

The genetic basis of classic nonketotic hyperglycinemia due to mutations in *GLDC* and *AMT*

Curtis R. Coughlin II, MS, MBe¹, Michael A. Swanson, PhD¹, Kathryn Kronquist, PhD^{1,2}, Cécile Acquaviva, PharmD, PhD³, Tim Hutchin, BSc, PhD⁴, Pilar Rodríguez-Pombo, PhD⁵, Marja-Leena Väisänen, PhD⁶, Elaine Spector, PhD^{1,2}, Geralyn Creadon-Swindell, BS^{1,2}, Ana M. Brás-Goldberg, BSc, PhD⁴, Elisa Rahikkala, MD, PhD⁷, Jukka S. Moilanen, MD, PhD⁷, Vincent Mahieu, MS⁸, Gert Matthijs, PhD⁸, Irene Bravo-Alonso, MS⁵, Celia Pérez-Cerdá, PhD⁵, Magdalena Ugarte, PhD⁵, Christine Vianey-Saban, PharmD, PhD³, Gunter H. Scharer, MD, PhD^{1,9} and Johan L.K. Van Hove, MD, PhD¹

Purpose: The study's purpose was to delineate the genetic mutations that cause classic nonketotic hyperglycinemia (NKH).

Methods: Genetic results, parental phase, ethnic origin, and gender data were collected from subjects suspected to have classic NKH. Mutations were compared with those in the existing literature and to the population frequency from the Exome Aggregation Consortium (ExAC) database.

Results: In 578 families, genetic analyses identified 410 unique mutations, including 246 novel mutations. 80% of subjects had mutations in *GLDC*. Missense mutations were noted in 52% of all *GLDC* alleles, most private. Missense mutations were 1.5 times as likely to be pathogenic in the carboxy terminal of *GLDC* than in the amino-terminal part. Intragenic copy-number variations (CNVs) in *GLDC* were noted in 140 subjects, with biallelic CNVs present in 39 subjects.

The position and frequency of the breakpoint for CNVs correlated with intron size and presence of Alu elements. Missense mutations, most often recurring, were the most common type of disease-causing mutation in *AMT*. Sequencing and CNV analysis identified biallelic pathogenic mutations in 98% of subjects. Based on genotype, 15% of subjects had an attenuated phenotype. The frequency of NKH is estimated at 1:76,000.

Conclusion: The 484 unique mutations now known in classic NKH provide a valuable overview for the development of genotype-based therapies.

Genet Med advance online publication 30 June 2016

Key Words: *AMT*; genotype; *GLDC*; glycine cleavage enzyme; nonketotic hyperglycinemia

INTRODUCTION

Nonketotic hyperglycinemia (NKH; OMIM 605899) is an autosomal recessive disorder of glycine metabolism that results from deficient activity of the glycine cleavage enzyme system (GCS). Most patients present in the first week of life with lethargy, hypotonia, and myoclonic jerks, and often progress to apnea requiring ventilator support.¹ The NKH phenotype is heterogeneous; although both severe and attenuated NKH can present neonatally, the neurodevelopmental outcomes can differ.¹ The severe form of NKH is characterized by profound developmental delay, spasticity, and intractable seizures. An attenuated phenotype is characterized by variable developmental progress, hyperactivity, chorea, intermittent ataxia and lethargy, and behavioral problems.^{2,3}

The GCS consists of glycine decarboxylase (P-protein), aminomethyltransferase (T-protein), hydrogen carrier protein (H-protein), and dihydrolipoamide dehydrogenase (L-protein). The P, T, and H proteins are encoded by *GLDC* (OMIM 238300), *AMT* (OMIM 238310), and *GCSH* (OMIM 238330) genes, respectively. Classic NKH is caused by defects in the genes that encode protein components of the GCS, and mutations have been reported in *GLDC* and *AMT*. In approximately 4% of patients with deficient GCS enzyme activity, no mutations are identified in *GLDC* or *AMT*; these cases are caused by defects in the synthesis of the cofactor lipoic acid.⁴ In classic NKH, a genotype–phenotype correlation has been established as mutations associated with residual enzyme activity result in an attenuated phenotype.^{5,6}

¹Section of Genetics, Department of Pediatrics, University of Colorado School of Medicine, Aurora, Colorado, USA; ²Molecular Genetics Laboratory, Department of Pathology and Laboratory Medicine, Children's Hospital Colorado, Aurora, Colorado, USA; ³Service Maladies Héritaires du Métabolisme, Centre de Biologie Est, CHU de Lyon, Lyon, France; ⁴Newborn Screening and Biochemical Genetics, Birmingham Children's Hospital, Birmingham, UK; ⁵Centro de Diagnóstico de Enfermedades Moleculares (CEDEM), CBM-SO, (UAM-CISC), Centro de Investigación en Red de Enfermedades Raras (CIBERER), IDIPAZ, Universidad Autónoma Madrid, Madrid, Spain; ⁶Research Group of Cancer Research and Translational Medicine, University of Oulu, Oulu, and Northern Finland Laboratory Centre Nordlab, Oulu, Finland; ⁷PEDEGO Research Unit, Medical Research Center, Department of Clinical Genetics, Oulu University Hospital and University of Oulu, Oulu, Finland; ⁸Center for Human Genetics, University of Leuven, Leuven, Belgium; ⁹Section of Medical Genetics and Genomics, Children's Hospitals and Clinics of Minnesota, Minneapolis, Minnesota, USA. Correspondence: Johan Van Hove (Johan.Vanhove@childrenscolorado.org)

Submitted 17 February 2016; accepted 25 April 2016; advance online publication 30 June 2016. doi:10.1038/gim.2016.74

