A founder *AGL* mutation causing glycogen storage disease type IIIa in Inuit identified through whole-exome sequencing: a case series

Isabelle Rousseau-Nepton MD, Minoru Okubo MD PhD, Rosemarie Grabs BsC, the FORGE Canada Consortium,* John Mitchell MD, Constantin Polychronakos MD, Celia Rodd MD MSc

ABSTRACT

Background: Glycogen storage disease type III is caused by mutations in both alleles of the *AGL* gene, which leads to reduced activity of glycogen-debranching enzyme. The clinical picture encompasses hypoglycemia, with glycogen accumulation leading to hepatomegaly and muscle involvement (skeletal and cardiac). We sought to identify the genetic cause of this disease within the Inuit community of Nunavik, in whom previous DNA sequencing had not identified such mutations.

Methods: Five Inuit children with a clinical and biochemical diagnosis of glycogen storage disease type IIIa were recruited to undergo genetic testing: 2 underwent whole-exome sequencing and all 5 underwent Sanger sequencing to confirm the identified mutation. Selected DNA regions near the *AGL* gene were also sequenced to identify a potential founder effect in the community. In addition, control samples from 4 adults of European descent and 7 family members of the affected children were analyzed for the specific mutation by Sanger sequencing.

Results: We identified a homozygous frameshift deletion, c.4456delT, in exon 33 of the *AGL* gene in 2 children by whole-exome sequencing. Confirmation by Sanger sequencing showed the same mutation in all 5 patients, and 5 family members were found to be carriers. With the identification of this mutation in 5 probands, the estimated prevalence of genetically confirmed glycogen storage disease type IIIa in this region is among the highest worldwide (1:2500). Despite identical mutations, we saw variations in clinical features of the disease.

Interpretation: Our detection of a homozygous frameshift mutation in 5 Inuit children determines the cause of glycogen storage disease type IIIa and confirms a founder effect.

G lycogen storage disease type III is a rare autosomal recessive disease characterized by recurrent hypoglycemia in childhood, as well as hepatomegaly with elevated transaminases and hyperlipidemia.[1] The disease involves a defect in the key glycogen debranching enzyme, which has 2 enzymatic activities (amylo-1,6-glucosidase and 4-α-glucanotransferase), resulting in reduced glycogen degradation, accumulation of limit dextrin in affected organs (primarily skeletal muscle, cardiac muscle and liver), organomegaly and dysfunction. Glycogen storage disease type IIIa involves the liver and cardiac and skeletal muscles, whereas glycogen storage disease type IIIb involves only the liver.

Mutations in the *AGL* gene encoding glycogen debranching enzyme have been described in many populations, including Northern European,[2] Egyptian,[3] Hispanic[2] and Asian;[2] a high prevalence of the disease was also found in the North African Jewish community (1/5400) and in the Faroe Islands (1/3600).[4,5] We previously described the presenting clinical characteristics of 4 Inuit children with putative glycogen storage disease type III and suspected the presence of a founder effect.[6] However, targeted genetic analysis had failed to identify a mutation in the *AGL* gene. The aim of our present study was to identify the genetic cause of glycogen storage disease type III in the Inuit population of Nunavik on the eastern coast of Hudson Bay. By using exome sequencing, which examines all protein-encoding DNA sequences (exons), we hoped to facilitate early diagnosis, prenatal and neonatal screening and screening of family members.

Methods

Participants

Five Inuit children (aged 3 mo to 5 yr) with a clinical and biochemical diagnosis of glycogen storage disease type IIIa followed at the Montreal Children's Hospital, 4 of whom were previously identified,[6] underwent genetic testing between 2009 and 2012. Two of the children underwent whole-exome sequencing, and all underwent confirmation of the identified mutation by conventional Sanger sequencing. In addition, the children had been included in the international study on glycogen storage disease type III (Christiaan P. Sentner, University Medical Centre Groningen, The Netherlands: unpublished data, 2006–2007). Control samples from 4 adults of European descent participating in another study and 7 family members of the affected children were analyzed for the specific mutation by Sanger sequencing. We obtained informed consent from all participants. The study was approved by the Montreal Children's Hospital Research Ethics Board.

