Broad variation in phenotypes for common GAA genotypes in Pompe disease

Monica Y. Niño1,2,3 | Stijn L. M. in’t Groen1,2,3 | Douglas O. S. de Faria1,2,3 | Marianne Hoogeveen-Westerveld2 | Hannerieke J. M. P. van den Hout1,3 | Ans T. van der Ploeg1,3 | Atze J. Bergsma1,2,3 | W. W. M. Pim Pijnappel1,2,3

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

Abstract

Patients with the common c.-32-13T > G/null GAA genotype have a broad variation in age at symptom onset, ranging from early childhood to late adulthood. Phenotypic variation for other common GAA genotypes remains largely unexplored. Here, we analyzed variation in age at symptom onset for the most common GAA genotypes using the updated and extended Pompe GAA variant database. Patients with the c.2647-7G > A/null genotype invariably presented symptoms at adulthood, while the c.32-13T > G/null, c.546G > T/null, c.1076-22T > G/null, c.2238G > C/null, and c.2173C > T/null genotypes led to presentations from early childhood up to late adulthood. The c.1309C > T/null genotype was associated with onset at early to late childhood. Symptom onset shifted toward higher ages in homozygous patients. These findings indicate that a broad variation in symptom onset occurs for various common GAA genotypes, suggesting the presence of modifying factors. We identified three new compound heterozygous c.-32-13T > G/null patients who carried the genetic modifier c.510C > T and who showed symptom onset at childhood. While c.510C > T acted by lowering GAA enzyme activity, other putative genetic modifiers did not at the group level, suggesting that these act in trans on processes downstream of GAA enzyme activity.

KEYWORDS

GAA deficiency, genotype–phenotype, glycogenosis type II, mutation database, Pompe disease

1 | INTRODUCTION

Pompe disease is a monogenic lysosomal storage disorder caused by disease-associated variants in the acid α-glucosidase (GAA) gene (van der Ploeg & Reuser, 2008). A large number of disease-associated GAA variants have been identified and are listed in the recently updated Pompe disease GAA variant database at http://www.pompevariantdatabase.nl. The database provides GAA genotypes and phenotypes of all patients in which these were reported (de Faria et al., 2021; 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 & Reuser, 2013; Herzog et al., 2012; M. A. Kroos et al., 2007; Laforêt

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. Human Mutation published by Wiley Periodicals LLC

Human Mutation. 2021;42:1461-1472. wileyonlinelibrary.com/journal/humu | 1461
et al., 2013; Reuser et al., 2018; van der Beek et al., 2009; van der Ploeg & Reuser, 2008; van der Ploeg et al., 2017). 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 present with generalized skeletal muscle weakness and hypertrophic cardiomyopathy shortly after birth, and die within the first year of life if left untreated (Kishnani et al., 2006; H. M. van den Hout et al., 2003). Patients with some residual GAA enzyme activity have a variable onset of symptoms due to the presence of at least one milder disease-associated GAA variant. These patients develop skeletal muscle weakness, leading to impaired mobility and respiration, resulting in a wheelchair and/or ventilator dependency at some point in their life (van der Beek et al., 2009, 2012). Cardiac symptoms are usually absent in this patient group.

Treatment of Pompe disease with enzyme replacement therapy (ERT) is available since 2006. ERT improves survival and is effective in normalizing cardiac hypertrophy and improving skeletal muscle function, resulting in increased mobility and stabilization of respiratory function (Amalfitano et al., 2001; Kishnani et al., 2007; Nicollino et al., 2009; Strothotte et al., 2010; van Capelle et al., 2010; H. van den Hout et al., 2000; J. M. van den Hout et al., 2001; van der Ploeg et al., 2010). The invasive nature of ERT treatment and its high costs are important factors that need to be taken into account in determining when treatment of patients with childhood or adulthood onset of symptoms should be started. Prediction of the phenotype that is associated with a particular genotype is an important aspect of decision-making.