The 113.15-kb *GLDC* gene on 9p24.1 contains a transcript of 3,767 base pairs encoding a protein of 1,021 amino acids.^{7–10} Most patients with classic NKH have biallelic mutations in *GLDC*.¹¹ Reported recurrent missense mutations include the p.R515S mutation observed in individuals of Caucasian ancestry¹² and the p.S564I mutation noted in the Finnish population.¹³ Most missense mutations in *GLDC* are private.¹⁴ Intragenic copy-number variants (CNVs), most commonly genomic deletions, have been reported in *GLDC*.¹⁵ The 5.97-kb *AMT* gene on 3p21.31 contains a transcript of 2,276 base pairs and encodes a protein of 403 amino acids.¹⁶ A recurrent mutation identified in *AMT* is the p.R320H mutation.¹²

To provide a comprehensive overview of the genetic basis of classic NKH, we compiled genetic results from clinical and research laboratories that provide diagnostic testing in patients suspected to have NKH. We report 410 mutations, including 246 novel mutations, in 578 unique families with identified mutations in *GLDC* or *AMT*. The frequency of mutations identified in this cohort was consistent with previous reports of the incidence of NKH and the incidence of attenuated NKH and had implications for the understanding of the genomic architecture and the protein structure.

MATERIALS AND METHODS

Subjects

Subjects were identified through clinical or research laboratories that provide targeted genetic testing for classic NKH. Data were retrospectively collected from five international centers. Ethical approval was obtained from each of the following sites: University of Colorado (COMIRB 05-0790, 15-0832), Centre de Biologie et Pathologie Est (IRB 0009118), Birmingham Children's Hospital (IRAS 182766), Universidad Autónoma Madrid (CEI 67-1192), and Oulu University Hospital (Regional Ethics Committee of the Northern Ostrobothnia Hospital District).

The inclusion criteria included any subject with a clinical suspicion of NKH, for which samples were sent to the DNA laboratory for diagnostic confirmation, and who was identified to have mutations in either *GLDC* or *AMT*. No patient had a mutation identified in *GCSH*. A family with two affected siblings was listed with a single identifier. Ethnic origin, gender, genetic testing results, and parental phase of results were recorded for each subject. Duplicate recording of individuals was excluded via identification of similar mutations and limited identifying information. Associations of specific mutations with ethnicity or gender were evaluated using a chi-squared test, incidence differences were compared using a proportion Z-test, and associations were evaluated with Pearson correlation and linear regression. A significant result was defined as $P \leq 0.01$.

Molecular genetics review

The *GLDC* and *AMT* genes were evaluated by Sanger sequencing at each individual laboratory; primers are available on request. Exonic CNVs were evaluated by multiplex ligation-dependent

probe amplification¹⁵ or custom array-based comparative genomic hybridization as previously described.⁵ Common Finnish missense mutations were analyzed by restriction enzyme digestion and agarose gel analysis.¹³ Sequence variants were compared with the sequences of *GLDC* (GenBank NM_000170.2, NP_000161.2, Ensembl ENSG00000178445) and *AMT* (GenBank NM_000481.2, NP_000472.2, Ensembl ENSG00000145020). Discrepancies in recorded mutations were reviewed by two investigators (C.R.C., J.V.H.) and verified by the appropriate laboratory. To estimate the general population frequency of mutations, we recorded for each mutation its frequency in the Exome Aggregation Consortium (ExAC) browser, which contains sequencing data for 60,706 individuals from disease-specific and population-based genomic studies.^{17,18} A minimum population frequency was estimated using the Hardy-Weinberg equation from the carrier rate in the ExAC browser of missense mutations identified in this study and in the literature. All nonsense mutations in ExAC were assumed to be pathogenic, and the overall carrier frequency in *GLDC* was corrected for the CNV rate. Mutated amino acids were mapped on the homology model built on the published structure from *Synechocystis* sp. as described.^{5,19} The residual activity of select missense mutations in *GLDC* was measured after expression in COS cells as described.⁵ The impact of the synonymous mutation c.921G>A, p.V307V on splicing was evaluated by sequencing cDNA made from mRNA extracted from Epstein-Barr virus-transformed lymphoblasts.

Literature review

A systematic literature search was independently performed by two investigators (C.R.C., J.V.H.) with the search criteria of NKH, glycine encephalopathy, *GLDC*, *AMT*, and glycine cleavage enzyme system. Publications that contained molecular data for patients with classic NKH were retained and systematically classified.

RESULTS

Subjects

Molecular genetic data for 578 unique families representing 302 males and 280 females identified a total of 1,130 mutations, including 410 unique mutations, of which 246 mutations are reported for the first time (**Supplementary Table S1** online). The majority of subjects (80%) had mutations identified in *GLDC* (464, of which 227 were female, 233 male, and 4 male-female sibling pairs), and the remaining 20% were in *AMT* (114, of which 49 were female and 65 male). For 17 subjects, testing was incomplete. For 8 subjects, only a common mutation panel or CNV analysis was performed, identifying only a single disease-causing mutation. In 7 subjects, a single mutation was identified through sequencing without CNV analysis. In two families without a proband sample, only one parental sample was available for analysis. In the other 561 families, sequencing and CNV analysis identified biallelic mutations in 554 cases, resulting in a 98.7% detection rate of pathogenic mutations.

