Int J Clin Exp Med 2019;12(6):7657-7662
www.ijcem.com /ISSN:1940-5901/IJCEM0090875

Original Article

Exome sequencing identifies a novel mutation in the GAA gene in a patient with glycogen storage disease type II

Shikai Chen, Peipei Jin, Shunchang Sun

Department of Clinical Laboratory, Ruijin Hospital North, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Received January 5, 2019; Accepted April 11, 2019; Epub June 15, 2019; Published June 30, 2019

Abstract: Glycogen storage disease is a group of inherited metabolic diseases resulting from a defect in the synthesis or degradation of glycogen. Glycogen storage disease type II is an autosomal recessive disorder caused by a deficiency of acid alpha-glucosidase, and the subsequent intralysosomal accumulation of glycogen, predominantly in the skeletal muscles. This article aimed to definitively diagnose a Chinese myopathy family, and to find the causative mutations of myopathy. Exome sequencing, with its high chromosomal coverage and accuracy, has been successfully used to identify pathogenic mutations. Exome sequencing followed by Sanger sequencing was performed for a family with myopathy to screen for the causative mutations described in the present study. Results: Two compound heterozygous mutations, c.2236insT and c.2238G>C (p. Trp746Cys), were identified at codon 746 on exon 16 of the *GAA* gene for the Chinese myopathy family. The patient was diagnosed with glycogen storage disease type II. The c.2236insT is a novel frameshift mutation predicted to truncate the distal C-terminal domain of acid alpha-glucosidase. The c.2236insT variant is more severe compared with the mild p.Trp746Cys mutation. The current study suggests that exome sequencing is a powerful tool to find causative mutations in patients with myopathy. The novel c.2236insT mutation expands the genetic spectrum of glycogen storage disease type II, so it is helpful for genetic diagnosis.

Keywords: Glycogen storage disease type II, exome sequencing, acid alpha-glucosidase gene

Introduction

Glycogen storage disease (GSD) is characterized by abnormal genetic glycogen metabolism in the muscles, liver, and brain and is divided into types 0 to X. Some patients present with the muscle symptoms [1]. Glycogen storage disease type II (GSD II) was first described in 1932 by the Dutch pathologist Pompe [2]. GSD II is a lysosomal disease due to the deficiency of acid alpha-glucosidase (GAA), and it is also a heterogeneous disease which varies by age at onset, the extent of organ involvement, and the rate of disease progression. GSD II is an autosomal recessive lysosomal storage disease caused by the deficiency of the lysosomal enzyme GAA, which is a key enzyme in hydrolyzation of lysosomal glycogen to glucose, and the subsequent intralysosomal accumulation of glycogen, predominantly in skeletal and cardiac muscle [3].

A wide spectrum of clinical presentations have been described. There are three types of GSD II: infantile, juvenile, and adult onset. The *GAA* gene has been mapped to chromosome 17q25.2-q25.3, is about 20 kb in length, and consists of 20 exons [4]. To date, more than 550 mutations in the *GAA* gene have been described in Pompe's Disease Mutation Database (http://www.erasmusmc.nl/klinische_genetica/research/lijnen/pompe_center/moleculaire_aspecten/). These mutations comprise missense, nonsense, splice-site mutations and small insertions/deletions. In the present study, we revealed a novel frameshift mutation, c.2236insT, in the *GAA* gene using exome sequencing in a patient with juvenile-onset GSD II.

Materials and methods

This study was approved by the ethics committee of Shenzhen Baoan Hospital, Southern Medical University (20150203). Informed patient consent was obtained before initiating this study. A Chinese myopathy family with one

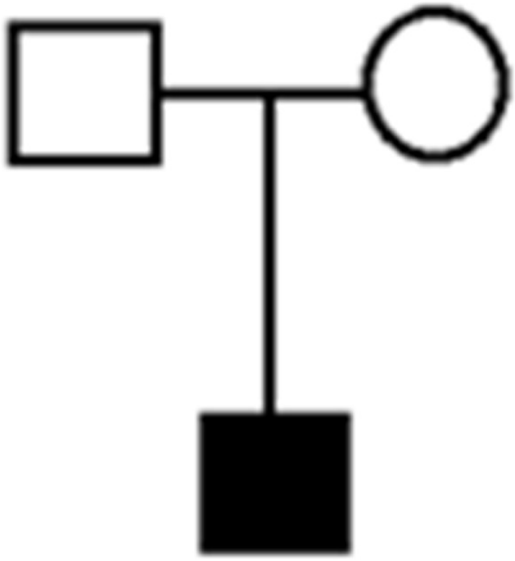

Figure 1. Pedigree of the family with myopathy.

affected child born to apparently unaffected parents was studied (**Figure 1**). The patient was diagnosed initially with myopathy by a clinical neurologist following a physical examination, an electromyography, and laboratory tests. The age of onset was 12 years old for the patient. Muscle weakness was the initial and the main symptom the patient complained of, and the muscle weakness occurred predominantly proximal to the extremities with hyper serum CK level (717 IU/L). The patient's parents were included in the present study in order to evaluate the segregation of the *GAA* gene mutations.