While most disease-associated GAA variants are rare, a number of variants occur more frequently. Common variants are enriched in distinct populations, including three variants in the Caucasian population: c.-32-13T > G (p.[0, p. =]) (also termed IVS1), c.525del (p.[Glu176Argfs*45]), and c.2481 + 102_2646 + 31del (p.[Gly828_Asn882del]) (also termed deleX18); two variants in the Asian population: c.1935C > A (p.(Asp645Glu)) and c.2238G > C (p.(Trp746Cys)); and one variant in patients from African-American descent: c.2560C > T (p.[Arg854=]) (Ausems et al., 2001; Becker et al., 1998; Dagnino et al., 2000; Huie et al., 1994; M. A. Kroos et al., 1995; Laforet et al., 2000; Lin & Shieh, 1996; Liu et al., 2014; Müller-Felber et al., 2007; Pittis & Filacomo, 2007; Reuser et al., 2018; Shieh & Lin, 1998; Van der Kraan et al., 1994; Wokke et al., 1995). The Caucasian c.-32-13T > G disease-associated variant is associated with a very broad range of age at symptom onset (Herzog et al., 2012; M. A. Kroos et al., 2007; Semplicini et al., 2018). This phenotypic variation is also observed within and between families as is illustrated by a study on 22 families with two or three siblings carrying the c.-32-13T > G variant (Wens et al., 2013). These findings have led to the hypothesis that modifying factors—genetic, epigenetic, or environmental—may alter the disease course in patients that carry the c.-32-13T > G variant (Herzog et al., 2012; Huie et al., 1994; Ko et al., 1999; M. Kroos et al., 2012; M. A. Kroos et al., 2007; Musumeci et al., 2015; Rainkar et al., 2017; Shieh & Lin, 1998; van der Ploeg & Reuser, 2008; Wens et al., 2013). It is, however, largely unknown whether phenotypic variation occurs in groups of patients with specific disease-associated variants other than c.-32-13T > G. We recently identified the GAA c.510C > T (p.(=)) variant as a genetic modifier that accelerates symptom onset in compound heterozygous and homozygous c.-32-13T > G patients (Bergsma et al., 2019), but additional putative variants that alter the course of Pompe disease remain enigmatic so far.

In the present study, we used the information included in the Pompe disease GAA variant database to compare variations in symptom onset for the most common GAA disease-associated variants in Pompe disease. We analyzed compound heterozygous and homozygous patients, investigated the effect of the second GAA allele, and made a prediction whether additional modifying factors for c.-32-13T > G patients may be present in cis or in trans. The findings on phenotypic variations associated with GAA genotypes will be important for diagnosis, genetic counseling, decision-making on the start of treatment with ERT, and for research into the identification of additional modifying factors.

2 | METHODS

2.1 | Patient information

To analyze genotypic variation, we used the most recent version of the Pompe disease GAA variant database (http://www.pompevariantdatabase.nl, update January 2020) (de Faria et al., 2021). References to the publications that contain the original patient information in this manuscript are listed in Tables S1 and S2. Geographical origins were indicated or were categorized in the following four 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 Asian (for patients from the Asian continent). For the analysis of GAA enzyme activity in fibroblasts from patients with the c.-32-13T > G variant and childhood or adulthood onset in the absence or presence of the modifier c.510C > T, we analyzed the patient cohort that was described in Bergsma et al. (2019).

2.2 | Nomenclature

Variant annotations and classification conform to recommendations of the Human Genome Variation Society (HGVS) (den Dunnen et al., 2016). NM_000152.5 was used as a reference sequence for GAA mRNA, LRG_673t1.1 was used for intronic regions, and NP_000143.2 for GAA protein. Position c.1 represents the first nucleotide of the translation start codon located in GAA exon 2.

2.3 | GAA enzyme activity assay

For the analysis of GAA enzyme activity in fibroblasts from patients with the c.-32-13T > G variant and childhood or adulthood onset in the absence or presence of the modifier c.510C > T, GAA enzyme activities were determined by the diagnostic department of Clinical Genetics of the Erasmus MC in fibroblasts using
4-methylumbelliferone-α-D-glucopyranoside (4-MU) (Sigma-Aldrich) as a substrate as previously described (M. A. Kroos et al., 2007).

## RESULTS AND DISCUSSION

### Phenotypic variation in compound heterozygous patients

The most frequent compound heterozygous genotypes in the Pompe disease GAA variant database (http://www.pompevariantdatabase.nl) 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 and that occurs in classic infantile patients that have the most severe form of Pompe disease (Niño et al., 2019). The ages at symptom onset were plotted for five types of GAA variants, based on their association with the following groups of phenotypes: adult, childhood or adult, childhood, classic infantile or childhood, and classic infantile (Figure 1 and Table 1). It should be noted that the definition of symptoms can vary between publications, and as a result could contribute to the variability in age at symptom onset, which is a limitation of this analysis. A minimum number of four patients per GAA genotype was required in order for a variant to be included in the analysis.

We note that other frequent compound heterozygous genotypes were also present in the database; in these cases, the second alleles were not null alleles. Since the clinical phenotype is the result of the combined action of two alleles, the absence of a null allele as the second variant precludes the classification of the variant in question.