GLDC mutations

In *GLDC*, there were 52% missense mutations, 11% nonsense mutations, 10% splice-site mutations, 6% small insertions/deletions (InDels), and 21% exonic CNVs (Table 1, Supplementary Table S2 online). Of a total of 172 unique missense mutations, 123 mutations (71%) were identified in only a single individual (Supplementary Table S2a online). Although the majority of mutations were private, several recurrent mutations were noted. Seventy percent of subjects were heterozygous for at least one missense mutation (Table 1), and the p.R515S mutation

Table 1 Genotype of all subjects reported by mutation type
Genotype of 464 subjects with mutations in *GLDC*

	Missense	CNV	In-frame stop	Splice	Indel
Missense	145				
CNV	77	40			
In-frame stop	37	7	22		
Splice	33	13	7	13	
Indel	18	8	4	10	7
Incomplete	11	3	1	1	1
Unknown	3	0	0	2	1
Total ^a	324	148	78	79	49

Genotype of 114 subjects with mutations in *AMT*

	Missense	Indel	Splice	In-frame stop	UTR
Missense	62				
Indel	16	12			
Splice	5	0	4		
In-frame stop	4	1	2	1	
UTR	3	0	1	0	1
Synonymous	1	0	0	0	0
Unknown	1	1	0	0	0
Total ^a	92	30	12	8	4

^aThe total number represents the number of subjects who were at least heterozygous for the type of mutation reported.

CNV, copy-number variations; UTR, untranslated region.

was the most common missense mutation identified. The p.R515S mutation was present in 6% of all *GLDC* alleles, and 50 (11%) patients had at least one p.R515S mutation. Recurrent missense mutations identified in >1% of all *GLDC* mutated alleles include the p.S564I (39 alleles), p.G761R (22 alleles), p.A389V (14 alleles), p.A802V (13 alleles), p.R790W (12 alleles), p.R905G (11 alleles), p.G771R (10 alleles), and p.R461Q (10 alleles) mutations. Pathogenic missense mutations were noted throughout *GLDC*, with enrichment in the 3' located exons. Pathogenic missense mutations were recorded as affecting more than 25% of the amino acids within exons 14, 16, and 24, as opposed to <5% of amino acids associated with a pathogenic mutation in exons 1, 2, 5, and 12 (Figure 1a). Overall, the rate at which amino acids were present in recorded pathogenic mutations for all the amino acids of each exon in the amino-terminal part of the gene (exons 1–12) was significantly lower at 11% (59/526) than in the carboxy-terminal part (exons 13–25 at 18%; 88/495; $P = 0.0024$). Thus, for each amino acid in the carboxy-terminal part, the likelihood of it being involved in a recorded pathogenic mutation in a patient is 1.5 times more likely than it is for an amino acid in the N-terminal part. The carboxy terminal of the mature *GLDC* protein is located around the active site, whereas the amino-terminal part is located in an upper region away from the active site, except for exon 8, which forms an upper border of the active site and has a higher pathogenic mutation rate (Figure 1b). No pathogenic missense mutations were found in exon 2, which is located away from the active site and outside the dimer interface region and away from the upper structure where mutants can be associated with protein instability. This structural basis may explain why missense mutations in exon 2 are benign and tend to be pathogenic in exon 8.

In 11% of alleles, single-nucleotide variants resulting in in-frame premature stop mutations were present, comprising 39 different nonsense mutations (Supplementary Table S2b online). Recurrent premature stop mutations identified in >1%

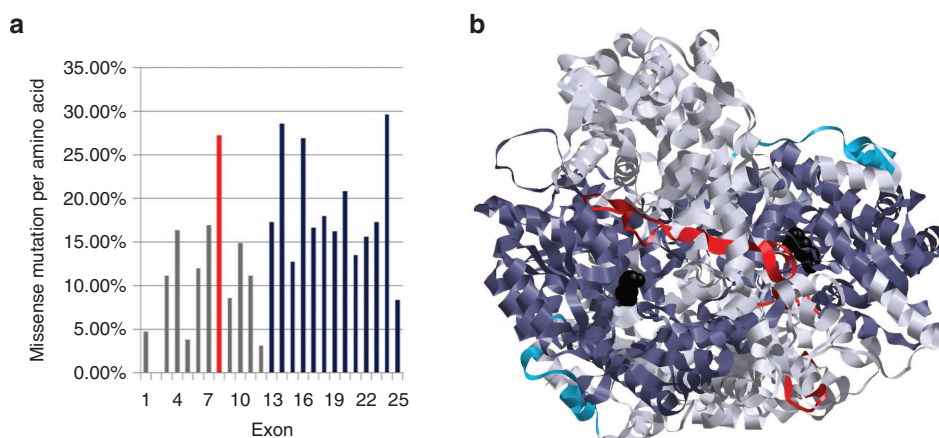


Figure 1 Missense mutations and location on the P-protein crystal structure. For each exon in *GLDC*, the percentage of amino acids identified as pathogenic missense mutation is shown (a). On the modeled crystal structure, the amino acids of the carboxy-terminal part of the protein are shown in dark blue and the amino acids of the amino-terminal part of the protein are shown in gray, except for the amino acids of exon 8, which are shown in red. The amino acids in exon 2 are shown in light blue. The active site pyridoxal-phosphate is shown in black (b).

