



National Tay-Sachs &
Allied Diseases Association

Variant (Mutation) Database

Please note that the term “mutation” has been updated and replaced with “variant.” More information regarding this change may be found in: “Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology.” PMID: 25741868.

Table of Contents

VARIANT (MUTATION) DATABASE	1
METHODS	3
GM2 GANGLIOSIDOSES	3
HEXA (TAY-SACHS) – BETA-HEXOSAMINIDASE A	3
Function:	3
Structure:	4
TAY-SACHS DISEASE	4
Genetics:	4
Inheritance:	4
Onset/Symptoms:	4
Genotype-Phenotype Correlations:	4
Therapeutic Approaches:	5
Prevalence:	5
Summary Table of Most Common Variants:	5
Other Variants of Note:	6
HEXB (SANDHOFF) – BETA-HEXOSAMINIDASE B	6
Function:	6
Structure:	6
SANDHOFF DISEASE	7
Genetics:	7
Inheritance:	7
Symptoms:	7
Genotype-Phenotype Correlations:	7
Therapeutic Approaches:	8
Prevalence:	8
Summary Table of Most Common Variants:	8
Other Variants of Note:	9
GM2A DEFICIENCY (AB VARIANT) – GANGLIOSIDE GM2 ACTIVATOR	9
Function:	9
Structure:	9
Genetics:	9
Inheritance:	9

Symptoms:.....	9
Therapeutic Approaches:	10
Prevalence:	10
ASPA (CANAVAN DISEASE) - ASPARTOACYLASE	10
Function:	10
Structure:.....	10
CANAVAN DISEASE.....	10
Genetics:	10
Inheritance:.....	10
Symptoms:.....	10
Genotype-Phenotype Correlations:	11
Therapeutic Approaches:	11
Prevalence:	11
Summary Table of Most Common Variants:	12
GLB1 (GM1 GANGLIOSIDOSIS) – BETA-GALACTOSIDASE.....	13
Function:	13
Structure:.....	13
GM1 GANGLIOSIDOSIS	13
Genetics:	13
Inheritance:.....	14
Symptoms:.....	14
Genotype-Phenotype Correlations:	14
Therapeutic Approaches:	14
Prevalence:	14
Summary Table of Most Common Variants:	15
Other Variants of Note:.....	15
REFERENCES.....	16
APPENDIX	16

Methods

GeniE, the Genetic Prevalence Estimator (<https://genie.broadinstitute.org/>) from the BROAD institute, was used for variant database construction. GeniE is a tool that uses data from the open source databases ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>) and gnomAD (<https://gnomad.broadinstitute.org/>) to estimate prevalence for recessive diseases.

- Variants for each gene were searched and compiled into an excel sheet. Next, variants with the highest allele frequencies that were either pathogenic or likely pathogenic (based on ClinVar) were selected and included in this document. They were ranked from highest allele frequency to lowest in the table of major variants for each gene.
- Pubmed IDs (PMIDs) were included for references because these can be used to find the relevant papers in PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) which is a free resource that supports biomedical and life sciences literature.
- Additional NTSAD internal documents, namely our GenePages, were also used to compile this information.
- A glossary describing many of the relevant terms used within this document can be found at: <https://ntsad.org/diseases/glossary/>

Disclaimer: This document is meant as a reference for general information, highlighting some of the more prevalent variants and basic disease information. It serves as a reference point for general information regarding these various disorders and their respective variants. If a variant of interest isn't covered within this document, please reach out and we can help connect you with additional resources. Additionally, information included was current at the time of creation of this document (May, 2024).

GM2 Gangliosidoses

GM2 gangliosidoses are a collection of lysosomal storage diseases which are characterized by the accumulation of GM2 gangliosides within the lysosome. Three diseases fall within this classification, including Tay-Sachs disease (TSD), Sandhoff disease (SD), and GM2 Activator Protein Deficiency (AB variant).

HexA (Tay-Sachs) – Beta-Hexosaminidase A

Function: HexA is a lysosomal enzyme responsible for breaking down various glycolipids, including GM2 gangliosides. GM2 gangliosides are lipid molecules which are enriched in lipid rafts of the cell membrane helping to modulate ion channels and membrane proteins. These glycosphingolipids are particularly abundant in the cells of the central nervous system, including neurons, and therefore, the enzymatic function of HexA is crucial in maintaining cellular function and homeostasis.