In patients with the noncoding splice variant c.2647-7G > A (Caucasian), which affects splicing of GAA exon 18 and is associated with the adult phenotype in combination with a null allele, age at symptom onset ranged from 35 years to 53 years of age (median, 40 years) (Figure 1 and Table 1). The patients with this variant all

![FIGURE 1](image-url) Variations in age at symptom onset in compound heterozygous patients. Age at symptom onset is indicated for patients that carry the most frequently occurring variants at compound heterozygous state in combination with a null allele (minimum of four patients). Clinical phenotype groups are shown. Median ages at symptom onset are indicated by long vertical lines; ranges by short vertical lines. 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 at symptom onset: 25 years, phenotype: adult), who was considered to be an enigmatic case (Hermans et al., 1993)
<table>
<thead>
<tr>
<th>DNA/protein nomenclature</th>
<th>Predicted effect</th>
<th>Phenotype with a null allele</th>
<th>Total number of patients</th>
<th>Phenotype of patients</th>
<th>Population*</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.2647-7G&gt;A</td>
<td>Splicing</td>
<td>Adult</td>
<td>7</td>
<td>Adult (7)</td>
<td>Italian (7)</td>
</tr>
<tr>
<td></td>
<td>Combined with 1 null allele</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.-32-13T&gt;G</td>
<td>Splicing</td>
<td>Childhood or adult</td>
<td>463</td>
<td>Classic infantile (1)</td>
<td>French (79), United States (74), Italian (56), Dutch (48), Caucasian (30), German (28), Polish (17), Brazilian (15), Austrian (6), Hispanic (5), Iranian (4), Turkish (4), Greek (2), Colombian (2), Danish (2), British (2), Moroccan (2), Romanian (1), Chinese (1), Costa Rican (1), Serbian (1), Canadian (1), Unknown (82)</td>
</tr>
<tr>
<td></td>
<td>Combined with 67 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2238G&gt;C, p.(Trp746Cys)</td>
<td>Missense</td>
<td>Childhood or adult</td>
<td>20</td>
<td>Childhood (9)</td>
<td>Chinese (14), Taiwanese (4), South Korean (1), Caucasian (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 9 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2173C&gt;T, p.(Arg725Trp)</td>
<td>Missense</td>
<td>Childhood or adult</td>
<td>5</td>
<td>Childhood (4)</td>
<td>Hispanic (2), French (2), British (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 3 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1076-22T&gt;G</td>
<td>Splicing</td>
<td>Childhood</td>
<td>4</td>
<td>Childhood (3)</td>
<td>United States (1), Caucasian (1), German (1), Austrian (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 3 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1309C&gt;T, p.(Arg437Cys)</td>
<td>Missense</td>
<td>Childhood</td>
<td>10</td>
<td>Childhood (10)</td>
<td>Chinese (5), Japanese (3), Korean (2)</td>
</tr>
<tr>
<td></td>
<td>Combined with 4 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.546G&gt;T</td>
<td>Splicing</td>
<td>Childhood</td>
<td>9</td>
<td>Childhood (5)</td>
<td>Japanese (7), Korean (2)</td>
</tr>
<tr>
<td></td>
<td>Combined with 6 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1857C&gt;G, p.(Ser619Arg)</td>
<td>Missense</td>
<td>Classic infantile or childhood</td>
<td>5</td>
<td>Classic infantile (2)</td>
<td>Japanese (4), Korean (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 3 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1935C&gt;A, p.(Asp645Glu)</td>
<td>Missense</td>
<td>Classic infantile</td>
<td>47</td>
<td>Classic infantile (41)</td>
<td>Taiwanese (32), Chinese (12), United States (2), Thai (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 28 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.525del</td>
<td>Frameshift</td>
<td>Classic infantile</td>
<td>27</td>
<td>Classic infantile (26)</td>
<td>Dutch (11), United States (7), Italian (4), Caucasian (2), Australian (2), British (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 19 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2560C&gt;T, p.(Arg854*)</td>
<td>Nonsense</td>
<td>Classic infantile</td>
<td>25</td>
<td>Classic infantile (23)</td>
<td>Dutch Guianese (9), United States (8), Brazilian (4), Caucasian (2), Colombian (1), Unknown (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 13 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.2481+102_2646+31del</td>
<td>Gross deletion</td>
<td>Classic infantile</td>
<td>27</td>
<td>Classic infantile (27)</td>
<td>Dutch (11), United States (7), Italian (4), Brazilian (2), Hispanic (1), German (1), Unknown (1)</td>
</tr>
<tr>
<td></td>
<td>Combined with 20 null alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Population is reported as indicated in the original publication. In certain cases, only Caucasian origin has been reported.
came from a single Italian family (Sampaolo et al., 2013). It is interesting that the age at symptom onset can vary greatly even within the same family. A similar broad range of age at symptom onset has previously been observed within families with another single GAA genotype c.-32-13T > G/c.525del (Wens et al., 2013). This indicates that phenotypic variation may occur between members of the same family with the same GAA genotype.

