null' *GAA* genotype maar vertoonden een breed spectrum van de ziekte (start van de ziekte, ernst van de ziekte en respons op ERT behandeling na 5 jaar behandeling). In onze studie konden we eerdere resultaten niet bevestigen: verschillen in het ACE polymorfisme konden de inter-familiaire en de intra-familiaire fenotypische diversiteit niet verklaren.

In **hoofdstuk 6** hebben we gezocht naar genen waarvan de expressie voorspellend zou kunnen zijn voor het ziekte-verloop van de ziekte van Pompe. De expressie

van genen werd bestudeerd op mRNA niveau in twee groepen patiënten die allen hetzelfde *GAA* genotype (c.32-13T>G / 'null', in 35/39 casussen) hebben. De ene groep patiënten had de eerste symptomen voor het 16e levensjaar en de andere groep patiënten na het 35ste levensjaar. Deze studies hadden als nadeel dat er veel experimentele verschillen waren tussen experimenten, maar er waren drie genen die daar geen last van hadden en er daardoor uit sprongen in alle experimenten. Daarom zouden die drie genen voorspellend kunnen zijn voor het verloop van de ziekte van Pompe: *MAOA* (Monoamine Oxidase Type A), dat voor versnelling van oxidatieve deaminatie van biogene en xenobiotische amines zorgt en een belangrijke rol speelt in het metabolisme van neuroactieve en vasoactieve amines in het centraal zenuwstelsel en perifere weefsels; *STC1* (Stanniocalcin-1), dat een rol speelt in de regulatie van calcium en fosfaat in de nieren en de darmen, fosfaat transport, in cel metabolisme en in homeostase van calcium en fosfaat; en *NRK* (Nik Related Kinase) gen, dat geassocieerd wordt met het induceren van actine polymerisatie tijdens embryogenese. In het algemeen spelen deze genen een rol in verschillende biologische processen, inclusief spier- en botontwikkeling, calcium/fosfaat homeostase, metabolisme en biogene amine metabole processen in de cel. Vele andere genen leken aanvankelijk ook een voorspellende rol te kunnen hebben, omdat zij anders tot expressie kwamen in gekweekte huid fibroblasten van aangedane Pompe patienten vergeleken met niet aangedane gezonde personen. Echter deze genen kwamen niet significant anders tot expressie in een onafhankelijk cohort, waardoor hun voorspellende rol in twijfel werd gebracht.

De bevindingen in dit proefschrift zijn een stap voorwaarts in het diagnosticeren en leiden tot meer begrip van de ziekte van Pompe, waarbij de nadruk ligt op de genotype-fenotype correlatie (**Hoofdstuk 7)**.

Resumen

El primer capítulo de esta tesis (**capítulo 1**) proporciona una visión histórica de los desarrollos que le dieron el nombre a la enfermedad de Pompe o enfermedad de almacenamiento de glucógeno tipo II. Además, describe la proteína llamada α-glucosidasa ácida (GAA), su biosíntesis, estructura, función y deficiencia que causa almacenamiento de glucógeno lisosomal, y cómo el almacenamiento de glucógeno lisosomal conduce al daño de las células musculares y los síntomas clínicos a través de una cascada de eventos patológicos. Los capítulos 2 a 6 describen las investigaciones o estudios que se realizaron para avanzar en la comprensión de la enfermedad de Pompe, la cual centra su atención en los métodos de diagnóstico (**capítulo 2**), la recolección y el análisis de las variantes en la secuencia del gen *GAA* 'La base de datos de la variante *GAA* de la enfermedad de Pompe (**capítulos 3 y 4**), el sugerido rol de *ACE* como modificador genético que podría contribuir a la variación clínica y a la respuesta de terapia de remplazo enzimático (ERT, de sus siglas en inglés) (**capítulo 5**), y la expresión génica diferencial de todo el genoma en la enfermedad de Pompe con énfasis en las diferencias entre los fenotipos de inicio temprano y tardío como posibles marcadores de la enfermedad (**capítulos 6**). El capítulo 7 incluye la discusión general de esta tesis. A continuación, se describe un breve resumen del contenido más esencial de cada capítulo.