Structure: The *HEXA* gene is a 40 Kb, 14-exon gene located on chromosome 15q23, and it directs the synthesis of the alpha-subunit of the heterodimeric HexA enzyme. The HexA heterodimer is comprised of an alpha and a beta subunit. The alpha subunit is encoded by *HEXA* while the beta subunit is encoded by *HEXB*. Together, these subunits form an active enzyme. Please note that by convention gene names are italicized, thus the terms *HEXA* or *HEXB* refer to the genes, and the term HexA to the enzyme.

Tay-Sachs Disease

Genetics: Tay-Sachs disease is a rare, lysosomal storage disorder that is a result of pathogenic variants in the *HEXA* gene, which impacts the functionality of the vital lysosomal enzyme, HexA. This enzyme is responsible for breaking down a group of glycosphingolipids called GM2 gangliosides, which are highly prevalent in neurons of the central nervous system. Pathogenic variants in the *HEXA* gene lead to a reduced or absent functional HexA, and this can cause GM2 gangliosides to accumulate within the lysosomes of neurons in the brain and spinal cord.

Inheritance: Tay-Sachs disease is inherited in an autosomal recessive pattern, indicating that two copies of the defective gene are necessary for disease development. Carriers of *HEXA* variants are typically asymptomatic.

Onset/Symptoms: Tay-Sachs is classified into three categories based on severity and age of onset.

- The classic infantile form is the most common and the most severe, with symptoms appearing within the first few months of life followed by a rapid and fatal progression. Early signs include developmental delays, loss of motor skills, weakness, exaggerated startle response, and a cherry-red spot on the macula.
- The juvenile onset manifests typically between ages 2 and 10 years, with patients initially presenting with clumsiness, muscle weakness, and loss of coordination. The cherry-red spot is not consistently observed in this population of patients. Disease progression is relatively rapid and becomes fatal around 15-18 years of age.
- The adult onset or late-onset Tay-Sachs disease (LOTS) is the least prevalent of the three and is less aggressive due to higher residual enzymatic activity (~5-20% of normal activity). Symptoms can emerge in adolescences but can also appear later in life, around 20-30 years of age. Patients typically present with progressive motor neuron disease, spinocerebellar degeneration and dystonia. Psychiatric manifestations are present in a subgroup of LOTS patients.

Genotype-Phenotype Correlations: In general, an inverse correlation between residual enzymatic activity and disease severity has been observed. This is reflected in the findings that individuals with two null alleles generally present with infantile Tay-Sachs, while individuals with one null and one missense variant usually have the juvenile-onset form, and finally, individuals with two missense variants typically have the milder, late-onset phenotype (PMID: 20301397).

Therapeutic Approaches: Currently there is no cure for Tay-Sachs disease, with therapeutic approaches targeting symptom management, including substrate reduction therapy and enzyme replacement therapy (PMIDs: 30524313, 34450229, 33304924). However, there are ongoing clinical trials exploring the therapeutic potential of gene therapy for this monogenic disease (PMID: 35145305).

Prevalence:

- Carrier Frequencies:
 - Ashkenazi Jewish: 1 in 30 individuals
 - French-Canadian: Gaspesie: 1 in 13 individuals
 - French-Canadian: Other: 1 in 73 individuals
 - Irish: 1 in 41 individuals
 - Caucasian: 1 in 90
 - Sephardic Jewish – Moroccan and Iraqi: 1 in 125 individuals
 - East Asian: 1 in 210 individuals
 - African: 1 in 216 individuals
 - Latino: 1 in 243 individuals
 - Finnish: 1 in 399 individuals
 - South Asian: 1 in 416 individuals
 - World Wide: 1 in 121 individuals
- Prevalence:
 - Ashkenazi Jewish:
 - Historically: 1 in 3,600 births
 - The implementation of widespread carrier screening programs and genetic counseling has reduced the number of babies born with Tay-Sachs over 90% in the last 50 years
 (<https://www.jewishgenetics.org/articles/tay-sachs-and-carrier-screening-how-they-shaped-the-jewish-community/>)
 - French-Canadian: 1 in 2,500 births
 - Non-Jewish Population: 1 in 320,000 to 1 in 350,000 births
 - Other Populations: 1 in 250,000 to 1 in 400,000 births