Genomic DNA was extracted from 5 ml peripheral blood samples using a blood genomic DNA purification kit (Qiagen, Germany) according to the manufacturer's protocol. One microgram of genomic DNA was randomly sheared into fragments with an average length between 250 and 300 bp by sonication. End repair, adenylation, and adapter ligation were performed for these fragments according to the Illumina protocol. Extracted DNA was amplified by ligation-mediated PCR (LM-PCR), purified, and hybridized to the SureSelect Human All Exon V5 (Agilent, USA) for enrichment. Quantitative PCR was performed to estimate the magnitude of enrichment for both captured and non-captured LM-PCR products. All captured libraries were then loaded on an Illumina Hiseq 4000 platform to generate paired end reads (Illumina, USA) according to the manufacturer's protocol. High-throughput sequencing was performed to meet the desired mean target sequencing depth. Clean reads that aligned to exon, splicing sites and immediate flanking intron regions were collected for mutation identification and subsequent analysis.

The raw data from the exome sequencing were processed using the Illumina Pipeline software (version 1.7) for base calling. Sequence reads were aligned to the human genome reference (NCBI build 37, hg19) for single nucleotide polymorphism (SNP) and insertion/deletion (indel) using a SOAP Aligner (version 2.21) and the Burrows-Wheeler Aligner tools. Low-quality SNP calls and indels were filtered out. All variants were categorized into coding, noncoding, and novel SNPs. All identified SNPs and indels were further filtered against the exome data from the 1000 Genome Project (ftp://www.1000genome.org) given that common variations were unlikely to be the causative mutations of myopathy. Synonymous mutations were removed. The functional effects of the mutations were predicted using the Sorting Intolerant From Tolerant tool.

Sanger sequencing was used on the samples from the myopathy family members to determine whether the potential mutations were causative, particularly the heterozygous and homozygous mutations co-segregated with the disease phenotype in the family. The primers for the polymerase chain reaction (PCR) and Sanger sequencing were designed with Primer3 software (http://frodo.wi.mit.edu/). The potential mutation region of the *GAA* gene was amplified by polymerase chain reaction using the following primers: 5'-tctgacctgagtcctccaag-3 (forward), 5'-caacacatacgttcctctttcc-3 (reverse). The direct sequencing of the PCR products was performed using ABI BigDye Dye Terminator Cycle Sequencing Kit (Applied Biosystems, USA) on an ABI 3730 Genetic Analyzer. The cycle sequencing of the PCR products was performed using Stratagene's RoboCycler Gradient 96 temperature cycler with Hot Top Assembly. The candidate causative mutations co-segregated with the disease phenotype were further estimated by Sanger sequencing of 200 unrelated healthy subjects with matched geographical ancestry.

Table 1. Variants identified by exome sequencing for the family with myopathy

Mutation	Father	Mother	Son
Missense variant	10367	10609	10623
Stop gained	86	93	92
Stop lost	41	38	40
Frameshift variant	275	296	281
Inframe insertion	147	152	163
Inframe deletion	187	172	178

Results

Exome sequencing was performed in the 3 individuals of the family with myopathy. Overall, 66, 68 and 61 million sequencing reads were produced for the genomic DNA from the father, mother, and affected son, respectively. An average of 6.59 billion bases of high-quality sequence was generated per simple, with an average sequencing depth of approximately 98.5 in the target region, which satisfied the requirements for calling SNPs and indels. The numbers of variants in the family with myopathy identified by exome sequencing are shown in **Table 1**. Most of variants were missense mutations. Coding region indels, splice acceptor, and donor site mutations are likely to be pathogenic mutations for the family with myopathy. Next, these variants were retained for further analysis using two filtering criteria: if a variant was within the allele frequency cutoff, which was less than 0.01 in the dbSNP142, HapMap, 1000 human genome, or if a variant was found in other individuals of the myopathy family, and if two variants were found in the same gene in the affected son of the myopathy family. These filtering conditions reduced the list of candidate variants to 99 mutations from 16 different genes. Only two compound heterozygous mutations, c.2236insT and c.2238G>C on exon 16 of the *GAA* gene, were found co-segregated with the myopathy phenotype. Then these two variants in the *GAA* gene were retained for further analysis. The other 97 mutations screened by exome sequencing in the other genes were not related to the myopathy.