La enfermedad de Pompe o la enfermedad de almacenamiento de glucógeno tipo II (GSDII) (capítulo 1) es una de más de 10 diferentes enfermedades de almacenamiento de glucógeno (GSD) y también una de más de 70 diferentes trastornos lisosomales (LSD). La enfermedad de Pompe es un trastorno autosómico recesivo causado por la deficiencia de la proteína denominada α-glucosidasa ácida (GAA). La enzima GAA normalmente degrada el glucógeno lisosomal en glucosa. Las variantes patogénicas en la secuencia del gen *GAA* son responsables de la deficiencia parcial o completa de la enzima GAA que causa que el glucógeno se acumule dentro de los lisosomas, causando patología muscular. Los lisosomas expandidos no solo interfieren con la contracción muscular sino también con otros tipos de procesos celulares, incluida la autofagia. En la etapa final de la patología, las células se dañan y se produce una degeneración muscular progresiva.

La enfermedad de Pompe presenta un espectro de fenotipos, que van desde un fenotipo rápidamente mortal en recién nacidos/lactantes, que suelen fallecer en su

primer año de vida debido a una falla cardiorrespiratoria, hasta una enfermedad más lenta y progresiva en niños mayores y adultos, en donde los músculos esqueléticos y respiratorios se ven afectados, pero el corazón no. La heterogeneidad clínica con respecto a la edad de inicio observada entre niños y adultos con el genotipo más común 'c.-32-13 T> G / 'nulo' ha llevado a la hipótesis de que factores modificadores juegan un rol en la severidad y progresión de la enfermedad (abordado en los **capítulos 4-6**).

A comienzos de 1932, la enfermedad de Pompe (glucogenosis tipo II) se conocía como una enfermedad de la infancia. Las variantes musculares se describieron mucho más tarde, pero aún antes de que se descubriera la deficiencia de α-glucosidasa ácida como la causa principal del almacenamiento de glucógeno lisosomal. En 1973, el espectro clínico completo de la glucogenosis tipo II había salido a la luz, y fue descrito por primera vez por el Dr. Andrew Engel y sus colegas. En ese momento, no era claro cuales factores eran la causa de la diversidad clínica observada. La diferencia entre el fenotipo clásico infantil descrito por el Dr. Pompe y los fenotipos de aparición tardía (juvenil y adulta) es tan sorprendente que podrían haber sido causados por mutaciones en diferentes genes; por ejemplo, un gen que codifica la estructura de GAA y un gen que regula la expresión de *GAA*. La última hipótesis fue abandonada ya que análisis de complementación no condujeron a la corrección de la deficiencia de la proteína GAA. Más tarde, fue generalmente aceptado que una variedad de mutaciones alélicas en el gen de la α-glucosidasa ácida (*GAA*) debería estar causando la heterogeneidad clínica, pero esto no se pudo probar en ese momento por falta de tecnología. El rol de la heterogeneidad genética ha salido a la luz después de la clonación del gen *GAA* permitiendo la búsqueda de variantes patogénicas en su secuencia.

El creciente número de variantes en la secuencia del gen que se han descubierto en los últimos 15 años han llevado a iniciar la base de datos de las mutaciones de la enfermedad de Pompe (www.pompecenter.nl), las cuales han aumentado a 558 variantes de acuerdo a la versión de mayo de 2016. Estas variantes fueron clasificadas desde 'muy severas ' a ' no patogénicas', ya sea bien por su ubicación y/o naturaleza, predicción *in silico* o por su efecto sobre la síntesis y función enzimática según los resultados de estudió *in vitro*. Sin embargo, La precisión de las predicciones nunca fueron cuestionadas, ni tampoco investigadas. Por consiguiente, actualizar la base de datos de mutaciones de la enfermedad de Pompe al vincular las variantes individuales de *GAA* con los genotipos de *GAA* y los fenotipos clínicos se convirtió en uno de los principales temas de ésta tesis, lo que dio como resultado la extensión de la base de datos de las variantes de *GAA* y que reemplazó a la base de datos anterior y actualmente

se puede encontrar a través del siguiente link: www.pompevariantdatabase.nl (**capítulos 3 y 4**).

