A Homozygous Mutation in GPT2 Associated with Nonsyndromic Intellectual Disability in a Consanguineous Family from Costa Rica

Tanya Lobo-Prada • Heinrich Sticht •
Sixto Bogantes-Ledezma • Arif Ekici • Steffen Uebe •
André Reis • Alejandro Leal

Received: 23 October 2016 / Revised: 24 November 2016 / Accepted: 15 December 2016
© SSIEM and Springer-Verlag Berlin Heidelberg 2017

Abstract Intellectual disability is a highly heterogeneous disease that affects the central nervous system and impairs patients' ability to function independently. Despite multiples genes involved in the etiology of disease, most of the genetic background is yet to be discovered. We used runs of homozygosity and exome sequencing to study a large Costa Rican family with four individuals affected with severe intellectual disability and found a novel homozygous missense mutation, p. 96G>R, c. 286G>A, in all affected individuals. This gene encodes for a pyridoxal enzyme involved in the production of the neurotransmitter glutamate and is highly expressed in the white matter of brain and cerebellum. Protein modeling of GPT2 predicted that the mutation is located in a loop where the substrate binds to the active site of the enzyme, therefore, suggesting that the catalytic activity is impaired. With our report of a second mutation we fortify the importance of *GPT2* as a novel cause

of autosomal recessive nonsyndromic intellectual disability and support the premise that *GPT2* is highly important for the neurodevelopment of the central nervous system.

Synopsis: The mutation p. 96G>R c. 286G>A in *GPT2*, located in a loop where the substrate binds to the active site of the enzyme, fortifies the importance of *GPT2* in the pathogenesis of nonsyndromic intellectual disability.

Intellectual disability (ID) is a disease that affects the central nervous system and it is diagnosed by three major criteria: IQ less than 70, deficits in at least two adaptive behaviors (communication, personal care, ability to live independently, work, health and/or personal safety) and the symptoms should present themselves before the age 18 (American Psychiatric Association 2013). ID is the most frequent disability in children and young adults (Leonard and Wen 2002). ID can be classified according to the IQ level (mild, moderate, severe, or profound) or according to the etiology of the disease, whether it is known (syndromic) or unknown (nonsyndromic).

Epidemiological studies have determined that about 25% of ID cases have a genetic background and this number increases with the severity of the disease (McLaren and Bryson 1987). Despite that numerous candidate genes have been involved in the etiology of the disease it is noteworthy that for about half of the patients it is still not possible to establish the cause of the disease. Nevertheless, recently, with the combination of exome sequencing and microarray technology it has become possible to identify rare genetic variants associated with this disease especially in consanguineous families.

The aim of this study was to combine autozygosity mapping and exome sequencing to determine the genetic

Communicated by: Ina Knerr, MD

T. Lobo-Prada • A. Leal (✉)
Section of Genetics and Biotechnology, School of Biology, University
of Costa Rica, 11501-2060 San Jose, Costa Rica
e-mail: alejandro.leal@ucr.ac.cr

T. Lobo-Prada • A. Leal
Neuroscience Research Center, University of Costa Rica, San Jose,
Costa Rica

H. Sticht
Institute of Biochemistry, Friedrich-Alexander-Universität Erlangen-
Nürnberg, Erlangen, Germany

A. Ekici • S. Uebe • A. Reis
Institute of Human Genetics, Friedrich-Alexander-Universität
Erlangen-Nürnberg, Erlangen, Germany

S. Bogantes-Ledezma
Neurology Department, Hospital San Juan de Dios, San Jose, Costa
Rica

cause of intellectual disability in a Costa Rican family with four affected siblings and whose parents are first-degree cousins. We report a novel homozygous missense mutation in *GPT2* – a reversible enzyme involved in the production of neurotransmitter glutamate. These findings contribute to elucidating genes that are critical for the adequate functioning of the central nervous system and understanding how mutations in these genes translate in different phenotypes.