Whole-exome sequencing

Whole-exome sequencing consists of sequencing all exons — the DNA that encodes proteins. These regions represent only about 1% of all genetic material, but they contain almost all known mutations causing monogenic diseases such as glycogen storage disease type III. The remaining 99% of genetic material consists of introns, which are large regulatory sequences.

Exome sequencing was conceived as a discovery tool for locating disease-causing genes and is increasingly used as a powerful diagnostic modality.[7] Unlike traditional Sanger sequencing (where a single targeted gene is sequenced at a time), many gene locations can be probed rapidly, and mutations in genes previously not associated with disease can be identified. Once the exome (i.e.,

the exonic material) is isolated, individual DNA molecules are sequenced multiple times to assure accuracy. Many other genetic variations are typically discovered in an individual, most of which are recessive or not associated with disease.

We performed whole-exome sequencing as part of the Finding of Rare Disease Genes in Canada (FORGE Canada) Consortium. We collected peripheral blood samples from 2 children with glycogen storage disease type III and extracted the DNA using a standard protocol. We captured exome sequences using the 50 Mb Agilent SureSelect bait library, a collection of synthetic DNA fragments complementary to the target exons, to which they bind. We then sequenced the exomes on the Illumina HiSeq instrument at 3 indexed samples per lane for a depth of about 100 (i.e., each nucleotide was sequenced about 100 times). We aligned the results of all whole-exome sequences for the 2 patients against the reference human sequence to look for differences that might explain why the children had the disease.

Sanger sequencing

All of the children with glycogen storage disease type IIIa and their available family members underwent DNA testing by Sanger sequencing to confirm the mutation detected by whole-exome sequencing. Sanger sequencing is the conventional method to detect DNA mutations, in which the gene is "read" from start to finish.

Haplotyping

We sequenced selected DNA regions close to the *AGL* gene. These regions would typically have been inherited from a common ancestor carrying the mutation, thus examining their similarity (haplotypes) helps to identify a potential founder effect in the community (i.e., identical haplotypes suggest highly shared DNA).[8]

Patient	Age at diagnosis, yr	Sex	Reason for referral	Physical findings
1	1.6	F	Hepatomegaly; unable to walk	Liver span 15 cm; heart murmur; hypotonia
2	1.6	F	Abdominal distension; suspected hypoglycemia	Liver span 15 cm; heart murmur
3	0.3	M	Hepatomegaly	Liver edge 4 cm below costal margin
4	1.9	F	Abdominal distension with hepatomegaly; sister of proband	Liver span 10 cm; heart murmur
5	1.0	M	Hepatomegaly; irritability; need for frequent feedings	Liver edge at right iliac fossa

Table 1: Clinical presentation of 5 Inuit children with glycogen storage disease type III*

Note: F = female, M = male.
*Adapted from Zimakas and Rodd, in which the first 4 patients' cases were detailed.[6]

Results

The 5 children involved in our study presented with hepatomegaly, difficulty walking or suspected hypoglycemia (Table 1). All of them had undergone laboratory investigations with results consistent with glycogen storage disease type IIIa (Table 2), although only 3 children had elevated creatine kinase. The children had no known consanguinity. The diagnoses were based on debranching enzyme activity (patients 1, 3 and 4), liver biopsy with biochemical profile consistent with glycogen storage disease type III (patient 2) and a classic biochemical profile with a positive family history (patients 4 and 5). Children's ages at last follow-up ranged from 5.5 to 16.5 years; hypoglycemia was still a problem for all of the children and was controlled with uncooked cornstarch (Table 3). The children ate regular meals and snacks; their diets were a blend of southern and traditional Inuit foods. Most of the children had clinically important end-organ damage. Sanger sequencing of all exons of the *AGL* gene was previously reported as showing no mutation in 4 of the 5 patients.