Sin lugar a dudas, uno de los eventos más emocionantes en la historia de la enfermedad de Pompe ha sido la introducción de la terapia de reemplazo enzimático en el año 2006. Sus beneficios y desafíos también se abordan en este primer capítulo, sin embargo, no han sido objeto de este proyecto de investigación.

Otra información que se describe en el capítulo 1 se relaciona con la frecuencia de la enfermedad de Pompe, la distribución geográfica, los síntomas y el espectro clínico. Las preguntas / objetivos de los estudios realizados en el contexto de esta tesis fueron formulados al final del capítulo 1 y son los siguientes:

1. ¿Cómo diagnosticar enzimáticamente la enfermedad de Pompe?
2. Actualización de la base de datos de mutaciones de la enfermedad de Pompe y facilitación del análisis de la correlación genotipo-fenotipo.
3. ¿El polimorfismo ACE contribuye a la heterogeneidad clínica de la enfermedad de Pompe?
4. ¿Pueden ser identificados marcadores de la enfermedad que puedan distinguir entre el tipo mas severo y entre el menos progresivo de la enfermedad? Los resultados de esos estudios están discutidos en un contexto mas amplio, las conclusiones generales están descritas, y las perpectivas están formuladas en el capitulo 7.

El Capítulo 2 está dedicado al diagnóstico enzimático de la enfermedad de Pompe y revisa el resultado de una serie de 2591 ensayos en 1709 casos clínicos obtenidos de diferentes tipo de muestras y utilizando diferentes procedimientos realizado durante los últimos 28 años en el departamento de Genética Clínica del Centro Médico Universitario Erasmus MC en Rotterdam . La actividad de la enzima GAA se midió en leucocitos, sangre (dried blood spots, DBS) y fibroblastos con el objetivo de establecer el mejor ensayo para fines particulares. Se descubrió que los fibroblastos cutáneos cultivados proporcionan con mucho los mejores resultados con respecto a la diferenciación entre las formas infantil clásica, juvenil y la adulta. Los leucocitos mixtos proporcionan el método de diagnóstico más rápido, por lo que la mezcla de reacción necesita contener acarbosa para inhibir la glucoamilasa. El glucógeno, así como el 4-metillumbeliferil-α-D-glucopiranosido (4MUG), son sustratos adecuados, pero el glucógeno discrimina mejor entre afectados y no afectados. La desventaja de usar glucógeno es que los portadores de la pseudodeficiencia Caucásica (*GAA2)* pueden calificarse como pacientes. En tales casos, el uso adicional de 4-methylumbelliferyl-

α-D-glucopyranoside (4MUG) resuelve el problema. Las gotas de sangre secas son la mejor fuente de muestra para fines de detección, pero no son adecuadas para el diagnóstico final, lo que requiere una segunda prueba de diagnóstico para confirmar el diagnóstico.