Summary Table of Most Common Variants:

cDNA Name & Alias (if applicable)	Protein Name	Allele Frequency	Clinical Significance ClinVar (2024)	VEP Consequence	Notes & Select References
c.1073+1G>A (IVS9+1G>A)		0.000603976	Pathogenic/ Likely pathogenic	splice donor	Common in non-Jewish Caucasians, PMIDs: 1301938, 9150157, 1837283, 21228398, https://www.ncbi.nlm.nih.gov/books/NBK1218/

c.1274_1277dupTATC (1278insTATC)	p.Tyr427IlefsTer5	0.00053781	Pathogenic	frameshift	Most common in Ashkenazi Jewish population, PMIDs: 8230592, 2848800, 22109873, 21228398, https://www.ncbi.nlm.nih.gov/books/NBK1218/
c.805G>A	p.Gly269Ser	0.000138175	Pathogenic/ Likely pathogenic	missense	Most common in late-onset Tay-Sachs (LOTS), PMIDs: 2522660, 28739864, 27896118, 31076878, 21228398, https://www.ncbi.nlm.nih.gov/books/NBK1218/
c.1496G>A	p.Arg499His	6.81501E-05	Pathogenic	missense	May be associated with juvenile Tay-Sachs, PMIDs: 2140574, 17015493, 14566483, 14577003
c.1421+1G>C (IVS12+1G>C)		6.75948E-05	Pathogenic	splice donor	Second-most common in Ashkenazi Jewish population PMIDs: 8230592, 22109873, 27879213, 12811781, https://www.ncbi.nlm.nih.gov/books/NBK1218/

To access the literature shown in the last column, enter the PMID number into the PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) search box.

Other Variants of Note:

1. c.207-2357_253+5128delinsG (7.6kb del): This is the most common pathogenic variant among French Canadian populations (PMID: 2824459, <https://www.ncbi.nlm.nih.gov/books/NBK1218/>).
2. c.533G>A, p.Arg178His: This is the most frequent B1 allele, which produces an inactive Hex A enzyme. A B1 variant produces a false negative by enzyme analysis, thus, DNA analysis must be used to determine if a patient is truly a carrier. A B1 variant is a pathogenic variant (PMID: 2973311, <https://www.ncbi.nlm.nih.gov/books/NBK1218/>).
3. c.739C>T, p.Arg247Trp and c.745C>T, p.Arg249Trp: Both variants are pseudodeficiency alleles. A pseudodeficiency allele produces a false positive by enzyme analysis, thus, DNA analysis must be used to determine if this patient is truly a carrier. A pseudodeficiency allele is a benign variant (PMIDs: 7902672, 1384323, 9150157, <https://www.ncbi.nlm.nih.gov/books/NBK1218/>).

HexB (Sandhoff) – Beta-Hexosaminidase B

Function: HexB is a lysosomal enzyme responsible for breaking down various glycolipids, including GM2 gangliosides. GM2 gangliosides are lipid molecules which are enriched in lipid rafts of the cell membrane helping to modulate ion channels and membrane proteins. These glycosphingolipids are particularly abundant in the cells of the central nervous system, including neurons, and therefore, the enzymatic function of HexB is crucial in maintaining cellular function and homeostasis.

Structure: The *HEXB* gene is 35-40 Kb, 14-exons, and located on chromosome 15q13.3. It directs the synthesis of the beta-subunit of the enzyme HexA. The HexA heterodimer is

comprised of an alpha and a beta subunit, while HexB is a homodimer of 2 beta subunits. The alpha subunit is encoded by *HEXA* while the beta subunit is encoded by *HEXB*. Together, these subunits form an active enzyme.

Sandhoff Disease

Genetics: Sandhoff disease is a rare, monogenic, lysosomal storage disorder which results from pathogenic variants in the *HEXB* gene, impacting the functionality of a vital lysosomal enzyme, beta-hexosaminidase. This enzyme is responsible for breaking down a group of glycosphingolipids called GM2 gangliosides, which is highly prevalent in neurons of the central nervous system. Pathogenic variants in the *HEXB* gene lead to a reduced or absent functional beta-hexosaminidase causing GM2 gangliosides to accumulate within the lysosomes of neurons in the brain and spinal cord. To date, more than 100 pathogenic variants have been reported in the *HEXB* gene in patients affected by Sandhoff disease (MIM #268800), including missense/nonsense, splicing, small deletions, small insertions, and gross deletions (PMID: 31428437).

Inheritance: Sandhoff disease is inherited in an autosomal recessive pattern, indicating that two copies of the defective gene are necessary for disease development. Carriers of *HEXB* variants are typically asymptomatic.

Symptoms: Sandhoff disease is classified into three categories based on severity and age of onset.