Given the small size and isolation of their community in Northern Quebec, we anticipated that all of the children would have inherited identical mutations from both parents (homozygous). The exome of any individual contains numerous DNA variants that could be disease-causing. Assuming recessive inheritance, we searched for genes carrying protein-altering changes on both copies (either homozygous or compound heterozygous) in both of the affected children who underwent whole-exome sequencing. We found 37 genes that satisfied these requirements; however, in 28 instances the changes were known polymorphisms with frequencies incompatible with a rare condition such as glycogen storage disease type III. Of the remaining 9 genes, the variants in all but 2 were seen in the 277 unaffected control sequences analyzed in the same centre (platform-specific variants), which again suggested that they were not disease-causing. Of the 2 remaining genes, one was *UNC13C* (unc-13 homologue *Caenorhabditis elegans*), a gene of unknown function, with 2 missense mutations (p.N100Y and p.D863G). The second gene was *AGL* (encoding glycogen debranching enzyme) with a homozygous frameshift deletion, c.4456delT, that mapped to location NM000_642 of exon 33.

We confirmed the single-base homozygous deletion of a thymine residue at this position of the *AGL* gene by Sanger sequencing (Appendix 1, www.cmaj.ca/lookup/suppl/doi:10.1503

Table 2: Biochemical profiles at presentation of 5 Inuit children with glycogen storage disease type III

Patient	AST, U/L (normal ≤ 45)	ALT, U/L (normal ≤ 19)	Creatine kinase, U/L (normal 53–310)	Triglycerides, mmol/L (normal 0.45–2.54)	Total cholesterol, mmol/L (normal 2.80–4.80)
1	604	358	633	1.88	3.68
2	4320	2655	167	1.97	6.27
3	281	160	546	2.22	3.39
4	271	256	573	2.11	NA
5	326	205	275	3.10	4.21

Note: ALT = alanine aminotransferase, AST = aspartate aminotransferase, NA = not available.

Table 3: Characteristics of 5 Inuit children with glycogen storage disease type IIIa at last follow-up visit

Patient	Fasting tolerance, h	Creatine kinase, U/L (normal 53–310)	Complications
1	9 without UCCS	1633	Cirrhosis, portal hypertension and gastroesophageal varices; stable muscle weakness
2	7 with UCCS	2557	Slight clinical muscle involvement
3	12 without UCCS	2755	Right ventricular hypertrophy, long corrected QT interval; clinical muscle involvement
4	4 with UCCS	2093	Mild clinical muscle involvement
5	7 with UCCS	1225	No clinical muscle involvement

Note: UCCS = uncooked cornstarch.

/cmaj.140840/-/DC1). All of the children with glycogen storage disease type IIIa had homozygous deletions. One sibling and 4 parents were heterozygous for the same mutation (i.e., one mutated copy and one normal copy), and 2 siblings had a normal *AGL* sequence. Thus, we confirmed complete segregation of the homozygous deletion with disease in all cases and the absence of the mutation in unaffected family members, which is a defining characteristic of disease-causing mutations.

To determine how genetically similar the children were, we examined the DNA near the *AGL* gene. All of the children had the same DNA sequences, or haplotypes (data not shown), which provides further evidence that c.4456delT is a founder mutation in Inuit people with glycogen storage disease type IIIa.

Interpretation

Our results confirm a homogeneous deletion of a thymine residue (c.4456delT) of the *AGL* gene in 5 Inuit children with glycogen storage disease type IIIa diagnosed using both clinical and biochemical approaches. The mutation identified has previously been described in 12 North African Jewish patients from 10 different families (all homozygous). This deletion causes the production of a shortened and unstable protein, resulting in rapid degradation of the enzyme, which leads to the clinical phenotype of excessive accumulation of glycogen in liver and muscle tissues. Our conservative estimate of prevalence is 1 in 2500 based on the population of Nunavik (about 10 750 in 2011), which corresponds to a carrier frequency of 1 in 25 (for implications for the population of Nunavik, see Appendix 2, www.cmaj.ca/lookup/suppl/doi:10.1503/cmaj.140840/-/DC1).