Two variants causing inherited myopathy, c.2236insT and c.2238G>C (p. Trp746Cys), occurred at codon 746 on exon 16 of the GAA gene for the patient. Sanger sequencing confirmed the transmission of two mutations and a complete co-segregation with the disease phenotype in the GSD II patient (**Figure 2**). The c.2238G>C mutation in the *GAA* gene was reported in some patients with GSD II in a previous study. The c.2236insT mutation was identified in a heterozygous state in the *GAA* gene for the mother, while the c.2238G>C mutation occurred in a heterozygous state for the father. The c.2236insT and c.2238G>C mutations were not found in the 200 healthy subjects.

The c.2236insT mutation in the *GAA* gene has not been reported in patients with GSD II. The novel c.2236insT mutation was not found in any of the healthy subjects. It was not also found in the dbSNP database or the 1000 Genomes data. The c.2236insT was a causative mutation leading to a premature stop (Glu 795 Stop).

Discussion

The exome sequencing approach is a rapid, high-throughput, and cost-effective approach to diagnose genetic disorders [5]. Myopathies are diseases that affect skeletal muscles caused by genetic defects, metabolic disorders, inflammation and medications. There are different types of genetic myopathies, most of which have similar symptoms. GSD is a kind of severe myopathy. GSD II is characterized by myopathy and abnormally diminished muscle tone, and additional abnormalities may include cardiomegaly, hepatomegaly, and/or macroglossia [6]. In the present study, whole exome sequencing was used to find the causative gene of a Chinese family with myopathy. Two variants, c.2236insT and c.2238G>C, were found at codon 746 on exon 16 of the *GAA* gene. GSD II is an autosomal recessive metabolic myopathy characterized by GAA deficiency. The GAA enzyme catalyzes the hydrolysis of the α-1,4 and α-1,6-glucosidic bonds of glycogen, and its deficiency causes lysosomal glycogen storage in skeletal muscles.

Patients with GSD II display a considerable clinical heterogeneity with regard to the onset age, the rate of progression and the severity of the disorder [7]. The *GAA* gene contains 20 exons. Human GAA is composed of five distinct regions: trefoil (P-type) domain, N-terminal β-sandwich, catalytic $(\beta/\alpha)_8$ barrel, proximal C-terminal and distal C-terminal domains [8]. The Asp518 and Asp616 residues are key catalytic sites for GAA [9]. GAA is synthesized as an