- The classic infantile form is the most common and the most severe, with symptoms appearing within the first few months of life followed by a rapid and fatal progression. Early signs include developmental delays, loss of motor skills, weakness, exaggerated startle response, and a cherry-red spot on the macula.
- The juvenile onset manifests typically between ages 2 and 10 years, with patients initially presenting with clumsiness, muscle weakness, and loss of coordination. The cherry-red spot is not consistently observed in this population of patients. Disease progression is relatively rapid and becomes fatal around 15-18 years of age.
- The adult onset or late-onset Tay-Sachs disease (LOTS) is the least prevalent of the three and is less aggressive due to higher residual enzymatic activity (~5-20% normal activity). Symptoms can emerge in adolescences but can appear later in life, around 20-30 years of age. Patients typically present with progressive motor neuron disease, spinocerebellar degeneration and dystonia.

Genotype-Phenotype Correlations: There are three variants that, when in the homozygous or compound heterozygous state with null variants, have been correlated with acute infantile Sandhoff disease; c.76delA, c.115delG, and c.445+1G>A. Generally, it is also acknowledged that individuals who have two null variants will present with the acute infantile phenotype, while individuals with one null and one missense or two missense variants have juvenile and late-

onset phenotypes, respectively. Overall, there is a strong, inverse correlation between residual enzymatic activity and disease severity (<https://www.ncbi.nlm.nih.gov/books/NBK579484/>).

Therapeutic Approaches: Currently there is no cure for Sandhoff disease, with available treatments targeting symptom management.

Prevalence:

- Carrier Frequencies:
 - Ashkenazi Jewish: 1 in 170 individuals
 - African: 1 in 895 individuals
 - East Asian: 1 in 385
 - Finnish: 1 in 2913 individuals
 - Caucasian: 1 in 202 individuals
 - Latino: 1 in 248 individuals
 - South Asian: 1 in 513 individuals
 - Northern Saskatchewan Metis: 1 in 15 individuals
 - Argentinian Creole: 1 in 26 individuals
 - World Wide: 1 in 286 individuals
- Prevalence:
 - General Populations: 1 in 500,000 births (PMID: 36835039)

Summary Table of Most Common Variants:

cDNA name & alias (if applicable)	Protein Name	Allele Frequency	Clinical Significance ClinVar (2024)	VEP Consequence	Notes & Select References
c.1250C>T	p.Pro417Leu	0.000776943	Pathogenic/ Likely pathogenic	missense	PMIDs: 1531140, 1386607, 24263030
c.115delG	p.Val39TrpfsTer25	5.53858E-05	Pathogenic	frameshift	Common variant in individuals with Métis ancestry in northern Saskatchewan, PMIDs: 24461908, 22191674, 23046579, https://www.ncbi.nlm.nih.gov/books/NBK579484/
c.1514G>A	p.Arg505Gln	4.33753E-05	Pathogenic/ Likely pathogenic	missense	Heat labile allele, can lead to false negative by enzyme analysis, may be associated with Late Onset, PMIDs: 8950198, 12027830, 23759947, 8357844
c.1509-26G>A (IVS12-26G>A)		3.80606E-05	Pathogenic/ Likely pathogenic	intron	PMIDs: 2522450, 17015493, 20798201, 27021291
c.850C>T	p.Arg284Ter	3.71958E-05	Pathogenic	stop gained	PMIDs: 23046579, 8162015, 18758829, 24613245, 26582265, 29448188
c.796T>G	p.Tyr266Asp	3.28929E-05	Pathogenic/ Likely pathogenic	missense	PMIDs: 17015493, 28895707

To access the literature shown in the last column, enter the PMID number into the PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) search box.

Other Variants of Note:

1. c.445+1G>A: This variant is common in the Creole population in northern Argentina (PMID: 8076944).
2. c.76delA: Common among the Maronite community in Cypress (PMID: 10982028).

GM2A Deficiency (AB Variant) – Ganglioside GM2 Activator

GM2A deficiency is an extremely rare form of GM2 gangliosidoses, with less than 30 cases reported in the literature (PMID: 35925454).

Function: The GM2 activator (*GM2A*) gene encodes a small, non-enzymatic, glycolipid transport protein, the GM2 activator protein, which acts as a substrate specific co-factor, working collectively with HexA to catalyze GM2 degradation, as well as the degradation of other molecules which contain terminal N-acetyl hexosamines. The GM2A protein binds to GM2 ganglioside to present it to HexA for its subsequent degradation.

Structure: *GM2A* is located on chromosome 5, locus 5q31.3-q33.1, and is a relatively small gene, <16 Kb in length (PMID: 10364519). On chromosome 3 there is a non-functional, processed *GM2A* pseudogene.