The prevalence of glycogen storage disease type III in North America is about 1 per 100 000.[1] As described by Shen and Chen,[9] 4 mutations are responsible for only a quarter of cases in the United States; otherwise, most children in North America have unique mutations.[2,10] Worldwide, more than 130 mutations have been reported in The Human Gene Mutation Database;[11] most of these are missense or nonsense mutations, small deletions or insertions. Reports document clusters of patients with homozygous mutations in the Faroe Islands (c.1222C > T),[5] in Turkey (c.3980G > A),[8] in Tunisia (c.3216_3217delGA),[12] and in the North African Jewish population (c.4555delT; with the new nomenclature convention, c.4556delT).[4] As described earlier, high prevalences are typical of founder effects; for example, 1 in 5400 in Jewish people from North Africa[4] and 1 in 3600 in

Faroe Islanders.[5] This effect has been confirmed in our population and others by showing little genetic variability on the chromosome on which the *AGL* gene is located (i.e., identical haplotype). Thus, the population shares much of the same genetic information and is at increased risk for otherwise rare diseases, particularly those with autosomal recessive inheritance.[13]

The question then arises as to why the Inuit and the North African Jewish population share the same mutation despite so many mutations having been described.[11] We know that c.4556delT has been detected as a known variant in a carrier state in the general population (in the 1000 Genomes database, www.1000genomes.org). Experts in glycogen storage disease propose that this shared deletion probably arises at a hotspot — a DNA site that is likely to be spontaneously mutated — as opposed to suggesting that the Inuit are of North African descent.[2] In addition, the absence of c.4556delT being reported in other populations with glycogen storage disease type IIIa could be secondary to difficulties and costs related to genetic testing, particularly in such a large gene (35 exons). Moreover, the mutation was missed in our earlier manual analysis of DNA from 4 patients by Sanger sequencing owing to an understandable human error (a single deleted thymine residue is easily missed within a thymine homopolymer), which highlights the complexity of the technique and its potential for errors. Whole-exome sequencing has advantages over conventional Sanger sequencing, because it repeats sequencing 100 times at each nucleotide. In several other disorders, conventional Sanger sequencing has failed to detect mutations that were later identified through whole-exome sequencing, showing the power of the latter technique.[14,15]

Despite genetic homogeneity, we found evidence of clinical heterogeneity among the Inuit children involved in our study. Three of the 5 children had muscle involvement with increased creatine kinase at presentation; with follow-up, 4 of the 5 children had clinical myopathy, and cardiac hypertrophy has developed in 1 child. Another child presented with cirrhosis and evolving portal hypertension with gastroesophageal varices. In contrast, no history of neuromyopathy was found in our patients as had previously been described for Jewish North African patients. Overall, myopathic changes are variable in both cohorts. Phenotypic variation is common among patients with different genotypes who have glycogen storage disease type IIIa.[1,3,11,16,17] In adults, the condition may only be diagnosed late once hepatomegaly, cirrhosis, hepatic carcinoma, cardiomyopathy or neuro-

muscular dysfunction has developed.[18] Clinical variability among patients with the same mutation could be related to the age at diagnosis, compliance with dietary recommendations, environment or unknown modifying genes.[2]

A possible reason for the clinical variability we saw might be the traditional Inuit diet of meat and fish rich in protein.[19] The optimal diet for glycogen storage disease type III is still controversial, but evidence suggests reversibility of muscle dysfunction and cardiomyopathy with a high-protein, and possibly ketogenic, diet.[20–22] At this time, we cannot ascribe the clinical variation to the traditional diet, because these children — similar to almost all Inuit from Nunavik — now eat a blended diet of southern and traditional food when available.[23]