Materials and Methods

Following informed consent genomic DNA was extracted from whole blood of all family members using the sucrose-phenol-chloroform method. We first excluded known micro-deletions associated with intellectual disability syndromes using MLPA P245-A2 kit and MLPA P297-B2 kit, a Fragile X syndrome and pathogenic CNVs using high-resolution chromosomal Affymetrix Cytoscan HD microarrays (Affymetrix, Santa Clara, California). Then we used polymorphic markers from the Cytoscan HD Arrays to perform a genome scan, followed by a linkage analysis considering an autosomal recessive pattern of inheritance, based on the disease segregation observed in the family. Other parameters were: a disease-gene frequency of 0.0001, uniform marker-allele frequencies and 100% penetrance. 10,000 SNPs were considered, and the EasyLINKAGE-plus v.5[®] (Hoffmann and Lindner 2005) Allegro[®] (Gudbjartsson et al. 2005) and HaploPainter[®] (Thiele and Nürnberg 2005) software were used. In order to make the linkage analysis possible, only data from all the affected individuals and three unaffected siblings were included. Genome positions were taken from the NCBI36/hg18 assembly (Karolchik et al. 2004, <https://genome.ucsc.edu/>). These SNPs were also taken into account to identify possible runs of homozygosity (ROH). The runs of homozygosity were calculated with PLINK v.1.07. (Purcell et al. 2007, <http://pngu.mgh.harvard.edu/purcell/plink/>) and were defined as genomic regions larger than 1 Mb with at least 1 SNP every 50 Kb. ROH shared only by affected individuals were considered candidate regions.

We used 5500x1 of Applied Biosystems (Applied BiosystemsTM, Foster City, California) for whole exome sequencing of two of the patients. Only non-synonymous, gain of function, insertions, and homozygous deletions were considered with special attention given to those inside the candidate region. We excluded common polymorphisms reported in dbSNP132, 1000 genome project (February 2012), 5400 Exome sequencing project (NHLBI), and 230 controls from the Institute of Human Genetics in Erlangen. A special attention was given to genes expressed in the central nervous system.

Confirmation and validation were performed through bidirectional direct sequencing, using the BigDye Terminator

Cycle Sequencing Kit (Applied Biosystems) on an ABI3730 capillary sequencer (Applied Biosystems). Sequence traces were evaluated using the DNASTar software package.

A crystal structure is available for human GPT2 (EC 2.6.1.2). However, this structure was crystallized without substrate and therefore the loops 95–104 and 143–152 are not resolved in the structure. To allow interpretation of the G96R and S153R mutations, the structure of barley GPT2 (PDB code: 3TCM; Duff et al. 2012) was used instead. This structure contains a PLP-cycloserine inhibitor and all active site residues are resolved in the crystal structure. Barley and human GPT2 exhibit 65% sequence similarity and the sites of mutation are conserved between both species. Modeling of the mutations was done with Swiss-PdbViewer (Guex and Peitsch 1997) and RasMol (Sayle and Milner-White 1995) was used for structure analysis and visualization.

Each affected individual was assessed for social competence by a neurologist using Vineland Social Maturity Scale. A complete clinical evaluation was also performed. Due to the severity of phenotype no common IQ evaluations could be performed. Structural brain abnormalities were evaluated by magnetic resonance imaging (MRI).

Results

We report a family with four patients, without any previous perinatal noxa, who could not reach an appropriate social level and had a severe delay in reaching different basic developmental milestones (Table 1). At the neurological exam they showed spastic diplegia, exalted osteotendinous reflexes, and bilateral extensor plantar reflex, suggesting pyramidal dysfunction. The two younger siblings affected had history of one status epilepticus during early childhood and presented a more severe phenotype, according to physical examination (i.e., sialorrhea and aggressive behavior) and lower social quotient (Table 2). Using MRI we found cortical atrophy thinning of corpus callosum and cerebellum and mild cerebellum atrophy; no other major structural findings were detected (Fig. 1).

The linkage analysis identified a locus on chromosome 16, segregating fully with the disease. Maximum two-point LOD scores were obtained for marker SNP_A-2159072 and others in the same region, $Z_{\max}(0.00) = 3.3803$. In addition, we found a ROH on chromosome 16 in all of the affected individuals but not in any of the healthy family members (chr16: 27157613–32483006). Nevertheless, after homozygosity mapping, this region appeared to be more extensive chr16: 27177795–48202503 and contained 226 genes of which 56 were noncoding genes (microRNA and pseudogenes) and 11 were reported as expressed in the brain. The WES coverage for this region was 80%, capturing 256 exons corresponding to 171 coding genes. After filtering and validating the exome sequencing results in this region,

Table 1 Vineland Social Maturity Scale measures and summary of important developmental milestones in the patients affected with intellectual disability of the consanguineous Costa Rican family