Genetics: GM2A deficiency (AB variant) is a very rare, monogenic, lysosomal storage disorder which results from pathogenic variants in the *GM2A* gene leading to a disfunction in the GM2 activator protein. This protein interacts with HexA, allowing it to facilitate the degradation of a group of glycosphingolipids called GM2 gangliosides, which is highly prevalent in neurons of the central nervous system. Pathogenic variants in the *GM2A* gene lead to a disfunctioning GM2A protein, causing GM2 gangliosides to accumulate within the lysosomes of neurons in the brain and spinal cord.

Inheritance: GM2A deficiency is inherited in an autosomal recessive pattern, indicating that two copies of the defective gene are necessary for disease development. Carriers of GM2A variants are typically asymptomatic.

Symptoms: GM2A deficiency (AB variant) is the rarest of the GM2 gangliosidoses and, to date, has primarily been documented in infants, with limited evidence demonstrating juvenile or late onset forms.

- The classic infantile form is the most common and the most severe, with symptoms appearing within the first few months of life followed by a rapid and fatal progression. Early signs include developmental delays, loss of motor skills, weakness, exaggerated

startle response, and a cherry-red spot on the macula. This disease is indistinguishable from Tay-Sachs disease on a phenotypic level.

Therapeutic Approaches: Currently there is no cure for GM2A deficiency, with available treatments targeting symptom management.

Prevalence:

- There are less than 30 reported cases and a likely incidence of <1 in every million live births (PMIDs: 34450229, 27402091, 26082327) from PMID: 37834060.
- There is a general lack of information due to the extremely limited number of cases reported. More information can be found at <https://www.ncbi.nlm.nih.gov/books/NBK583219/>

ASPA (Canavan Disease)- Aspartoacylase

Function: Aspartoacylase (ASPA) is a lysosomal enzyme responsible for catalyzing the conversion of N-acetylaspartic acid (NAA) to aspartate and acetate. NAA is the second most abundant amino acid derivative in the brain, localizing primarily in oligodendrocytes, and its metabolism plays a significant role in maintaining white matter and proper brain function. However, it's precise role has yet to be identified.

Structure: The *ASPA* gene is ~30 kb, has six exons intervened by five introns, and is located on chromosome 17p13.2 (PMID: 8088831). It directs the synthesis of aspartoacylase, a biologically active monomer, which acts as a zinc-dependent hydrolase, and is composed of 313 amino acids.

Canavan Disease

Genetics: Canavan disease is a rare, monogenic, lysosomal storage disorder which results from pathogenic variants in the *ASPA* gene. These variants impact the functionality of a vital lysosomal hydrolase, aspartoacylase, and lead to progressive brain atrophy and spongy degeneration of the white matter (<https://pubmed.ncbi.nlm.nih.gov/20301412/>).

Inheritance: Canavan disease is inherited in an autosomal recessive pattern, indicating that two copies of the defective gene are necessary for disease development. Carriers of *ASPA* variants are typically asymptomatic.

Symptoms: Canavan disease is classified into two categories based on the phenotypic spectrum, typical and atypical Canavan disease.

- o In typical Canavan disease (infantile), which accounts for about 85-90% of cases, neurological phenotypes become evident by about 3 to 5 months of age. This is followed by severe neurodegeneration and developmental decline along with various other

neurological-related symptoms including seizures, poor visual tracking, and reduced muscle tone.

- Atypical Canavan disease (juvenile) is less common, accounting for only about 10-15% of cases, with neurodevelopmental impairments becoming evident later (within the first few years of life). Similar to the typical form, this is followed by developmental regression, however it is much milder. The clinical course of the atypical form is broader compared to the typical form.

Genotype-Phenotype Correlations: There have been genotype-phenotype correlations proposed for various classes of variants. In the Ashkenazi Jewish population, the p.Tyr231Ter and p.Glu285Ala variants result in complete loss of enzymatic activity, which translates to the more severe phenotype of typical Canavan disease. Additionally, the p.Ala305Glu variant results in very low levels of ASPA activity and has been reported in both typical and atypical Canavan patients. Finally, there have been several variants that result in higher residual enzyme activity (p.Arg71His, p.Asp204His, p.Pro257Arg, p.Tyr288Cys) and therefore are associated with atypical Canavan disease (<https://www.ncbi.nlm.nih.gov/books/NBK1234/>).