Limitations

We are unable to confirm that the Inuit do not have an admixture of North African Jewish heritage. We are working to obtain the haplotype material from the article describing this original deletion.[4] Given the lack of other disorders commonly found in the North African population, we support the current theory of mutations arising at hotspots. We have most likely identified all living Inuit children with glycogen storage disease type IIIa. However, our prevalence may be an underestimate because there may be early sudden infant deaths secondary to glycogen storage disease, adults who might have a milder form of th disorder and inaccuracy in the number of people at risk in the communities of the eastern Hudson Bay region.

Conclusion

With the identification of c.4456delT as the disease-causing mutation of the *AGL* gene in Inuit children with glycogen storage disease type IIIa, we are now able to offer comprehensive genetic counselling and testing to this remote population. In addition, local health care professionals will be more aware of this condition in patients with atypical presentations. We anticipate that this information will allow earlier dietary intervention and close follow-up of hepatic and cardiac function. Knowledge of this mutation could also facilitate the implementation of neonatal screening for this disorder if communities so wish. These steps are likely fiscally sound given the relatively inexpensive cost of screening for this mutation in this small population. This type of program has been studied in other communities where founder mutations exist.[24] Neonatal screening would allow for early identification of patients before damage from hypoglycemia and accumulated glycogen, poten-

tially improving the long-term health of children living in the eastern Hudson Bay region.