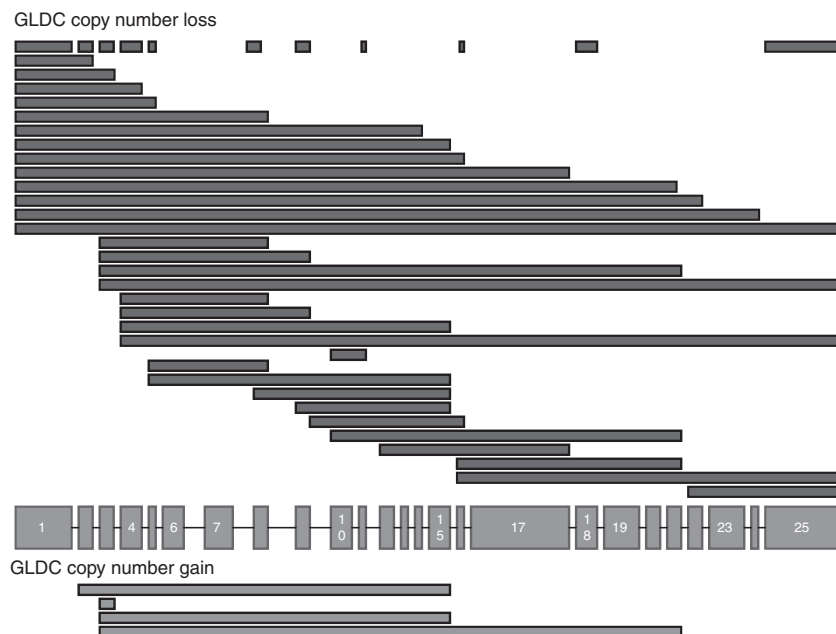


Figure 2 Copy-number variants identified in *GLDC*. The extent of the copy-number variants is shown on the genomic structure of the *GLDC* gene. Copy-number losses range from single-exon deletions (top row) to multiexon deletions and duplications (intron size not to scale).

of all *GLDC* mutated alleles include the p.R337X (15 alleles), p.R424X (13 alleles), and p.E167X (11 alleles) mutation. A similar percentage (10%) of splice-site mutations was noted, comprising 37 different splice-site mutations, including several recurrent splice-site mutations, such as c.2316-1G>A (IVS19-1G>A), c.2665+1G>C (IVS22+1G>C), and c.2919+1G>A (IVS24+1G>A), each noted in >1% of all *GLDC* alleles (**Supplementary Table S2c** online). Mutations affected each exon–intron boundary, with the exception of introns 5, 8, 9, 11, 14, and 20. Small insertion and deletion (InDel) mutations were identified in 7% of all *GLDC* mutations comprising 42 different mutations (**Supplementary Table S2d** online). As expected, the majority of the InDel mutations were private mutations, with the few recurrent mutations being homozygous in a given individual.

Intragenic CNVs were noted in 21% of all *GLDC* alleles, with deletions (181) of exons within the *GLDC* gene more frequent than duplications (7) (**Supplementary Table S2e** online, **Figure 2**). The length of the CNVs varied from a single exon to encompassing the entire gene. Biallelic CNVs were present in 39 subjects with compound heterozygous CNVs in 10 subjects, including 4 subjects with nonoverlapping CNVs (**Table 1**). The majority of CNVs were located in the 5' end of the gene, with the first exon involved in 117 CNVs, exon 2 in 104 CNVs, exon 3 in 77 CNVs, exon 4 in 78 CNVs, and exon 5 in 73 CNVs. By contrast, exons 24 and 25 were involved in only 6 and 7 CNVs, respectively. The distribution of where the deletion break occurred (at either the start or the end of the deletion) varied remarkably. Most frequent breaks occurred before exon 1 (117 alleles, 62% of all deletion alleles) and in intron 2 (67 alleles), intron 15 (32 alleles), and introns 3, 8, and 4 (in 18, 27, and 17 alleles), whereas no breaks occurred at all in introns 6,

12, 13, 19, and 20 (**Figure 3a**). A correlation exists between the number of breaks and the interrelated variables of intron size ($r^2 = 0.786$, $P < 0.001$) and the number of Alu repeat elements within each intron ($r^2 = 0.781$, $P < 0.001$) (**Figure 3b**).

AMT mutations

The majority of mutations were single-nucleotide variants resulting in missense mutations (67%), premature stop mutations (3%), or splice-site mutations (8.4%) (**Table 1**, **Supplementary Table S3** online). In contrast to *GLDC*, the majority of missense mutations in *AMT* are recurrent as all but nine mutations are identified in multiple subjects (**Supplementary Table S3a** online). The p.R320H mutation was the most common missense mutation identified in 16% of all *AMT* alleles, and 27 (24%) patients had at least one p.R320H mutation. The p.R222C (12 alleles), p.R94W (9 alleles), p.R73C (8 alleles), p.R296H (7 alleles), and p.M1T (7 alleles) were present in >3% of all *AMT* alleles. As opposed to *GLDC*, only moderate differences were seen between different exons regarding the relative frequency with which amino acids appear as mutated into pathogenic missense mutations in *AMT*. Approximately 8–12% of the amino acids in exons 2, 3, 5, and 6 were associated with a pathogenic missense mutation.

Six single-nucleotide variants resulted in a premature stop mutation (**Supplementary Table S3b** online). Recurring splice-site mutations included c.471+2T>C (IVS4+2T>C) and c.878-1G>A (IVS7-1G>A) (**Supplementary Table S3c** online). The silent mutation c.921G>A, p.V307V created a new donor splice site, as evidenced by the presence of the mRNA sequence that spliced into exon 9 at c.1036, resulting in a frameshift and early stop codon. InDel mutations resulting in complex rearrangements were present in 20% of *AMT* alleles (**Supplementary**