Los capítulos 3 y 4 describen cómo la base de datos online de las mutaciones de la enfermedad de Pompe, adscrita al anterior link: www.pompecenter.nl, y que actualmente se puede encontrar en el nuevo enlace: www.pompevariantdatabase.nl fue actualizada y modificada logísticamente para proporcionar más y mejor información sobre los efectos funcionales de las variantes de la secuencia *GAA* mediante la vinculación de cada variante (un alelo) con genotipos patógenos (una combinación de 2 alelos) y sus fenotipos clínicos asociados. El enlace directo profundiza la comprensión de la correlación genotipo-fenotipo y facilita el diagnóstico, la predicción de la gravedad de la enfermedad y el asesoramiento genético. La versión mejorada de la base de datos (**capítulo 3**) proporciona acceso rápido a todos los fenotipos clínicos asociados con un genotipo de *GAA* particular y delimita fácilmente la diversidad clínica dentro de los genotipos (**capítulo 4**). Se han descrito tipos similares de análisis en otros trastornos, pero se limitan principalmente a pacientes individuales por genotipo. Nuestro análisis genotipo-fenotipo ilustra que el 49% de las 422 variantes de *GAA* asociadas a la enfermedad están vinculadas a los fenotipos infantil clásica, la juvenil y la adulta. Además, se predice que el 38% de las variantes asociadas a la enfermedad que figuran en la base de datos son CRIM (cross-reactive immunological material, de sus siglas en ingles) negativos, mientras que la mayoría de las variantes de tipo missense son CRIM positivo (**capítulo 3**). Las nuevas características que se agregan a la base de datos de las mutaciones de la enfermedad de Pompe proporcionan una herramienta invaluable para el diagnóstico, el pronóstico de la progresión de la enfermedad y la toma de decisiones del tratamiento de la enfermedad de Pompe. El **Capítulo 4** es una primera demostración de cómo se puede usar la versión actualizada de la base de datos para fines de investigación. En este estudio analizamos y comparamos la variación fenotípica entre pacientes con los genotipos c-32-13T> G más comunes y otros genotipos de *GAA* bastante comunes. Lo cual lleva a la conclusión de que debe haber factores genéticos o ambientales que conjuntamente determinan el curso de la enfermedad de Pompe en el caso de pacientes que no tengan una total deficiencia de actividad enzimática.

En el capítulo 5 se exploró el efecto de uno de los pocos factores modificadores previamente reportados. Se trata del polimorfismo *ACE* I/D. Basado en un número bastante pequeño de pacientes, los pacientes con el genotipo ACE DD en comparación

con el genotipo ACE II parecían tener una progresión más rápida de la enfermedad y una respuesta reducida a la terapia de reemplazo enzimático. Sin embargo, algunos de los hallazgos informados son contradictorios, y se recomendaron estudios en una población de pacientes más grande. Por lo tanto, investigamos el papel del polimorfismo *ACE* en una cohorte relativamente grande de 131 niños y adultos cubriendo el completo espectro fenotípico de pacientes con el genotipo *GAA* c.-32-13T> G / 'nulo' con respecto a la edad de inicio de la enfermedad, severidad de la enfermedad y respuesta a la terapia de remplazo enzimático (ERT) durante un período de seguimiento de 5 años. En nuestro estudio no pudimos confirmar resultados anteriores: las diferencias en el polimorfismo **ACE** no pudo ni explicar la diversidad fenotípica inter-familiar ni la intra-familiar.

En el **capítulo 6**, realizamos una búsqueda de marcadores de enfermedad en la enfermedad de Pompe que podrían usarse como lectura o pronóstico de la severidad de la enfermedad. El método elegido fue analizar la expresión génica a nivel de ARN mensajero (mARN) en dos grupos de pacientes con genotipos de *GAA* similares (c.-32-13T> G / "nulo" en 35/39 casos); uno con inicio de síntomas antes de los 16 años y otro con el inicio de los síntomas después de los 35 años. Estos estudios se vieron obstaculizados por la variación inter-experimental, sin embargo tres genes se destacaron como marcadores tentativos de la enfermedad: *MAOA* (monoamina oxidasa tipo A), que cataliza la desaminación oxidativa de las aminas biogénicas y xenobióticas y tiene funciones importantes en el metabolismo de aminas neuroactivas y vasoactivas en el sistema nervioso central y tejidos periféricos, *STC1* (Stanniocalcin-1) que desempeña un papel en la regulación del transporte renal e intestinal de calcio y fosfato, el metabolismo celular o la homeostasis celular de calcio / fosfato, y *NRK* (Nik Related Kinase) que se ha asociado con la inducción de la polimerización de actina en la embriogénesis tardía. En general, estos genes se han implicado en varios procesos biológicos, incluyendo el desarrollo óseo y muscular, la homeostasis celular de calcio / fosfato, el metabolismo celular y proceso metabólico de aminas celulares. Adicionalmente, una gran cantidad de genes podrían asignarse como posibles marcadores de la enfermedad, ya que se expresaron de manera diferencial en fibroblastos de piel cultivados de pacientes afectados con la enfermedad de Pompe en comparación con los controles no afectados.