However, genotype/phenotype correlation is unclear and reports are conflicting. Some studies have tried to characterize variants as severe or mild, but children sharing the same genotype may have different clinical courses (PMID: 9568915). Other studies have indicated that variants with highest residual enzyme activities (about 10% of normal) were found in patients with more mild clinical symptoms (<https://doi.org/10.1002/humu.23181> and PMID: 22850825 and PMID 28101991). "...However, for less severely affected patients of our cohort we observed molecular genetic findings such as p.Glu285Ala;p.Ala305Glu/p.Ala305Glu;p.Ala305Glu/p.Ala287Thr;p.Ala305Glu. These patients achieved and sustained a higher number of abilities and more complex functions. We also observed similar courses of the disease among siblings. Individual reports of ambulatory patients without macrocephaly suggest that milder phenotypes may be more abundant." PMID: 34011350.

Therapeutic Approaches: Currently there is no cure for Canavan disease, with available treatments being symptomatic and supportive.

Prevalence:

- Carrier Frequencies:
 - African: 1 in 741
 - Ashkenazi Jewish: 1 in 40 to 1 in 82 individuals
 - Finnish: 1 in 241 individuals
 - Caucasian: 1 in 486
 - Latino: 1 in 899
 - South Asian: 1 in 1923
 - World Wide: 1 in 393 individuals
- Prevalence:

- 1 in 100,000 births in general population (PMID: 35636725)
- Ashkenazi Jewish:
- Historically: 1 in 5,000 to 1 in 6,700 births in Ashkenazi Jewish population (PMID: 28101991)
- Now: Estimated to be 1 in 6,400 to 1 in 27,000 births with the implementation of widespread carrier screening programs and genetic counseling (PMID: 35636725)

Summary Table of Most Common Variants:

cDNA Name & Alias if applicable)	Protein Name	Allele Frequency	Clinical Significance ClinVar (2024)	VEP Consequence	Notes & Select References
c.914C>A	p.Ala305Glu	0.000301737	Pathogenic	missense	Common in Europeans of non-AJ ancestry, complete loss of ASPA activity, typical or atypical form, PMIDs: 7668285, 16437572, 28101991, https://www.ncbi.nlm.nih.gov/books/NBK1234/
c.854A>C	p.Glu285Ala	0.000220556	Pathogenic	missense	Founder in AJ population. Complete loss of ASPA activity, typical form, PMIDs: 28101991, 8037206, 8252036, 7485179, https://www.ncbi.nlm.nih.gov/books/NBK1234/
c.212G>A	p.Arg71His	0.000112149	Pathogenic/ Likely pathogenic	missense	Relatively higher residual ASPA activity, atypical form, PMIDs: 18070137, 16437572, https://www.ncbi.nlm.nih.gov/books/NBK1234/
c.693C>A	p.Tyr231Ter	4.15589E-05	Pathogenic	stop gained	Founder in AJ population, complete loss of ASPA activity, typical form, PMIDs: 7485179, 8023850, https://www.ncbi.nlm.nih.gov/books/NBK1234/
c.503G>A	p.Arg168His	3.28444E-05	Pathogenic/ Likely pathogenic	missense	PMIDs: 28101991, 10909858, 22750302

To access the literature shown in the last column, enter the PMID number into the PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) search box.

Update September, 2025: There is a newly discovered Canavan variant, a deep intronic SVA_E insertion. This variant may possibly be the most common pathogenic variant and could account for those individuals with clinical and biochemical diagnoses where only one or no pathogenic ASPA variants could be identified.

The intronic SVA_E insertion is described in 2 publications: PMIDs: 40257001 and 40995055. There is also an editorial highlighting the importance of the findings described in PMID: 40995055 (PMID: 40995054). These 2 publications describe 13 patients identified to date with this variant. Functional analysis showed that it results in aberrant splicing and transcript degradation because this insertion creates a novel splice acceptor site. The variant was not present in ClinVar or dbSNP. The overall allele frequency across all populations was 0.0005 in gnomAD v4.1. The HGVS nomenclature for this variant is seq[GRCh37] NC_000017.10:g.3393511_3393512insSVA_E.

GLB1 (GM1 Gangliosidosis) – Beta-Galactosidase

Function: The *GLB1* gene encodes the β -galactosidase enzyme that breaks down several substances, including GM1 ganglioside (a glycosphingolipid) and keratan sulfate (a large sugar molecule). Pathogenic *GLB1* variants can cause two phenotypically distinct lysosomal storage disorders, GM1 gangliosidosis and mucopolysaccharidosis type IVB (MPS IVB), also known as Morquio B disease.