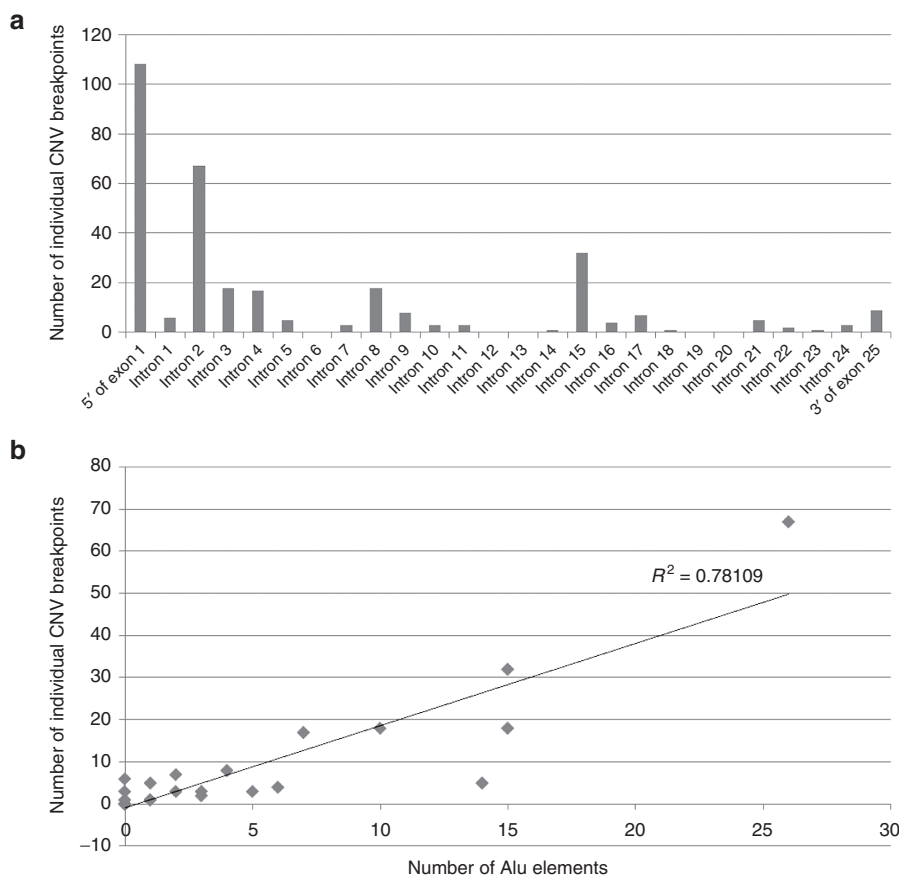


Figure 3 Distribution of break points of intragenic deletions within *GLDC*. The number of CNV boundaries within each intron of an intragenic deletion (a). A close correlation is shown between the number of alleles with a break point for a deletion and the number of intronic Alu repeat elements (b).

Table S3d online); the majority of private mutations with most of the few recurrences were contributed by homozygous individuals. As opposed to *GLDC*, no intragenic CNVs were noted. Interestingly, five variants were identified in the 5' untranslated region clustered within a span of 11 nucleotides, and the frequency was either at a minor allele frequency <0.0001 or unavailable in the ExAC browser (Supplementary Table S3e online).

Ethnic variation

Specific founder mutations were identified in this cohort. The frequency of the p.R515S mutation in *GLDC* in this study (6%) was similar to previous general population estimates of 5%.¹² However, this mutation was significantly enriched in subjects identified from the United Kingdom ($P \leq 10^{-13}$); 38% of all subjects from the United Kingdom had at least one p.R515S mutation. Consequently, subjects identified from the United Kingdom were more likely to have mutations in *GLDC* (93%) ($P = 0.003$). The p.I106T mutation in *AMT* was identified only in subjects from The Netherlands, where at least one allele containing the p.I106T mutation was identified in every subject with mutations in the *AMT* gene. The p.S132L mutation was most commonly identified in Pacific people; six of seven subjects had at least one copy of the p.S132L mutation, whereas

in our cohort it was identified in only one other subject. Three mutations in *GLDC* represented 91% of disease-causing alleles in subjects originating from Finland. The p.S564I (53% of Finnish alleles) was identified only in subjects from Finland, whereas the p.G761R (26% of Finnish alleles) was identified in three other subjects. A recurrent deletion of exons 1–8 (12% of Finnish alleles) was also noted. There were no Finnish subjects with mutations in *AMT*.

Literature review

We retrieved 239 unique mutations in *GLDC* and *AMT* reported in the literature, including large intragenic CNVs (Supplementary Table S4 online). The frequency of these mutations was not recorded because the authors could not ensure that a single patient was unique to each report. Three reported sequence variants c.1705G>A (p.A569T), c.1229G>A (p.R410K), and c.2113G>A (p.V705M) had been reported as pathogenic mutations^{11,20,21} but were noted at high allele frequency in the ExAC database at, respectively, 474/121,400, 346/121,404, and 443/121,392 alleles, and they were even present in homozygous forms in ExAC. In our cohort, the asymptomatic mother of an affected child was homozygous for p.R410K, and three patients with two other known pathogenic mutations carried one or two copies of p.V705M. The amino acid R410 is located away from

the active site on the outer margin of the protein, and the mutation does not confer a charge change. The V705 amino acid is located at the lateral side of the active cleft but not at the surface of the protein, where it could interfere with binding of the H-protein. Expression studies of these mutations showed $78\% \pm 16.4$, $77\% \pm 1.5$, and $191\% \pm 24.7$ residual activity for mutations p.R410K, p.A569T, and p.V705M, respectively. Taken together, these data indicate that these mutations are not pathogenic.

Population frequency

We utilized this comprehensive mutation review and literature review to evaluate the frequency of pathogenic mutations in the ExAC browser. The allele frequency of missense mutations and splice-site mutations as well as all frameshift and premature stop mutations reported in the ExAC browser suggest a carrier frequency of 0.50% (1/200) for classic NKH. In the reported cohort, 17% of all disease-causing alleles were large CNVs, which would not be identified in ExAC data. Assuming a similar rate of CNVs, the carrier rate for *GLDC* would be 1/171 (0.58%) and for *AMT* it would be 1/660 (0.15%). Combined, the carrier rate for classic NKH would be 1/138, giving a minimal incidence of classic NKH at approximately 1:76,000 live births based on ExAC data alone. The incidence of NKH will be increased in geographic areas with a founder mutation.