Five variants that were associated with onset during either childhood or adulthood, if combined with a null allele, were subjected to further study. The combination of the c.-32-13T > G variant with a null allele was encountered most frequently. In total, 312 compound heterozygous patients in combination with a null allele and from which age at symptom onset was described were included from the Pompe disease GAA variant database. Approximately 90% of Caucasian patients with childhood or adulthood onset Pompe disease have been reported to carry this variant (see Bergsma et al., 2019; M.A. Kroos et al., 2007; Montalvo et al., 2006; Semplicini et al., 2018, and references therein). The c.-32-13T > G variant causes aberrant splicing of GAA exon 2, resulting in at least eight distinct aberrant splice products (Boerkoel et al., 1995; Huie et al., 1994; van der Wal et al., 2017). It also allows a low level of normal splicing, resulting in the expression of 10%-15% functional GAA protein compared to average healthy control values. The median age at symptom onset of compound heterozygous c.-32-13T > G patients combined with a null allele was 30 years, and an exceptional GAA protein compared to average healthy control values. The median age at symptom onset was described were included from the Pompe disease GAA variant database. Approximately 90% of Caucasian patients with childhood or adulthood onset Pompe disease have been reported to carry this variant (see Bergsma et al., 2019; M.A. Kroos et al., 2007; Montalvo et al., 2006; Semplicini et al., 2018, and references therein).

The c.-32-13T > G variant causes aberrant splicing of GAA exon 2, resulting in at least eight distinct aberrant splice products (Boerkoel et al., 1995; Huie et al., 1994; van der Wal et al., 2017). It also allows a low level of normal splicing, resulting in the expression of 10%-15% functional GAA protein compared to average healthy control values. The median age at symptom onset of compound heterozygous c.-32-13T > G patients combined with a null allele was 30 years, and an exceptionally broad range between <1 and 71 years of age was found, similarly as reported previously (Herzog et al., 2012; M.A. Kroos et al., 2007; Semplicini et al., 2018). We further tested whether there were sex-related differences in age at symptom onset within this group of patients. No significant differences between males and females could be detected (Figure S1; Kaplan–Meier log-rank test: p = .3).

The c.546G > T (p.(=)) variant, which is common among Japanese patients (Fukuhara et al., 2018), also showed a particularly broad range of age at symptom onset ranging from <1 to 72 years of age (median 17 years, n = 9 patients). This variant also causes aberrant splicing of exon 2, in which some aberrant splicing products are identical to those caused by the c.-32-13T > G variant (leaky wild-type splicing, complete skipping of exon 2), while other splicing products are unique to c.546G > T (Bergsma et al., 2021; Maimaiti et al., 2009).

Patients with the missense variants c.2238G > C (p.(Trp746Cys)) (Asian) and c.2173C > T (p.(Arg725Trp)) (Caucasian) and the splice variant c.1076-22T > G (Caucasian) also showed a broad variation in age at symptom onset from <1 to 41 years of age (ranges <1–41, 3–22, and <1–25 years, and medians 22.5, 8.5, and 8 years, respectively). These data indicate that patients carrying GAA variants that are associated with childhood or adult onset of the disease show a minimum range in age at symptom onset of 19 years and a maximum of 72 years, in both Asian and Caucasian populations. We note that relatively low patient numbers have been reported to date for these variants, except for c.-32-13T > G, and that it is possible that the inclusion of more patients will further broaden the range of age at symptom onset.

For only one variant, the missense variant c.1309C > T (p.(Arg437Cys)) (Asian), sufficient cases were described (n = 9) to predict that this variant in combination with a null allele is associated with childhood onset of symptoms. The age at symptom onset ranged from <1 to 10 years (median 6 years). It remains to be seen in a larger patient cohort whether this variant is always associated with childhood onset or whether it can also occur in adult-onset patients when combined with a null allele.

The c.1857C > G (p.(Ser619Arg)) missense variant was found in combination with a null allele in five Asian patients. The patients diagnosed with this combination presented either with the classic infantile form or a slightly milder form (Fukuhara et al., 2018). In compound heterozygous patients with the c.1857C > G/null genotype, there was very little phenotypic variation with symptom onset <1 year of age in all five patients. However, homozygous patients showed more variation in symptom onset (see below), indicating that the c.1857C > G (p.(Ser619Arg)) variant allows for the production of some residual GAA enzymatic activity, and can thus be classified as a classic infantile or childhood variant.