References

1. Dagli A, Sentner CP, Weinstein DA. Glycogen Storage Disease Type III. In: Pagon RA, Adam MP, Bird TD, et al., editors. *GeneReviews®*. Seattle: University of Washington, Seattle; 1993.
2. Goldstein JL, Austin SL, Boyette K, et al. Molecular analysis of the AGL gene: identification of 25 novel mutations and evidence of genetic heterogeneity in patients with Glycogen Storage Disease Type III. *Genet Med* 2010;12:424-30.
3. Endo Y, Fateen E, El Shabrawy M, et al. Egyptian glycogen storage disease type III - identification of six novel AGL mutations, including a large 1.5 kb deletion and a missense mutation p. L620P with subtype IIId. *Clin Chem Lab Med* 2009;47:1233-8.
4. Parvari R, Moses S, Shen J, et al. A single-base deletion in the 3′-coding region of glycogen-debranching enzyme is prevalent in glycogen storage disease type IIIA in a population of North African Jewish patients. *Eur J Hum Genet* 1997;5:266-70.
5. Santer R, Kinner M, Steuerwald U, et al. Molecular genetic basis and prevalence of glycogen storage disease type IIIA in the Faroe Islands. *Eur J Hum Genet* 2001;9:388-91.
6. Zimakas PJ, Rodd CJ. Glycogen storage disease type III in Inuit children. *CMAJ* 2005;172:355-8.
7. Wang Z, Liu X, Yang BZ, et al. The role and challenges of exome sequencing in studies of human diseases. *Front Genet* 2013;4:160.
8. Aoyama Y, Ozer I, Demirkol M, et al. Molecular features of 23 patients with glycogen storage disease type III in Turkey: a novel mutation p.R1147G associated with isolated glucosidase deficiency, along with 9 AGL mutations. *J Hum Genet* 2009;54:681-6.
9. Shen JJ, Chen YT. Molecular characterization of glycogen storage disease type III. *Curr Mol Med* 2002;2:167-75.
10. Endo Y, Horinishi A, Vorgerd M, et al. Molecular analysis of the AGL gene: heterogeneity of mutations in patients with glycogen storage disease type III from Germany, Canada, Afghanistan, Iran, and Turkey. *J Hum Genet* 2006;51:958-63.
11. Cooper DNBE, Stenson PD, Phillips AD, et al. *The Human Gene Mutation Database*. Available: www.hgmd.org (accessed 2014 Dec. 10).
12. Mili A, Ben Charfeddine I, Amara A, et al. A c.3216_3217delGA mutation in AGL gene in Tunisian patients with a glycogen storage disease type III: evidence of a founder effect. *Clin Genet* 2012;82:534-9.
13. International HapMap Consortium. The International HapMap Project. *Nature* 2003;426:789-96.
14. Aten E, Sun Y, Almomani R, et al. Exome sequencing identifies a branch point variant in Aarskog-Scott syndrome. *Hum Mutat* 2013;34:430-4.
15. Omoyinmi E, Melo Gomes S, Standing A, et al. Brief Report: whole-exome sequencing revealing somatic NLRP3 mosaicism in a patient with chronic infantile neurologic, cutaneous, articular syndrome. *Arthritis Rheumatol* 2014;66:197-202.
16. Mili A, Ben Charfeddine I, Mamai O, et al. Molecular and biochemical characterization of Tunisian patients with glycogen storage disease type III. *J Hum Genet* 2012;57:170-5.
17. Okubo M, Spengos K, Manta P, et al. Phenotypical variability in glycogen storage disease type III with a recurrent AGL mutation c.750-753delAGAC. *Muscle Nerve* 2011;43:451.
18. Fellows IW, Lowe JS, Ogilvie AL, et al. Type III glycogenosis presenting as liver disease in adults with atypical histological features. *J Clin Pathol* 1983;36:431-4.
19. Kuhnlein HV, Soueida R, Receveur O. Dietary nutrient profiles of Canadian Baffin Island Inuit differ by food source, season, and age. *J Am Diet Assoc* 1996;96:155-62.
20. Dagli AI, Zori RT, McCune H, et al. Reversal of glycogen storage disease type IIIa-related cardiomyopathy with modification of diet. *J Inherit Metab Dis* 2009;32(Suppl 1):S103-6.
21. Sentner CP, Caliskan K, Vletter WB, et al. Heart failure due to severe hypertrophic cardiomyopathy reversed by low calorie, high protein dietary adjustments in a glycogen storage disease type IIIa patient. *JIMD Rep* 2012;5:13-6.
22. Valayannopoulos V, Bajolle F, Arnoux JB, et al. Successful treatment of severe cardiomyopathy in glycogen storage disease type III With D,L-3-hydroxybutyrate, ketogenic and high-protein diet. *Pediatr Res* 2011;70:638-41.
23. Blanchet C, Dewailly E, Ayotte P, et al. Contribution of selected traditional and market foods to the diet of Nunavik Inuit women. *Can J Diet Pract Res* 2000;61:50-9.
24. Carleton SM, Peck DS, Grasela J, et al. DNA carrier testing and newborn screening for maple syrup urine disease in Old Order Mennonite communities. *Genet Test Mol Biomarkers* 2010;14:205-8.

Affiliations: Department of Pediatrics (Rousseau-Nepton, Mitchell, Polychronakos), Montreal Children's Hospital, Montréal, Que.; Okinaka Memorial Institute for Medical Research (Okubo), Tokyo, Japan; Endocrine Genetics Laboratory (Grabs), Montreal Children's Hospital, McGill University Health Centre, Montréal, Que.; Department of Pediatrics and Child Health (Rodd), Winnipeg, Man.

Contributors: Isabelle Rousseau-Nepton has participated substantially in the analysis and interpretation of the data, ensured dissemination of findings to participants, and drafted and revised the article for important intellectual content. Minoru Okubo and Rosemarie Grabs acquired, analyzed and interpreted the data and revised the article for important intellectual content. John Mitchell analyzed and interpreted the data and revised the article for important intellectual content. Constantin Polychronakos analyzed and interpreted the data, helped draft the article and revised it for important intellectual content. Celia Rodd contributed to the conception of the study, recruited participants, analyzed and interpreted the data, helped draft the article and revised it for important intellectual content. All of the authors approved the final submitted version and agree to act as guarantors of the work.