DISCUSSION

We report a comprehensive review of mutations in *GLDC* and *AMT* identified in patients with classic NKH. We attempted to capture the majority of potential subjects through contributed data from five international centers. Using this complete set of mutations, previously published mutations and the allele frequency in the ExAC data set enabled us to calculate an estimate of the prevalence of carrier pathogenic mutations in this population. This is an unbiased minimum disease incidence of 1:76,000, which relates surprisingly well with a previous estimate of 1:63,500 (ref. 22).

Missense mutations were the most common type of mutation identified in both *GLDC* and *AMT*. In *GLDC*, most missense mutations were identified in a single individual, which limits the benefit of a common mutation testing, except in populations in which founder mutations have been identified, such as the p.R515S mutation in the United Kingdom, the p.S564I and p.G761R mutations and the deletion of exons 1–8 in Finland, and p.S132L in Pacific people. Recurrent missense mutations were more common in *AMT*, with the p.R222C and p.R320H accounting for 20% of all disease alleles. Other founder mutations may also be present in other specific ethnic groups not represented in this study, such as the c.2T>C, p.M1T in Palestinians in Jerusalem.²³ General population sequencing data provided valuable information in the interpretation of suspected disease-causing variants. Reviewing the frequency of all missense and splice mutations in the ExAC browser confirmed our suspected reclassification of three variants as benign, which was further validated by documentation of substantial residual activity, in contrast to previous suggestions.^{11,20,24}

Protein structure

In *GLDC*, missense mutations were enriched in exons that encoded the carboxy-terminal part of the mature protein over the amino-terminal part. Previous studies had noted increased frequency of mutations in exon 19, which contains the binding site for pyridoxal phosphate.¹¹ The mutation rate per amino acid was lower in the amino-terminal part of the protein than in the carboxy-terminal part. Whereas in eukaryotes and some prokaryotes the P-protein is a homodimeric form α_2^{NC} , in many prokaryotes the protein is a heterotetrameric form of $\alpha_2^{\text{N}}\beta_2^{\text{C}}$ corresponding to the N-terminal and C-terminal halves, with pyridoxal-phosphate binding localized to the C-terminal half.²⁵ The role of the N-terminal half is less clear, and the lower frequency of amino acids involved in pathogenic mutations implies less stringent functionality, except for the amino acids in exon 8, which appeared to be part of the active site.

Genomic architecture

Intragenic CNVs were noted in 21% of *GLDC* alleles, consistent with a previous report.¹⁵ Although CNVs were noted throughout *GLDC*, specific areas were enriched with CNV events. Correlating break point frequency with intron size and, particularly, the number of Alu elements, the distribution of the CNVs is probably due to the genomic architecture of *GLDC*. Previous breakpoint analysis had identified Alu-mediated recombination in three of five cases of intragenic deletions in *GLDC*.¹⁵ Intragenic CNVs have never been reported in *AMT*, and the genomic architecture of *AMT* involves small introns with only three Alu elements, in contrast to the 115 Alu elements noted between intron 1 and intron 24 of *GLDC*. Large contiguous gene deletions encompassing *AMT* have only rarely been reported.²⁶

Of specific note, five patients were identified as having variants in the 5' untranslated region within nucleotides c.1–55 to c.1–66, which are very rare in the ExAC database supporting the deleterious nature. These data suggest the identification of a regulatory region for *AMT*. Further functional studies are required to confirm the pathogenic nature of these variants. The molecular basis for the regulation of expression of the GCS components has not yet been studied.

Genotype–phenotype correlations

Missense mutations associated with residual enzyme activity have been associated with an attenuated phenotype. Specifically, the mutations p.A202V, p.A389V, p.A802V/E, p.T269M, p.L548V, and p.G607S in *GLDC* and p.I106T in *AMT* have been identified in patients with attenuated NKH who made relatively improved developmental progress.⁵ Forty-eight subjects were identified as having at least one of these seven mutations, representing 8.3% of the cohort. In a previous study, these seven mutations were present in 57% of patients with a documented attenuated phenotype,⁵ suggesting that 14.5% of subjects in this cohort would have the potential for an attenuated phenotype. Previous studies suggested that 20% of patients with NKH have attenuated NKH.^{2,3} These studies were limited to natural-history

questionnaires, which may have overestimated the incidence of attenuated NKH because life expectancy is expected to be longer in this cohort of patients. The present study is relatively unbiased because genetic testing is performed for both diagnostic confirmation and reproductive counseling regardless of the patient's clinical status. Although the list of expressed mutations with determination of residual enzyme activity is steadily growing, the very large number of missense mutations will make it unrealistic to have a truly comprehensive data set. Biochemical and clinical indicators, including *in vivo* biochemical testing of residual activity such as the ¹³C-glycine breath test, will remain important tools for early recognition of residual enzyme activity and the potential for attenuated phenotype.²⁷

Diagnostic relevance

The paradigm for the diagnosis of genetic disorders has changed to a genome-first approach. Clinical laboratories are increasingly reliant on various levels of evidence before reporting that a mutation is pathogenic.²⁸ This can be problematic for clinicians and families who are often making clinical decisions based on the laboratory report and certainty of diagnosis. The reported data are generated from patients with a clinical suspicion of classic NKH. Classic NKH has a characteristic biochemical finding of elevated serum and CSF glycine that enables confirmation of the clinical diagnosis. This report increases the number of reported mutations identified in patients by 50%, with 484 unique mutations now reported in *GLDC* and *AMT*. These mutations will be available at <http://www.lovd.nl/>, facilitating assignment of pathogenicity to sequence variants. Particular caution must be taken with respect to silent missense mutations that affect splicing, particularly in *AMT*: c.339G>A p.Q113Q, which affects the last nucleotide of exon 3, causing missplicing of intron 3, and c.921G>A p.V307V, which creates a new donor splice site.