By definition, the classic infantile variants c.1935C > A (p.(Asp645Glu)) (missense), c.525del (frameshift causing deletion), c.2560C > T (p.(Arg854*)) (nonsense), and c.2481 + 102_2646 + 31del (in-frame deletion) were invariably associated with an age at symptom onset of <1 year of age.

3.2 Phenotypic variation in homozygous patients

Patients with specific homozygous variants can help to provide more insight into phenotypes of these variants since these can be assessed for only a single disease-associated variant without the presence of a second different GAA disease-associated variant. A priori, one would expect that homozygous patients with childhood or adulthood onset would have a median age at symptom onset that is shifted toward a higher age relative to their compound heterozygous counterpart in which the variant is combined with a null allele. This shift is driven by a higher level of residual GAA enzymatic activity produced from the second allele compared to activity in compound heterozygous patients that carry the one copy of the variant combined with a null allele.

One such example is the c.-32-13T > G variant. Twenty-two symptomatic homozygous c.-32-13T > G patients are described in the database. The median age at symptom onset of these patients was 37 years, which was 7 years older compared to patients with the c.-32-13T > G/null allele GAA genotype (Figure 2 and Table 2). The range of age at symptom onset was very broad, from <1 year to 58 years of age, which was comparable to the range of age at symptom onset in compound heterozygous patients (range <1–71 years).

As asymptomatic individuals (e.g., from newborn screening, or siblings of patients, or individuals who are homozygous for c.-32-13T > G, see below) were not taken into account in the present analysis and often remain undiagnosed or are not reported, the median age at symptom onset in Figure 2 and Table 2 is likely biased and is in reality higher. The common c.-32-13T > G variant leads to a residual activity of 10%–15% per allele (Boerkoel et al., 1995; Huie et al., 1994; van der Wal et al., 2017). When present at homozygous state, patients with the c.-32-13T > G variant should have 20%–30%
residual enzyme activity, which is just above the disease threshold of 20% (threshold value according to patient ranges used at the Erasmus MC assessed in fibroblasts with 4MU as substrate). This leads to the prediction that the majority of homozygous c.-32-13T > G patients remain asymptomatic. Indeed, it was previously noted that the frequency of the common c.-32-13T > G variant in the Dutch population is 1 in 154, which implies the existence of 170 individuals in the Netherlands that carry c.-32-13T > G in a homozygous state (Ausems et al., 1999). Currently, our center does not see any patients with this genotype, which suggests that it most often does not lead to a disease phenotype. Others have drawn similar conclusions (Musumeci et al., 2015; Rairikar et al., 2017; Semplicini et al., 2018). The broad range of age at symptom onset in the present analysis formally excluded the possibility that the second allele was responsible for the broad range of age at symptom onset in compound heterozygous c.-32-13T > G patients in this study and implies the involvement of modifying factors that delay or accelerate disease progression in c.-32-13T > G patients, as suggested previously (Herzog et al., 2012; Ko et al., 1999; M. Kroos et al., 2012; M. A. Kroos et al., 2007; Musumeci et al., 2015; Rairikar et al., 2017; Semplicini et al., 2018; Shieh & Lin, 1998; van der Ploeg & Reuser, 2008; Wens et al., 2013). A previous analysis of homozygous c.-32-13T > G patients showed that the c.510C > T variant acts as such a modifying factor and is overrepresented in patients with this genotype (Bergsma et al., 2019). The c.510C > T variant lowers enzymatic activity and can lead to a residual GAA enzymatic activity below the disease threshold. Given the spread of homozygous c.-32-13T > G patients that develop symptoms across several European countries, the United States, Colombia, Algeria, and Turkey (Table 2), it will be interesting to assess whether genetic modifiers, such as c.510C > T, have a particular geographical distribution. Because the minor allele frequency of c.510C > T is too low to be reported, this will require a more extensive genetic analysis of a larger cohort of patients.

A similar conclusion can be drawn for other variants that were not associated with the classic infantile phenotype. In the case of c.546G > T, the median age at symptom onset in homozygous patients was higher than in compound heterozygous patients in combination with a null allele (37.5 vs. 17 years, respectively) but the difference in age between the earliest and latest onset of disease was similar (67.5 vs. 72 years, respectively). The c.-32-3C > A splicing variant was associated with childhood onset when combined with a null allele, but with onset at adulthood when present at homozygous state (median 35, range 31–43 years). The range of age at symptom onset in patients with compound heterozygous c.-32-3C > A in combination with a null allele could not be assessed with certainty since there have been insufficient patients reported with this GAA genotype. Patients with the c.1857C > G (p.(Ser619Arg)) missense variant had a broad range of age at symptom onset in homozygous patients that ranged from <1 to 17 years (median 7 years), while it was only found in early childhood patients (median <1 year, range 2 years) when present at compound

![FIGURE 2](https://example.com/figure2.png)
The c.1064T > C (p.(Leu355Pro)) variant was found in six homozygous patients but the information in compound heterozygous patients who carried a null allele was lacking; homozygous patients showed a median age at symptom onset of <1 year (range <1 – 6 years).