β -galactosidase is a lysosomal hydrolase that cleaves β -linked galactose residues from the non-reducing end of glycan moieties found in various glycoconjugates. Reduction or absence of β -Galactosidase activity leads to the accumulation of GM1 ganglioside and its asialo derivative GA1, primarily in lysosomes of neuronal tissue, and severe progressive neurological disease (PMID: 34539759).

The *GLB1* gene encodes the elastin binding protein (EBP), in addition to the enzyme β -galactosidase, via alternative splicing. EBP is responsible for chaperoning the deposition of elastin fibers in the extracellular matrix (ECM). Although β -galactosidase is involved in multiple catabolic pathways, its two main substrates are GM1 ganglioside and keratan sulfate (PMID: 34539759).

Structure: The *GLB1* gene is a 62.5kb, 16-exon gene located on chromosome 3p21.33 (PMIDs: 7693577, 2511208).

GM1 Gangliosidosis

Genetics: GM1 gangliosidosis is a rare, monogenic, lysosomal storage disorder which results from pathogenic variants in the *GLB1* gene, impacting the functionality of a vital lysosomal hydrolase, beta-galactosidase 1. There have been more than 200 pathogenic or likely pathogenic variants reported in the literature (PMID: 34539759). Ultimately, the dysfunction in the *GLB1* gene leads to toxic accumulation of GM1 gangliosides within the lysosomes and drives progressive neurodegeneration.

Mucopolysaccharidosis type IVB (MPS IVB), also known as Morquio B disease, is another lysosomal storage disease that can also be caused by *GLB1* variants (typically caused by mutations in the 3' end of gene, PMID: 34539759). These patients have predominantly skeletal abnormalities without the neurodegenerative aspects of GM1 gangliosidosis (PMID: 29152458). The variants that cause MPS IV disrupt the breakdown of keratan sulfate by β -galactosidase and the breakdown of GM1 ganglioside is not affected by these variants. Because keratan sulfate is predominantly found in cartilage and the cornea, the buildup of this substance causes skeletal abnormalities and cloudy corneas (<https://medlineplus.gov/genetics/gene/glb1/>). Cases with phenotypic features of both GM1 gangliosidosis and Morquio B disease have also been described (PMID: 34539759).

Inheritance: GM1 gangliosidosis is inherited in an autosomal recessive pattern, indicating that two copies of the defective gene are necessary for disease development. Carriers of *GLB1* variants are typically asymptomatic.

Symptoms: GM1 gangliosidosis is classified into three categories based on age of onset.

- Type I (Infantile): The infantile form of GM1 gangliosidosis is the most severe form, which has an age of onset before 12 months of age and is characterized by a macular cherry-red spot and progressive neurological and neuromuscular dysfunction.
- Type II (Late-Infantile and Juvenile): Type II GM1 gangliosidosis can be divided into 2 subcategories; late-infantile, with an age of onset ranging between 1 and 3 years of age, and juvenile, which typically presents between 3 and 10 years of age. With this type, progressive central nervous system impairment is observed, with declines in motor and speech occurring.
- Type III (Chronic/Adult): This is the mildest form of GM1 gangliosidosis, typically not presenting until between late childhood and the third decade of life. This type presents with generalized dystonia and, as the disease progresses, cognitive impairment, with some patients also presenting with akinetic-rigid parkinsonism. There has been a direct correlation demonstrated between disease prognosis and the degree of neurological decline (<https://www.ncbi.nlm.nih.gov/books/NBK164500/>).

Genotype-Phenotype Correlations: No clear correlations have been reported due to the high molecular heterogeneity of the disease. Some literature suggests that pathogenic variants that impact the surface of beta-galactosidase are typically associated with type III GM1, meaning individuals presents with a milder phenotype. On the other hand, variants that involve the enzyme active site or the protein core typically invoke a more severe phenotype, like what is seen with type I (<https://www.ncbi.nlm.nih.gov/books/NBK164500/>).

Therapeutic Approaches: Currently there is no cure for GM1 gangliosidosis, with available treatments being symptomatic and supportive. However, there are various ongoing trials exploring the therapeutic potential of substrate reduction therapy, enzyme enhancement therapy, enzyme replacement therapy, and gene therapy for this monogenic disease.