Characteristics of patients	PZ43	PZ35	PZ33	PZ26
Perinatal history	AGA, no perinatal noxa	AGA, no perinatal noxa	AGA, no perinatal noxa	AGA, no perinatal noxa
<i>Vineline Social Maturity Scale</i>				
Age equivalent (years)	5.6	5.6	3.7	2.3
Estimated social age (years)	22.4	22.4	14.8	9.2
Age of independent walking (years)	>2 (drags feet)	>2 (tiptoes and kicks)	>3 (stagger)	>3 (stagger)
Seizures, age, further treatment	No	No	One status epilepticus at 7 years old, no treatment	One status epilepticus at 8 months old, no treatment
Sialorrhea	No	No	Yes	Yes
Language				
First words age	>5	>5	>5	>5
Simple phrases (age in years)	>10	>10	>10	>10
Comprehension	Limited	Limited	Very limited	Extremely limited
Sphincter control age in years	4	4	6	Failed
Awareness of danger	Yes	Yes	No	No
Aggressive behaviors	No	No	Yes	Yes
Loss of acquired skills	No	No	No	No

AGA appropriate for gestational age

Table 2 Summary of clinical signs found on physical examination in the patients affected with intellectual disability of the consanguineous Costa Rican family

Clinical signs	PZ43	PZ35	PZ33	PZ26
Facial dysmorphism				
Incisors protrusion	+	+	+	+
High arched palate	+	+	+	+
Left double antihelix	—	—	—	+
Adduct thumb	—	Left	Both, pred. left	Both, pred. left
Fine motor skill				
Voluntary use of hand as clamp	+	+	+	+
Cutaneous stigmata				
Hypomelanosis guttata	+	+	+	+
café au lait spots	+	—	—	—
Becker nevus	—	+	—	+
Others				
Mitral systolic murmur	+	—	—	—
Mild right scoliosis		+	+	+
Cranial nerves	NAD	NAD	NAD	NAD
Motor activity				
Spastic diplegia	+	+	+	+
Frequent falls	—	—	—	+
Exalted osteotendinous reflexes	+	+	+	+
Plantar reflex	Bilateral extensor	Bilateral extensor	Bilateral extensor	Bilateral extensor
Cerebellum	NAD	NAD	NAD	NAD

NAD no abnormality detected

Fig. 1 Magnetic resonance imaging showing thinning of corpus callosum (white star) and mild cerebellum atrophy in four affected patients (two white stars)

we found three nonsynonymous variants in *ZNF646*, *ZNF629* and *GPT2*. The three genetic changes segregated with the disease in the family. Nevertheless, the genetic variations in *ZNF646* and *ZNF629* had been previously reported in dbSNP database with an allele frequency of 1:2,000 and therefore were considered unlikely as disease causing mutations.

After validation through Sanger sequencing, we found that genes *SMPD3* (Sphingomyelin Phosphodiesterase 3), *SETD1A* (Set domain-containing protein 1A), *TAOK2* (TAO kinase 2) y *KIAA0556* (mRNA not yet characterized) that had heterozygote genetic changes and were in the region of interest were false positives. We also sequenced exon 4 of *SBK1* (*Homo sapiens* SH3-binding domain kinase 1) that was not covered by the NGS and had been previously related to neurodevelopment but did not find any genetic changes.

We found a mutation in exon 5 of *GPT2* (OMIM #616281) that changed a guanine for an adenine (c. 286G>A), which resulted in an amino acid change in position 96 of the protein (p. 96G>R) (Fig. 2). This mutation was detected only once in 121,290 alleles, with a frequency of 0.0000008245. According to the latest version of Exome Aggregation Consortium (ExAC) this single allele was not found in 11,574 Latinos (Lek et al. 2016). The crystal structure of an inhibited GPT2 shows that G96 is in the immediate vicinity

of the substrate-binding site and makes contacts with the substrate (Fig. 1a, b). In the G96R variant, the arginine sidechain protrudes into the ligand-binding site and would result in steric clashes with a bound substrate (Fig. 3c) thereby affecting catalytic activity.

Bioinformatics prediction programs (SIFT, PhyloP, PolyPhen, and Mutation Taster) determined that this position is evolutionarily highly conserved and that the variant is predicted to be pathogenic. We did not find any pathogenic CNVs or other chromosomal structural or numerical abnormalities associated with the disease.