Confirming their classification as null alleles, the variants c.1935C > A (p.(Asp645Glu)), c.2560C > T (p.(Arg854*)), and c.2481 + 102_2646 + 31del showed invariable symptom onset before the age of 1 year when present at the homozygous state or at compound heterozygous state. In addition, 10 patients were homozygous for the null allele c.1933G > A (p.(Asp645Asn)), and all patients had an age at symptom onset of <1 year.

We conclude from these analyses that the common GAA variants examined display phenotypic variation in age at symptom onset of up to 71 years between the youngest and the oldest patient when combined with a null allele and/or when present at a homozygous state. Broad phenotypic variation with respect to age at symptom onset and severity of disease was seen both in patients with splicing variants (c.−32−13T > G, c.546G > T, c.1076−22T > G, c.−32−3C > A) and missense variants (c.2238G > C (p.(Trp746Cys)), c.2173C > T (p.(Arg725Trp)), c.1309C > T (p.(Arg437Cys)), c.1857C > G (p.(Ser619Arg)), c.1064T > C (p.(Leu355Pro))) and was independent of ethnicity. This suggests that disease-associated and potentially variant-specific modifying factors exist for Pompe disease that can delay or accelerate the progression of the disease course.

### 3.3 Phenotypic variation in patients with the c.-32-13T > G variant: Effect of c.510C > T and of the second allele

We previously reported that the GAA c.510C > T variant is a silent variant that occurs in cis with the c.-32-13T > G variant in a subset of c.-32-13T > G patients. The c.510C > T variant was found in 8/136 compound

---

**TABLE 2** Information on variants and patients described in Figure 2

<table>
<thead>
<tr>
<th>DNA/protein nomenclature</th>
<th>Predicted effect</th>
<th>Phenotype with a null allele</th>
<th>Total number of patients</th>
<th>Phenotype of the patients</th>
<th>Population*</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.-32-13T &gt; G</td>
<td>Splicing</td>
<td>Childhood or adult</td>
<td>47</td>
<td>Childhood (7) Adult (16)  Unknown (15)</td>
<td>German (10), French (3), Italian (3), Portuguese (2), United States (2), Colombian (1), Caucasian (1), Algerian (1), Turkish (1), Unknown (14)</td>
</tr>
<tr>
<td>c.546G &gt; T</td>
<td>Splicing</td>
<td>Childhood</td>
<td>8</td>
<td>Childhood (1) Adult (7)</td>
<td>Japanese (8)</td>
</tr>
<tr>
<td>c.-32-3C &gt; A</td>
<td>Splicing</td>
<td>Childhood</td>
<td>7</td>
<td>Adult (7)</td>
<td>Brazilian (7)</td>
</tr>
<tr>
<td>c.1857C &gt; G, p.(Ser619Arg)</td>
<td>Splicing</td>
<td>Classic infantile or childhood</td>
<td>5</td>
<td>Classic infantile (1) Childhood (4)</td>
<td>Japanese (5)</td>
</tr>
<tr>
<td>c.1064T &gt; C, p.(Leu355Pro)</td>
<td>Missense</td>
<td>Classic infantile or childhood</td>
<td>10</td>
<td>Classic infantile (4) Childhood (6)</td>
<td>Italian (3), Arab (3), Syrian (1), Colombian (1), Portuguese (1), Unknown (1)</td>
</tr>
<tr>
<td>c.2015G &gt; A, p.(Arg672Gln)</td>
<td>Missense</td>
<td>Classic infantile or childhood</td>
<td>5</td>
<td>Classic infantile (1) Childhood (4)</td>
<td>Japanese (3), Chinese (1), Unknown (1)</td>
</tr>
<tr>
<td>c.1933G &gt; A, p.(Asp645Asn)</td>
<td>Missense</td>
<td>Classic infantile</td>
<td>11</td>
<td>Classic infantile (10) Childhood (1)</td>
<td>United States (4), Italian (3), Indian (2), Unknown (2)</td>
</tr>
<tr>
<td>c.1935C &gt; A, p.(Asp645Glu)</td>
<td>Missense</td>
<td>Classic infantile</td>
<td>37</td>
<td>Classic infantile (34) Childhood (2) Unknown (1)</td>
<td>Taiwanese (29), Thai (3), Chinese (2), Japanese (1), Unknown (2)</td>
</tr>
<tr>
<td>c.2560C &gt; T, p.(Arg854*)</td>
<td>Nonsense</td>
<td>Classic infantile</td>
<td>23</td>
<td>Classic infantile (17) Unknown (6)</td>
<td>United States (12), Brazilian (4), African (2), Pakistani (1), Arab (1), Nigerian (1), French Guianese (1), Unknown (1)</td>
</tr>
<tr>
<td>c.525del</td>
<td>Frameshift</td>
<td>Classic infantile</td>
<td>11</td>
<td>Classic infantile (11)</td>
<td>Dutch (6), Italian (2), Caucasian (2), Unknown (1)</td>
</tr>
</tbody>
</table>