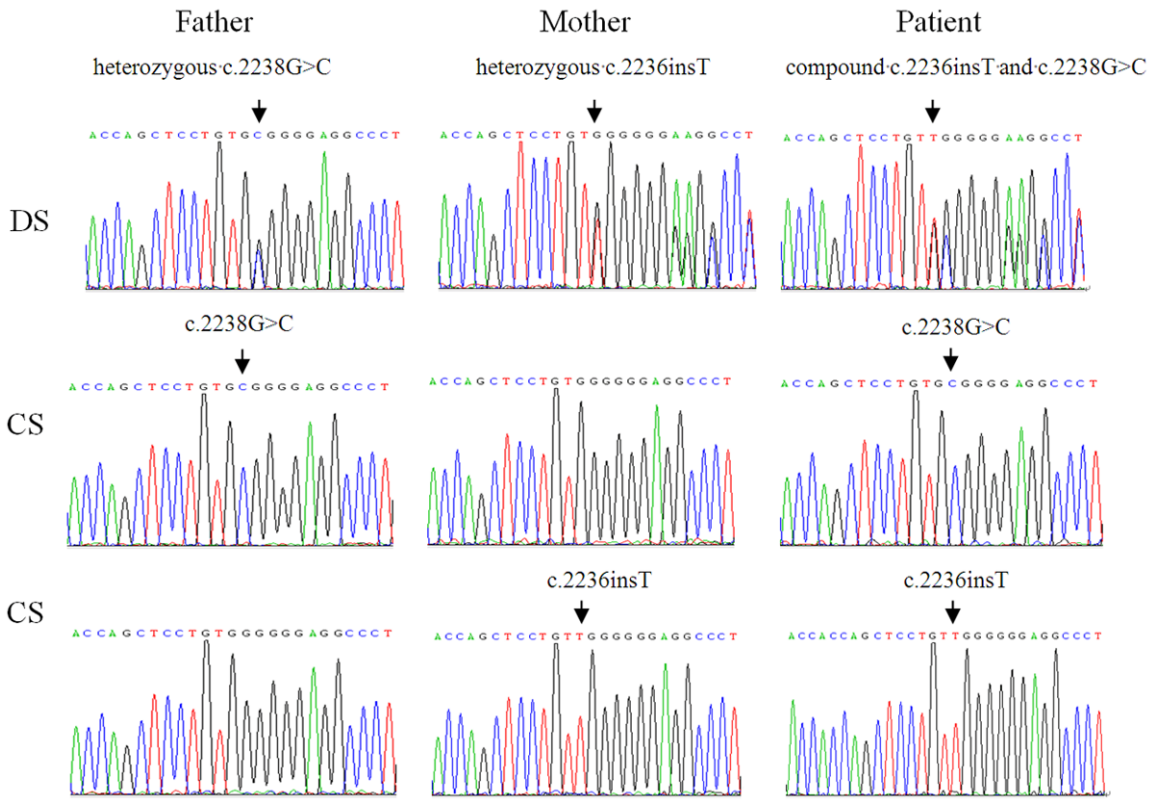

Figure 2. Mutations on exon 16 of the GAA gene. DS: direct sequencing of the PCR product; CS: cycle sequencing of the PCR product.

inactive enzyme precursor of 110 kD which is subsequently transported into pre-lysosomal and lysosomal compartments by the mannose 6-phosphate receptor. In lysosomal compartments, it is processed to a 95 kD intermediate and eventually to fully active enzymes of 76 and 70 kD [10]. Clinical heterogeneity of GSD II is caused by the *GAA* mutations that determine the degree of GAA deficiency [11].

Some studies have proposed different molecular mechanisms underlying GSD II etiopathology in vitro. The types of mutations causing GSD II include missense and splice site mutations, nonsense mutations, and partial insertions/deletions [12]. Pathogenic missense mutations usually affect posttranslational processing, folding, and the intracellular transport of GAA, which partially or entirely abolishes its function [13]. As of now, 557 GAA sequence variations have been listed in the Pompe Disease Mutation database (www.pompecenter.nl), and of these 85 variations are of unknown effect, 99 variations are considered as non-pathogenic or presumably non-pathogenic, 169 variations are considered as less severe or potentially less severe, 22 variations are potentially mild, and 182 variations are considered very severe. The p.Trp746Cys variant can be classified as a potentially mild mutation since the 110 kD GAA precursor as well as the processed forms of GAA were detected with just a small amount [14]. Certainly, some controversy exists concerning the p.Trp746Cys pathogenicity since it was described as a polymorphism or a causative mutation. Some studies have shown that p.Trp746Cys is the most common mutation in the Chinese late-onset Pompe patients [15]. The present study demonstrated the p.Trp746Cys variant was deleterious. The study argued in favor of a potentially less severe variation for the p.Trp746Cys substitution.

The c.2236insT mutation is a possibly causative mutation because it is a frameshifting mutation predicting a premature stop (Glu 795 Stop). The frameshifting mutation can lead to a premature termination codon on exon 16 of the *GAA* gene. Truncating mutations often lead to instability of the enzymes. Some studies have

suggested that a premature termination codon may cause the nonsense-mediated decay of mRNA and the production of no fully active enzymes [16]. Certainly, if the premature termination codon occurs on the last exon of a gene or the 3' end of the penultimate exon, the decay machinery mediated by the nonsense appears to miss it, and some mature protein is made [17]. The c.2236insT mutation occurs on exon 16 of the *GAA* gene consisting of 20 exons. Therefore, the c.2236insT variant should be a very severe mutation. Further studies are necessary to explore how the c.2236insT mutation contributes to the clinical features of GSD II.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (31571294).

Disclosure of conflict of interest

None.

Address correspondence to: Shunchang Sun, Department of Clinical Laboratory, Ruijin Hospital North, Shanghai Jiao Tong University School of Medicine, 999 Xiwang Road, Jiading, Shanghai 201801, China. Tel: 8613636359946; E-mail: shunchangsun@aliyun.com

References

[1] Kuperus E, Kruijshaar ME, Wens SCA, de Vries JM, Favejee MM, van der Meijden JC, Rizopoulos D, Brusse E, van Doorn PA, van der Ploeg AT, van der Beek NAME. Long-term benefit of enzyme replacement therapy in pompe disease: a 5-year prospective study. Neurology 2017; 89: 2365-2373.