Discussion

We found a novel mutation, c. 286G>A p. 96G>R, in *GPT2* associated with autosomal recessive severe nonsyndromic intellectual disability in four children of a Costa Rican family of consanguineous marriage. *GPT2* encodes for a pyridoxal dependent enzyme involved in a reversible reaction that transfers an alpha amino group from oxoglutarate to L-alanine and the formation of pyruvate and L-glutamate. Therefore, *GPT2* has an important role in gluconeogenesis, amino acid metabolism and neurotransmitter production (Yang et al. 2002). We suggest that this mutation is the cause of ID in this family.

Fig. 2 Family pedigree and location of homozygous recessive mutation c. 286G>A in *GPT2* in four affected patients (black), heterozygous in both parents and dominant homozygous in a healthy sibling

GPT2 also known as *ALT2* is highly expressed in muscle, adipose tissue, kidney and brain and lower quantities have been found in liver and breast (Yang et al. 2002). The other isoform, *GPT1* or *ALT1* is expressed in most of the same organs except brain and it is known as an indicator of liver damage. Deregulations in either *GPT* isoforms have been associated with liver damage (Sherman 1991), obesity (Strauss et al. 2000) and muscle disease (Rutledge et al. 1985).

Recently, a missense homozygous mutation in *GPT2* was found to be associated with severe developmental encephalopathy in a Jewish origin family with three affected siblings (Celis et al. 2015). They present a *GPT2* nonsynonymous homozygous mutation, c. 459C>G, p. 153S>R, associated with static encephalopathy related to severe ID. Celis et al. (2015) reported that their mutation was near a catalytic residue highly conserved in evolution and, after functional assays that mutation resulted in a severe loss of enzymatic function. Inspection of the *GPT* three-dimensional structure shows that p. 153S>R is located in a loop of one subunit that is placed in close proximity to the active site of the second subunit (Fig. 3a). In the p. 153S>R variant, the longer sidechain of the

arginine causes steric clashes with S372 (Fig. 3e) that are not present in the wildtype due to the shorter S153 sidechain (Fig. 3d). The clashes formed in the p. 153S>R variant will cause misfolding of the loop thus offering a structural explanation for the experimentally observed reduced catalytic activity.

In the case of the Costa Rican family, the mutation was also found close to the substrate binding site (Fig. 3a) suggesting that both mutations affect catalytic activity. In contrast to the p. 153G>R variant, the p. 96G>R variant has a direct structural effect on substrate binding by protruding into the active site (Fig. 1c). This might indicate that both mutations affect catalytic activity to a different extent.

When comparing the Celis et al. patients with ours, some different clinical signs could be noticed. Both affected families share clinical signs such as a severe developmental delay, non-verbal language (except for two of the Costa Rican patients) and some clinical findings suggesting pyramidal dysfunction (e.g., brisk reflexes and bilateral extensor plantar responses, high arched palate, and sialorrhea). Nevertheless, when compared to the Costa Rican, Celis et al. family had characteristic facial dysmorphic features (e.g., bifrontal narrowing, hypotelorism, low or rotated set ears and epicanthal folds), microcephaly or

Fig. 3 (a) Structure of dimeric GPT2 showing the subunits as *blue* and *cyan* ribbons, respectively. The sites of the p. 96G>R (Costa Rican family) and p. 153S>R (Mizrahi Jewish origin family) exchange are highlighted in *magenta* and *yellow*, respectively. A PLP-cycloserine inhibitor present in the active site is shown in *space-filled* presentation and *colored* according to the atom types. (b, c) Effect of p. 96G>R variant. (b) p. 96G>R is located in close proximity to the substrate binding site and forms interactions with the

substrate-analogous inhibitor PLP-L-cycloserine. (c) The larger side-chain of arginine causes steric clashes with the substrate which will block substrate access to the active site. (d, e) Effect of a p. 153S>R variant. (d) p. 153S>R is located in a loop in close proximity to S372. (e) The larger sidechain of arginine causes steric clashes with S372, which will cause a rearrangement of the loop. Clashes are indicated by *magenta arrows*

dolichocephaly, developmental encephalopathy, seizures, axial hypertonia or hypotonia, and an incapacity to walk and talk. In our family, after an exhaustive clinical evaluation, we found some specific clinical features such as summarized in Table 2. We also found that the severity of the disease progresses with maternal age. The wide spectrum of clinical features in both families suggests that *GPT2* is highly important in brain functioning but much is still to be learnt

about the interactions of *GPT2* with other proteins and its role in brain development and physiology.