En conjunto, los hallazgos descritos en esta tesis son un paso adelante en el diagnóstico y la comprensión de la enfermedad de Pompe con énfasis en la correlación genotipo-fenotipo, como es discutido en el capítulo 7.

List of Publications

Monica Y. Niño Martinez, Mark Wijgerde, Douglas Oliveira Soares de Faria, Nadine A.M.E. van der Beek, Hannerieke J.M. van den Hout, Ans T. van der Ploeg, Frans W. Verheijen and W.W.M. Pim Pijnappel. 'Enzymatic Diagnosis of Pompe Disease: lessons from 28 years experience', *Under review.*

Monica Y. Niño Martinez, Stijn L.M. in 't Groen, Atze J. Bergsma, Nadine A.M.E. van der Beek1, Marian Kroos, Marianne Hoogeveen-Westerveld, Ans T. van der Ploeg and W.W.M. Pim Pijnappel. 'Extension of the Pompe mutation database by linking disease-associated variants to clinical severity'. Human Mutat. 2019; 40: 1954– 1967. Jun 29. doi: 10.1002/humu.23854. PMID: 31254424. *Top 10% most dowloaded papers of Human Mutation in 2018-2019.*

Monica Y. Niño, Stijn L.M. in 't Groen, Atze J. Bergsma, Marianne Hoogeveen-Westerveld, Ans T. van der Ploeg and W.W.M. Pim Pijnappel. 'The extended Pompe mutation database reveals distinct phenotypic spectra of Pompe disease-associated GAA variants', *Submitted*

Monica Yasmin Niño, Heidi Eliana Mateus, Dora Janeth Fonseca, Marian A Kross, Sandra Yaneth Ospina, Juan Fernando Mejía, Jesus Alfredo Uribe, Arnold J.J. Reuser, Paul laissue. 'Identification and Functional Characterization of *GAA* mutations in Colombian Patients affected by Pompe Disease'. Journal Inherited Metabolic Disease Reports-JIMD, 2013;7:39-48. doi: 10.1007/8904_2012_138. Epub 2012 Apr 19.

Esther Kuperus, Jan C. van der Meijden, Stijn L. M. in 't Groen, Marian A. Kroos, Marianne Hoogeveen-Westerveld, Dimitris Rizopoulos, **Monica Yasmin Nino Martinez**, Michelle E. Kruijshaar, Pieter A. van Doorn , Nadine A. M. E. van der Beek, Ans T. van der Ploeg, W. W. M. Pim Pijnappel. 'The ACE I/D polymorphism does not explain heterogeneity of natural course and response to enzyme replacement therapy in Pompe disease'. PLoS One. 2018 Dec 7;13(12):e0208854. doi: 10.1371/journal.pone.0208854. eCollection 2018. PMID: 30532252

Reuser, A.J.J., van der Ploeg, A.T., Chien, Y.-H., Llerena, J., Jr., Abbott, M.-A., Clemens, P.R., Kimonis, V.E., Leslie, N., Maruti, S.S., Sanson, B.-J., Araujo, R., Periquet, M., Toscano, A., Kishnani, P.S., on behalf of the Pompe Registry Sites, & **Niño, M.Y.**, in 't Groen, S.L.,

Bergsma, A.J., van der Beek, N.A., Kroos, M., Hoogeveen-Westerveld, M., van der Ploeg, A.T. and Pijnappel, W.P. (2019), Front Cover, *Human Mutation*, Volume 40, Issue 11, Nov 2019. doi:10.1002/humu.23934.