*Population is reported as indicated in the original publication. In certain cases, only Caucasian origin has been reported.
heterozygous c.-32-13T > G patients and in 2/4 homozygous c.-32-13T > G patients (Bergsma et al., 2019). Here, we report three additional c.-32-13T > G/null allele patients with Pompe disease who carried c.510C > T and who developed symptoms at a relatively young age during childhood at 3, 9, and 12 years of age. GAA enzyme activities in fibroblasts were determined for two of the three patients and averaged 8.9 nmol/h/mg, which was lower than in fibroblasts from patients with c.-32-13T > G/null GAA genotypes that lacked c.510C > T (12.4 nmol/h/mg) (Bergsma et al., 2019). Our findings support our earlier conclusion that the c.510C > T variant is associated with a childhood disease presentation of symptoms in patients with the c.-32-13T > G/null allele genotype. The updated ages at symptom onset in compound heterozygous c.-32-13T > G + c.510C > T patients collected from the Pompe disease GAA variant database are plotted in Figure 3. Seven of these patients carried a null allele as the second allele. The median age at symptom onset of these patients was 3 years (range <1–13), similar to what was reported previously (Bergsma et al., 2019).

To assess a possible effect of the second allele in the total population of compound heterozygous c.-32-13T > G patients present in the database, we analyzed symptom onset in all patients with the same second allele (c.525del), any null allele, an unknown (disease-associated) allele, and a verified non-null allele (i.e., an allele that is confirmed to produce at least some residual enzymatic activity). Importantly and in agreement with previous reports (Wens et al., 2013), patients with the GAA c.-32-13T > G/c.525del genotype showed a similar median age at symptom onset and age range (median 30 years for both groups and an age range of <1–70 and <1–71 years, respectively). This confirmed that phenotypic variation in either patient group was not caused by differences in severity of the second allele (Bergsma et al., 2019; Wens et al., 2013). Patients with the genotype c.-32-13T > G/unknown (disease-associated) also showed a similar median age at symptom onset (28.5 years) and range (79 years). This suggested that the majority of unknown (disease-associated) variants in the database are null alleles. Identification of classic infantile patients that carry one of these alleles is required to confirm this. c.-32-13T > G patients in which the second allele was not a null allele showed a slightly higher median age at symptom onset of 37 years compared to c.-32-13T > G patients with a null allele (age range <1–70) (Figure 3).

We conclude that the very broad range of age at symptom onset in patients with the c.-32-13T > G variant is not caused by the second allele, and that most c.-32-13T > G patients carry a second null allele. Furthermore, when the c.510C > T variant occurs in cis with c.-32-13T > G, it acts as a negative modifying factor that is associated with an onset of symptoms during childhood.

3.4 In search of additional modifying factors for c.-32-13T > G patients

The c.510C > T variant occurs in a subset of c.-32-13T > G patients and lowers GAA enzyme activity by worsening splicing outcomes. The presence of the c.510C > T variant predicted symptom onset at childhood, but not all patients with childhood onset carried c.510C > T (Bergsma et al., 2019). This suggested that additional modifying factors exist that may explain childhood presentation in c.-32-13T > G patients that do not carry c.510C > T. To test whether such putative modifying factors would act by lowering GAA enzyme activity, like c.510C > T does, we compared GAA enzyme activities in fibroblasts from compound heterozygous c.-32-13T > G patients with childhood onset but without c.510C > T with compound heterozygous patients with adulthood onset without c.510C > T (Figure 4). This showed no difference in GAA enzyme activity between the two groups. We conclude that putative modifying factors associated with early onset of symptoms other than c.510C > T do not lower GAA enzyme activity in fibroblasts at the group level, suggesting that the majority of such putative modifying factors likely act in trans, for the following reason. While cis-acting