Prevalence:

- Carrier Frequencies:
 - African: 1 in 356
 - East Asian: 1 in 305
 - Finnish: 1 in 246
 - Caucasian: 1 in 277
 - Latino: 1 in 431
 - South Asian: 1 in 285
 - Roma: 1 in 50
 - South Brazilian: 1 in 58
 - World Wide: 1 in 297 individuals

- Prevalence:
 - The estimated prevalence of GM1 gangliosidosis is 1 in 100,000 to 1 in 300,000 worldwide (PMID: 36541412, 30094186). Higher prevalence estimates in Brazil, Roma ancestry, and Maltese islands are likely representative of founder effects (PMID: 18524657). Chronic/adult GM1 gangliosidosis is higher in the Japanese population (<https://www.ncbi.nlm.nih.gov/books/NBK164500/>).
 - The prevalence of MPS IVB has been reported as 1 in 250,000 to 1 in 1,000,000 live births (PMID: 30094186).

Summary Table of Most Common Variants:

cDNA Name & Alias (if applicable)	Protein Name	Allele Frequency	Clinical Significance ClinVar (2024)	VEP Consequence	Notes & Select References
c.75+2dupT		0.00014954	Pathogenic	splice region	PMIDs: 8198123, 8199591, 29160035, 21497194, 25936995, 36629845, 34539759
c.1733A>G	p.Lys578Arg	0.000113378	Pathogenic/ Likely pathogenic	missense	PMIDs: 8213816, 21497194, 25557439, 34539759
c.442C>A	p.Arg148Ser	8.17857E-05	Pathogenic/ Likely pathogenic	missense	PMIDs: 10839995, 34539759, 10841810
c.442C>T	p.Arg148Cys	7.12579E-05	Pathogenic/ Likely pathogenic	missense	PMIDs: 17221873, 23151865, 21497194

To access the literature shown in the last column, enter the PMID number into the PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) search box.

Other Variants of Note:

cDNA Name & Alias (if applicable)	Protein Name	Notes & Select References
c.1577dupG (1622_1627insG)	p.Trp527LeufsTer5	High prevalence in Brazilian population, associated with type I GM1 (PMID: 21637542, https://www.ncbi.nlm.nih.gov/books/NBK1234/)
c.176G>A	p.Arg59His	High prevalence in Roma & Brazilian populations, associated with type I /type II (juvenile) GM1 (PMID: 21637542, https://www.ncbi.nlm.nih.gov/books/NBK1234/)
c.495_497delTCT	p.Leu166del	High prevalence in Chinese population and is associated with type I / type II (late-infantile) GM1 (PMID: 30267299, https://www.ncbi.nlm.nih.gov/books/NBK1234/)
c.152T>C	p.Ile15Thr	High prevalence in Japan; associated w/GM1 adult form (https://www.ncbi.nlm.nih.gov/books/NBK1234/)
c.1343A>T	p.Asp448Val	The most common pathogenic variant in persons of Korean ancestry with type II (late-infantile) GM1, also reported in persons of Chinese & Turkish ancestry with type I GM1 (PMID: 29439846, https://www.ncbi.nlm.nih.gov/books/NBK1234/)
c.817_818delTGinsCT (851-852TG>CT)	p.Trp273Leu	Most frequent variant in MPS IVB, causes skeletal dysostosis without neurologic involvement (PMID: 33266180, https://www.ncbi.nlm.nih.gov/books/NBK1234/)

To access the literature shown in the last column, enter the PMID number into the PubMed (<https://pubmed.ncbi.nlm.nih.gov/>) search box.

References

Carrier frequencies: https://www.jewishgeneticdiseases.org/wp-content/uploads/Sema4_Carrier-Frequencies-1.pdf, unless otherwise stated.

ClinVar: <https://www.ncbi.nlm.nih.gov/clinvar/>

GeniE: the Genetic Prevalence Estimator: <https://genie.broadinstitute.org/>

GnomAD: <https://gnomad.broadinstitute.org/>

PubMed: <https://pubmed.ncbi.nlm.nih.gov/>

Appendix

Selections made when running a variant list on GeniE (<https://genie.broadinstitute.org/>):

Notes

gnomAD version *

4.1.0 (GRCh38)

2.1.1 (GRCh37)

[What's the difference between gnomAD v2 and v4?](#)

Gene *

PCSK9

Transcript *

Select a gene to see transcripts for that gene.

Include variants based on clinical significance in ClinVar? *

Include pathogenic and likely pathogenic variants

Include pathogenic and likely pathogenic variants and variants with conflicting interpretations of pathogenicity *

Do not include variants based on clinical significance

* Where at least one of the conflicting classifications is pathogenic or likely pathogenic

Include gnomAD variants based on variant type? *

High Confidence Loss of Function only

High Confidence Loss of Function and Missense Variants with strong REVEL score ($\geq .932$)