Curriculum Vitae

Name:	Monica Yasmin Niño Martinez
Place of birth:	Tunja-Colombia

EDUCATION

August 2013- March 2020	**PhD program in Biomedical Sciences** Center for Lysosomal and Metabolic Diseases, Department of Pediatrics, Department of Clinical Genetics Erasmus University Medical Center, Rotterdam, The Netherlands Thesis: *The genotype-phenotype correlation in Pompe Disease studied from an enzymatic and molecular perspective*
January 2007- July 2013	**Master in Human Genetics** Master in Human Genetics, School of Medicine and Health Sciences Universidad del Rosario, Bogotá, Colombia Thesis: *Identification and Functional Characterization of GAA Mutations in Colombian Patients Affected by Pompe Disease*
August 2010- Sept 2011	**Master of Science** Master in Health Sciences, Specialization Genetic Epidemiology Erasmus University Medical Center, Rotterdam, The Netherlands Thesis: *Investigation into the genotype-phenotype correlation in Colombian patients with Pompe disease*

January 1999- March 2005	**Bachelor in Bacteriology** Pontificia Universidad Javeriana, Bogotá, Colombia Thesis: *Parasitic agent study causing acute diarrhea disease in children under five years of age at Cartagena and Quibdó cities"*

Research / Professional Experience

August 2013- March 2020	PhD program in Biomedical Sciences, Erasmus MC, Rotterdam The genotype-phenotype correlation in Pompe disease studied from an enzymatic and molecular perspective
July 2012- July 2013	Full time Teacher, Universidad de Boyacá, Tunja, Colombia Function: Teaching courses such as genetics, microbiology, bacteriology, and immunology in Medicine, Bacteriology and Respiratory Therapy programs. Additionally, supervision of student's research projects
January 2007- July 2013	Master in Human Genetics, Universidad del Rosario, Bogotá, Colombia *Identification and Functional Characterization of GAA* *Mutations in Colombian Patients Affected by Pompe Disease*
February- July 2011	Pompe Center (Research internship) Erasmus MC, Clinical Genetics Department, Rotterdam, The Netherlands Under supervision of Arnold J. Reuser

February 2009- Dec 2009	Teacher, Universidad de Boyacá, Tunja, Colombia Function: Teaching courses in genetics and immunology as part of Medicine and Bacteriology programs
January 2006- July 2006	Bacteriologist, E.S.E. Health Center, Buenavista, Colombia Function: Bacteriologist (laboratory of first level of complexity), which offers service in different areas such as hematology, blood chemistry, parasitology, urine test, immunology, and mycobacteria
July 2005- January 2006	Mandatory Social Service, E.S.E. Health Center, Buenavista, Colombia Function: analysis, process and report of human samples test as part of diagnosis, and decision making about therapeutic intervention
January- June 2004	Human Genetics Institute (Research internship) Pontificia Universidad Javeriana, Bogotá, Colombia Under supervision of Sandra Ramirez Clavijo, PhD
June- December 2003	Hospital Universitario San Ignacio. (clinical laboratory internship) Pontificia Universidad Javeriana, Bogotá, Colombia

AWARDS

April, 2020	**Honorable Acknowledgments for being an author of one of the top 10% most downloaded papers of *Human Mutation* in 2018-2019** 'Extension of the Pompe mutation database by linking disease-associated variants to clinical severity'. Human Mutat. 2019; 40: 1954– 1967. Jun 29. doi.org/10.1002/humu.23854.
October, 2019	**Front Cover, *Human Mutation*, Volume 40, Issue 11, Nov 2019. doi:10.1002/humu.23934.**

August 2013- July 2017	**Scholarship to study abroad provided by Colciencias-Colombia** Colciencias is the administrative department of science, technology and innovation that supports and leads the national system of science, technology and innovation
July, 2013	**Honor distinction for excellence in graduation work** Laureate Thesis Award. Highest honor awarded for outstanding research work. Master in Human Genetics, Universidad del Rosario, Colombia
August 2010- Sep 2011	**Scholarship to study abroad provided by** ERACOL, which is an academic exchange program between Europe and Latin America

COURSES, CONFERENCES

September 2019	SSIEM Annual Symposium 2019, Rotterdam, The Netherlands
May 2017	Research Integrity Course, Erasmus MC, Rotterdam, The Netherlands
Nov 2016- Feb 2017	Course in English Biomedical Writing and Communication, Erasmus MC, Rotterdam, The Netherlands
November, 2016	8th European Symposium on Step Forward in Pompe Disease, Amsterdam, The Netherlands
October 2016	Muscles2Meet Neuromuscular Young Talent Symposium, Utrecht, The Netherlands
June, 2016	The MGC PhD workshop 2016, Dortmund, Germany (oral presentation)