![Figure 3](image-url) Effects of the second allele in patients carrying the c.-32-13T > G variant. Age range is indicated with a black horizontal line. Median ages at symptom onset are indicated by long vertical lines; ranges by short vertical lines.
genetic factors could still play a role in patients lacking c.510C > T, such putative factors, for example, distant enhancers, can be expected to alter GAA expression, which would be reflected at the level of GAA enzyme activity. The fact that GAA enzyme activity was not different between patients with childhood onset that lack c.510C > T and with adulthood onset that lack c.510C > T led us to conclude that at the group level, trans-acting factors that act downstream of regulating GAA activity, for example, factors that regulate lysosomal or skeletal muscle homeostasis, seem to be the most likely explanation for the differences in symptom onset among such patients. This conclusion does not exclude the possibility that for individual patients, cis-acting factors could still play a role. In addition, our conclusion relies on the assumption that enzyme activity in fibroblasts in vitro is a good proxy of enzyme activity in skeletal muscles in vivo, which remains to be determined. It will be interesting to verify these results in muscle cells derived from primary biopsies or patient-derived induced pluripotent stem cells (van der Wal et al., 2018) because muscle cells rather than fibroblasts better represent the primary affected cell type. Genome-wide studies using large patient cohorts will help to identify putative modifying factors, and advanced in vitro model systems, such as 2D or 3D models, derived from induced pluripotent stem cells along with gene-corrected controls (Juliano et al., 2020; van der Wal et al., 2018) are required to unequivocally demonstrate their biological effect and to study their mechanism of action. We speculate that candidate modifying factors may act in trans and may modulate cellular pathways that are known to be involved in the progression of Pompe disease, such as glucose metabolism, lysosomal biogenesis, autophagy, and skeletal muscle strength and endurance.

4 | CONCLUSION

Previous work has reported on the large variability in age at symptom onset of patients with the c.-32-13T > G GAA variant. Here, we compared disease onset of patients carrying the c.-32-13T > G variant with patients carrying other variants and conclude that nine distinct GAA disease-associated variants (c.2647-7G > A; c.546G > T; c.2238G > C (p.Trp746Cys)); c.2173C > T (p.Arg725Trp)); c.1076-22T > G; c.1309C > T (p.Arg437Cys)); c.-32-3C > A; c.1857C > G (p.Ser619Arg)); c.1064T > C (p.Leu355Pro)) are associated with a large variation in age at symptom onset when present at compound heterozygous state in combination with a null allele and/or at homozygous state. Altogether, these findings suggest a strong influence of modifying factors that affect the age at symptom onset in patients with Pompe disease. Furthermore, additional patients carrying c.510C > T in cis with the c.-32-13T > G variant were reported, confirming that the c.510C > T variant is a disease modifier that accelerates age at symptom onset when combined with the c.-32-13T > G variant. Importantly, the obtained evidence showed that the majority of putative additional genetic modifiers besides c.510C > T do not lower GAA enzyme activity in fibroblasts, suggesting that these putative modifiers act in trans.

ACKNOWLEDGMENTS

We thank Dr. Arnold J. J. Reuser for his critical review of the manuscript and Ton Verkerk for his invaluable assistance with manuscript and Ton Verkerk for his invaluable assistance with generating the database. We thank Philip Lijnzaad for his help with the statistical analysis. This study was supported by grants from the Administrative Department of Science, Technology and Innovation (Colciencias, Colombia) and Sanofi Genzyme. The collaboration project is co-funded by the PPP Allowance made available by Health-Holland, Top Sector Life Sciences & Health, to the Princes Beatrix Spierfonds to stimulate public–private partnerships (project numbers: LSHM17075 and LSHM19015). Research on Pompe disease at Erasmus MC is financially supported by “Princes Beatrix Spierfonds” (project numbers: 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 Council of Technological and Scientific Development,” Brasil (grant number: 234407/2014-0); Colciencias and Sanofi Genzyme. 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.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

The data described in this study are freely accessible in the Pompe disease GAA variant database at http://www.pompevariantdatabase.nl.
WEB RESOURCE

Pompe disease GAA variant database: http://www.pompevariant database.nl

ORCID

Atze J. Bergsma http://orcid.org/0000-0002-8733-5078
W. W. M. Pim Pijnappel http://orcid.org/0000-0002-7042-2482

REFERENCES


storage disease type II: Frequency of three common mutant alleles and their associated clinical phenotypes studied in 121 patients. *Journal of Medical Genetics*, 32(10), 836–837.


SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.