[2] Roig-Zamboni V, Cobucci-Ponzano B, Iacono R, Ferrara MC, Germany S, Bourne Y, Parenti G, Moracci M, Sulzenbacher G. Structure of human lysosomal acid α-glucosidase-a guide for the treatment of Pompe disease. Nat Commun 2017; 8: 1111.

[3] McIntosh P, Austin S, Sullivan J, Bailey L, Bailey C, Viskochil D, Kishnani PS. Three cases of multi-generational pompe disease: are current practices missing diagnostic and treatment opportunities? Am J Med Genet A 2017; 173: 2628-2634.

[4] Tsai AC, Hung YW, Harding C, Koeller DM, Wang J, Wong LC. Next generation deep sequencing corrects diagnostic pitfalls of traditional molecular approach in a patient with prenatal onset of Pompe disease. Am J Med Genet A 2017; 173: 2500-2504.

[5] Mologni L, Piazza R, Khandelwal P, Pirola A, Gambacorti-Passerini C. Somatic mutations identified at diagnosis by exome sequencing can predict response to imatinib in chronic phase chronic myeloid leukemia (CML) patients. Am J Hematol 2017; 92: E623-E625.

[6] Vega AI, Medrano C, Navarrete R, Desviat LR, Merinero B, Rodríguez-Pombo P, Vitoria I, Ugarte M, Pérez-Cerdá C, Pérez B. Molecular diagnosis of glycogen storage disease and disorders with overlapping clinical symptoms by massive parallel sequencing. Genet Med 2016; 18: 1037-1043.

[7] Barba-Romero MA, Garcia-Cuartero I. Importance of clinical suspicion in the diagnosis of late-onset Pompe disease. Rev Neurol 2016; 236: 236-238.

[8] Doerfler PA, Todd AG, Clément N, Falk DJ, Nayak S, Herzog RW, Byrne BJ. Copackaged AAV9 vectors promote simultaneous immune tolerance and phenotypic correction of pompe disease. Hum Gene Ther 2016; 27: 43-59.

[9] Güngör D, Kruijshaar ME, Plug I, Rizopoulos D, Kanters TA, Wens SC, Reuser AJ, van Doorn PA, van der Ploeg AT. Quality of life and participation in daily life of adults with Pompe disease receiving enzyme replacement therapy: 10 years of international follow-up. J Inherit Metab Dis 2016; 39: 253-260.

[10] Jung JW, Huy NX, Kim HB, Kim NS, Van Giap D, Yang MS. Production of recombinant human acid α-glucosidase with high-mannose glycans in gnt1 rice for the treatment of Pompe disease. J Biotechnol 2017; 249: 42-50.

[11] Rastall DP, Seregin SS, Aldhamen YA, Kaiser LM, Mullins C, Liou A, Ing F, Pereria-Hicks C, Godbehere-Roosa S, Palmer D, Ng P, Amalfitano A. Long-term, high-level hepatic secretion of acid α-glucosidase for Pompe disease achieved in non-human primates using helper-dependent adenovirus. Gene Ther 2016; 23: 743-752.

[12] Angelini C, Savarese M, Fanin M, Nigro V. Next generation sequencing detection of late onset pompe disease. Muscle Nerve 2016; 53: 981-983.

[13] Musumeci O, Thieme A, Claeys KG, Wenninger S, Kley RA, Kuhn M, Lukacs Z, Deschauer M, Gaeta M, Toscano A, Gläser D, Schoser B. Homozygosity for the common GAA gene splice site mutation c.-32-13T>G in Pompe disease is associated with the classical adult phenotypical spectrum. Neuromuscul Disord 2015; 25: 719-724.

[14] Turaça LT, de Faria DO, Kyosen SO, Teixeira VD, Motta FL, Pessoa JG, Rodrigues E Silva M, de Almeida SS, D'Almeida V, Munoz Rojas MV,

Martins AM, Pesquero JB. Novel GAA mutations in patients with Pompe disease. Gene 2015; 561: 124-131.

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

[16] Todd AG, McElroy JA, Grange RW, Fuller DD, Walter GA, Byrne BJ, Falk DJ. Correcting neuromuscular deficits with gene therapy in Pompe disease. Ann Neurol 2015; 78: 222-234.

[17] Montagnese F, Barca E, Musumeci O, Mondello S, Migliorato A, Ciranni A, Rodolico C, De Filippi P, Danesino C, Toscano A. Clinical and molecular aspects of 30 patients with late-onset Pompe disease (LOPD): unusual features and response to treatment. J Neurol 2015; 262: 968-978.