Complete testing for *GLDC* and *AMT*, including CNV analysis, is essential to evaluate classic NKH; 101 subjects had a single CNV that traditional sequencing does not identify. If CNV analysis had not been performed, these patients might have been mistakenly identified as carriers of NKH. Despite reports to the contrary,²⁹ none of the heterozygous carrier parents known to the investigators have symptoms of NKH, and all symptomatic patients with a single mutation identified on sequence were later identified as having a CNV. The four subjects with nonoverlapping CNVs and normal sequencing of *GLDC* might have gone undiagnosed or misclassified as variant NKH if CNV analysis had not been performed (Table 1). Parental testing in the setting of a homozygous mutation may reveal the need for CNV analysis. Parental testing is also recommended to evaluate the possibility of a *de novo* mutation. In two subjects, *de novo* mutations were identified, with a p.G135A mutation identified on a maternal allele and a p.G771R mutation on a paternal allele in *GLDC*, respectively. Although the *de novo* rate in this cohort is <1%, recognition of this important mechanism significantly reduces a couple's recurrence risk. Gonadal mosaicism was not identified or evaluated in our cohort.

Clinical relevance

Genetic testing is playing an increasing role in medical decision making, especially as genotype-phenotype correlations increase and molecular-based therapies are evaluated. This is especially important in the neonatal period, when decisions regarding medical interventions are often made.³⁰ An attenuated phenotype of NKH can be predicted when a mutation with residual enzyme activity is identified. A mutation with residual activity is necessary but not sufficient for an attenuated phenotype.⁵ Early treatment must be initiated to achieve the best developmental outcome.³¹

Future prospects

The pathophysiology of NKH is not well defined, and therapeutic strategies have not changed since introduction of *N*-methyl-D-aspartate-receptor antagonist therapy in 1992.³² The genetic basis has now been well defined and treatment strategies based on the genetic basis can be pursued. Because increased residual activity is associated with improved clinical outcome, strategies that enhance the residual activity of a mutation can be pursued.⁵ Patients with in-frame premature stop codons, estimated as 11% of *GLDC* alleles, could benefit from medications designed to promote translation through a premature stop codon in mRNA in order to produce functional proteins, as suggested for other inborn errors of metabolism.^{33,34} For unstable missense mutants, chemical chaperones to restore protein folding have been suggested as a therapeutic strategy in various genetic disorders, including inborn errors of metabolism.³⁵ Specifically, pyridoxal 5'-phosphate has been investigated as a chaperone for various vitamin B₆-dependent enzymes, similar to GCS.³⁶ This report provides a valuable overview of mutations that cause classical NKH as genotype-based therapies are being developed.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/gim>

ACKNOWLEDGMENTS

This work was supported, in part, by Fundación Ramón Areces (CIVP16A1853) and Spanish Ministerio de Economía y Competitividad and el Fondo Europeo de Desarrollo Regional (FEDER) (PI12/02078) (PRP) MINECO-FEDER. The study was funded by donations from the following organizations: NKH Crusaders Fund, Hope for NKH Fund, CU Nonketotic Hyperglycinemia Fund, Joseph's Goal Fund, Brodyn's Friends Fund, Smiles for Miles NKH Research Fund, Lucy's BEElievers Fund, Madi's Mission to Find a Cure for NKH Fund, and Les Petits Bourdons. These funding sources had no role or influence in the study design, the collection and interpretation of the data, or the writing of the manuscript.

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Hennermann JB, Berger JM, Grieben U, Scharer G, Van Hove JL. Prediction of long-term outcome in glycine encephalopathy: a clinical survey. *J Inher Metab Dis* 2012;35:253–261.