June, 2015	Course on Gene Expression Data Analysis Using R: How to make sense out of your RNA-seq/ microarray data, Erasmus MC, Rotterdam, The Netherlands
May, 2015	Course on R, Erasmus MC, Rotterdam, The Netherlands
February, 2015	The CLC workbench / Ingenuity Variant Analysis Workshop, Erasmus MC, Rotterdam, The Netherlands
February, 2015	The Ingenuity Pathway Analysis (IPA) Workshop, Erasmus MC, Rotterdam, The Netherlands
June, 2014	13th International Postgraduate Course on LSDs by IMS GmbH, Johannes Gutenberg-University Mainz, Nierstein, Germany
April - May 2014	Genetics Course, Erasmus MC, Rotterdam, The Netherlands
January- March 2014	Cell and Developmental Biology, Erasmus MC, Rotterdam, The Netherlands
October- Nov 2013	Biochemistry and Biophysics, Erasmus MC, Rotterdam, The Netherlands
November, 2010	The SNP Course VII, Erasmus MC, Rotterdam, The Netherlands

PhD Portfolio

Courses:	Year	ECTS
Biochemistry and Biophysics	2013	3
Cell and Developmental Biology Course	2014	3
Genetics Course	2014	3
13th International Postgraduate Course on Lysosomal Storage Disorders	2014	2
Introduction to Next Generation Sequencing Technologies, applications and data analysis	2014	2
Safely working in the lab	2014	0.5
Gene expression data analysis using R	2015	2
The Course on R	2015	1.4
Research integrity	2017	0.3
Course in English Biomedical Writing and Communication	2017	3
Workshops:	**Year**	**ECTS**
21th MGC PhD Workshop, Munster	2014	1
MGC PhD Workshop, Maastricht	2015	1
MGC PhD Workshop, Dortmund	2016	1
The Ingenuity Pathway Analysis Workshop	2015	0.5
The Ingenuity Variant Analysis Workshop	2015	0.5
The workshop on Microsoft Excel 2010 Basic	2015	0.3
(Inter) National Meetings:	**Year**	**ECTS**
SSIEM Annual Symposium 2019	2019	1
Muscles2Meet Neuromuscular Young Talent Symposium	2016	1
8th European Symposium on Step Forward in Pompe Disease	2016	1
Sophia Research Day, Rotterdam	2014	0.5
Sophia Research Day, Rotterdam	2015	0.5
Sophia Research Day, Rotterdam	2017	0.5
Sophia Research Day, Rotterdam	2020	0.5
Lectures, Erasmus MC	2014-19	2
Clinical Genetics Meetings	2013-19	2
Presentations:	**Year**	**ECTS**
Oral Presentation 21th MGC PhD Workshop, Munster	2014	0.25
Oral Presentation 13th International Postgraduate Course on Lysosomal Storage Disorders	2014	0.5
Poster presentation SSIEM 2019 Symposium	2019	0.25
	Total ECTS	**34.5**

A

Acknowledgements

Dedicated to my family and the patients, who are the motivation of this arduous work. In memory of my mom and brother in law.

'Challenges are conquered by working together'

No words suffice to express my gratitude, but from the deepest of my heart:

"I wish to thank all the people whose assistance, support and guidance were a milestone in the completion of this project".

Thanks for all your encouragement!

Agradecimientos

Dedicado a mi familia y pacientes, quienes son la motivacion de tan ardua labor. En memoria a mi ma y cuñado.

Quiero dar un reiterado agradecimiento a todos mis amigos por sus consejos, palabras de motivacion, abrazos llenos de energia y memorables momentos compartidos. Recuerdos que permaneceran en mi memoria! En especial a esa fuerza spiritual que me acompaño, me brindo fortaleza y me guio a recorrer el camino de dificultades.

Gracias a todos por su enorme apoyo y motivacion!