- Hoover-Fong JE, Shah S, Van Hove JL, Applegarth D, Toone J, Hamosh A. Natural history of nonketotic hyperglycinemia in 65 patients. *Neurology* 2004;63:1847–1853.
- Hennermann JB. Clinical variability in glycine encephalopathy. *Future Neurology*. 2006;1:621–630.
- Baker PR 2nd, Friederich MW, Swanson MA, et al. Variant non ketotic hyperglycinemia is caused by mutations in LIAS, BOLA3 and the novel gene GLRX5. *Brain* 2014;137(Pt 2):366–379.
- Swanson MA, Coughlin CR Jr, Scharer GH, et al. Biochemical and molecular predictors for prognosis in nonketotic hyperglycinemia. *Ann Neurol* 2015;78:606–618.
- Korman SH, Boneh A, Ichinohe A, et al. Persistent NKH with transient or absent symptoms and a homozygous GLDC mutation. *Ann Neurol* 2004;56:139–143.
- Burton BK, Pettenati MJ, Block SM, Bensen J, Roach ES. Nonketotic hyperglycinemia in a patient with the 9p- syndrome. *Am J Med Genet* 1989;32:504–505.
- Isobe M, Koyata H, Sakakibara T, Momoi-Isobe K, Hiraga K. Assignment of the true and processed genes for human glycine decarboxylase to 9p23-24 and 4q12. *Biochem Biophys Res Commun* 1994;203:1483–1487.
- Takayanagi M, Kure S, Sakata Y, et al. Human glycine decarboxylase gene (GLDC) and its highly conserved processed pseudogene (psiGLDC): their structure and expression, and the identification of a large deletion in a family with nonketotic hyperglycinemia. *Hum Genet* 2000;106:298–305.
- Kume A, Koyata H, Sakakibara T, Ishiguro Y, Kure S, Hiraga K. The glycine cleavage system. Molecular cloning of the chicken and human glycine decarboxylase cDNAs and some characteristics involved in the deduced protein structures. *J Biol Chem* 1991;266:3323–3329.
- Kure S, Kato K, Dinopoulos A, et al. Comprehensive mutation analysis of GLDC, AMT, and GCSH in nonketotic hyperglycinemia. *Hum Mutat* 2006;27:343–352.
- Toone JR, Applegarth DA, Coulter-Mackie MB, James ER. Biochemical and molecular investigations of patients with nonketotic hyperglycinemia. *Mol Genet Metab* 2000;70:116–121.
- Kure S, Takayanagi M, Narisawa K, Tada K, Leisti J. Identification of a common mutation in Finnish patients with nonketotic hyperglycinemia. *J Clin Invest* 1992;90:160–164.
- Applegarth DA, Toone JR. Nonketotic hyperglycinemia (glycine encephalopathy): laboratory diagnosis. *Mol Genet Metab* 2001;74:139–146.
- Kanno J, Hutchin T, Kamada F, et al. Genomic deletion within GLDC is a major cause of non-ketotic hyperglycinaemia. *J Med Genet* 2007;44:e69.
- Nanao K, Takada G, Takahashi E, et al. Structure and chromosomal localization of the aminomethyltransferase gene (AMT). *Genomics* 1994;19:27–30.
- Analysis of protein-coding genetic variation in 60,706 humans. *bioRxiv*. <http://biorxiv.org/content/early/2015/10/30/030338>. Accessed November 18, 2015.
- Zech M, Boesch S, Jochim A, et al. Large-scale TUBB4A mutational screening in isolated dystonia and controls. *Parkinsonism Relat Disord* 2015;21:1278–1281.
- Hasse D, Andersson E, Carlsson G, et al. Structure of the homodimeric glycine decarboxylase P-protein from *Synechocystis* sp. PCC 6803 suggests a mechanism for redox regulation. *J Biol Chem* 2013;288:35333–35345.
- Toone JR, Applegarth DA, Kure S, et al. Novel mutations in the P-protein (glycine decarboxylase) gene in patients with glycine encephalopathy (non-ketotic hyperglycinemia). *Mol Genet Metab* 2002;76:243–249.
- Conter C, Rolland MO, Cheillan D, Bonnet V, Maire I, Froissart R. Genetic heterogeneity of the GLDC gene in 28 unrelated patients with glycine encephalopathy. *J Inher Metab Dis* 2006;29:135–142.
- Applegarth DA, Toone JR, Lowry RB. Incidence of inborn errors of metabolism in British Columbia, 1969-1996. *Pediatrics* 2000;105:e10.
- Boneh A, Korman SH, Sato K, et al. A single nucleotide substitution that abolishes the initiator methionine codon of the GLDC gene is prevalent among patients with glycine encephalopathy in Jerusalem. *J Hum Genet* 2005;50:230–234.
- Narisawa A, Komatsuzaki S, Kikuchi A, et al. Mutations in genes encoding the glycine cleavage system predispose to neural tube defects in mice and humans. *Hum Mol Genet* 2012;21:1496–1503.
- Nakai T, Nakagawa N, Maoka N, Masui R, Kuramitsu S, Kamiya N. Structure of P-protein of the glycine cleavage system: implications for nonketotic hyperglycinemia. *EMBO J* 2005;24:1523–1536.
- Wong KK, deLeeuw RJ, Dosanjh NS, et al. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet* 2007;80:91–104.
- Kure S, Korman SH, Kanno J, et al. Rapid diagnosis of glycine encephalopathy by 13C-glycine breath test. *Ann Neurol* 2006;59:862–867.
- Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015;17:405–423.
- Brenton JN, Rust RS. Late-onset nonketotic hyperglycinemia with a heterozygous novel point mutation of the GLDC gene. *Pediatr Neurol* 2014;50:536–538.
- Boneh A, Allan S, Mendelson D, Spriggs M, Gillam LH, Korman SH. Clinical, ethical and legal considerations in the treatment of newborns with non-ketotic hyperglycinaemia. *Mol Genet Metab* 2008;94:143–147.
- Bjoraker KJ, Swanson MA, Coughlin CR 2nd, et al. Neurodevelopmental outcome and treatment efficacy of benzoate and dextromethorphan in siblings with attenuated nonketotic hyperglycinemia. *J Pediatr* 2016;170:234–239.
- Hamosh A, McDonald JW, Valle D, Francomano CA, Niedermeyer E, Johnston MV. Dextromethorphan and high-dose benzoate therapy for nonketotic hyperglycinemia in an infant. *J Pediatr* 1992;121:131–135.
- Buck NE, Wood LR, Hamilton NJ, Bennett MJ, Peters HL. Treatment of a methylmalonyl-CoA mutase stopcodon mutation. *Biochem Biophys Res Commun* 2012;427:753–757.
- Sánchez-Alcudia R, Pérez B, Ugarte M, Desviat LR. Feasibility of nonsense mutation readthrough as a novel therapeutical approach in propionic acidemia. *Hum Mutat* 2012;33:973–980.
- Majtan T, Liu L, Carpenter JF, Kraus JP. Rescue of cystathionine beta-synthase (CBS) mutants with chemical chaperones: purification and characterization of eight CBS mutant enzymes. *J Biol Chem* 2010;285:15866–15873.
- Cellini B, Montioli R, Oppici E, Astegno A, Voltattorni CB. The chaperone role of the pyridoxal 5'-phosphate and its implications for rare diseases involving B6-dependent enzymes. *Clin Biochem* 2014;47:158–165.