Glutamate – an excitatory neurotransmitter – is highly abundant in brain and it has been proven to be implicated with different processes involved in complex adaptive functioning: cognition, memory, and learning (McEntee and Crook 1993). Evidence has shown the importance of glutamate availability in brain homeostasis. Numerous neurodegenerative diseases

such as Parkinson, Alzheimer's, epilepsy, Huntington's, multiple sclerosis, and nonsyndromic ID have been associated with modifications in glutamate receptors or free glutamate levels. In our family, we suspect that c. 286G>A is probably affecting *GPT2* enzymatic activity by diminishing the available glutamate in the synaptic cleft and, therefore, causing severe ID.

Our findings of a second family carrying an independent missense mutation in *GPT2* support the hypothesis that homozygous missense mutations near catalytic residues of *GPT2* are associated with severe ID. The mutation c. 286G>A p. 96G>R found in our patients probably affects substrate binding to the active site, altering the enzymatic activity, therefore, glutamate metabolism. However, these findings must be interpreted with caution, however, since functional analysis is lacking. Functional studies are required to further confirm the importance of *GPT2* in neural development. In addition, regardless of high accuracy of exome sequencing coverage, there are coding regions that are not fully covered by this method. In our study the coverage was 80%, leaving some exons where possible mutations could be queried. Moreover, with the increasing understanding of the importance of noncoding regions in disease pathogenesis, we cannot rule out other mutations might be contributing to the phenotype seen in our family.

In conclusion, this study fortifies the importance of *GPT2* in the pathogenesis of ID. Given the severity of this disease, future studies into the molecular pathways in which this protein is involved are of crucial importance for a better understanding of the condition and to find potential targets to improve care for these patients.

Acknowledgments We thank the patients and their family for the kind cooperation.

Conflict of Interest

T L-P, H S, S B-L, A E, S U, A R, A L declare that we have no conflict of interest. We declare that this article has not been published previously and all the authors approve its publication.

Authors' Contribution

Tanya Lobo-Prada was involved in the conception design analysis interpretation and drafting of the article. PhD. Alejandro Leal and Prof. André Reis contributed to the conception and design and revised the manuscript critically for important intellectual content. Heinrich Sticht contributed with in silico modeling of the enzyme and contributed critically to the main article draft. Sixto Bogantes did

valuable to contributions to the clinical characterization of the patients and revised the manuscript critically. Arif Ekici and Steffen Uebe contributed to analysis and interpretation of the data and revised the article critically.

Guarantor and Corresponding Author

Alejandro Leal accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Details of Funding

This research was funded by German Research Foundation (DFG) grant AB393/2-2, University of Costa Rica (project 111-B1-349), German Academic Exchange Program (DAAD) and Graduate Studies System of University of Costa Rica. The authors confirm independence from the sponsors; the content of the article has not been influenced by the sponsors.

Details of Ethical Approval

The study was performed as per protocol of the University of Costa Rica's Institutional Review Board.

Patient Consent Statement

All authors declare that informed consent was obtained from all the subjects on this research in accordance with the Code of Ethics of the World Medical Association. The privacy rights of the human subjects were also protected.

References

- American Psychiatric Association (2013) Diagnostic and statistical manual of mental disorders, 5th edn. American Psychiatric Association, Washington
- Celis K, Shuldiner S, Haverfield EV, Cappell J, Yang R, Gong DW, Chung WK (2015) Loss of function mutation in glutamic pyruvate transaminase 2 (*GPT2*) causes developmental encephalopathy. *J Inher Metab Dis* 38:941–948
- Duff SM, Rydel TJ, McClerren AL, Zhang W et al (2012) The enzymology of alanine aminotransferase (*AlaAT*) isoforms from *Hordeum vulgare* and other organisms, and the *HvAlaAT* crystal structure. *Arch Biochem Biophys* 528:90–101
- Gudbjartsson DF, Thorvaldsson T, Kong A, Gunnarsson G, Ingolfsdottir A (2005) Allegro version 2. *Nat Genet* 37:1015–1016
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723
- Hoffmann K, Lindner TH (2005) easyLINKAGE-Plus – automated linkage analyses using large-scale SNP data. *Bioinformatics* 21:3565–3567

- Karolchik D, Hinrichs AS, Furey TS, Roskin KM, Sugnet CW, Haussler D, Kent WJ (2004) The UCSC Table Browser data retrieval tool. *Nucleic Acids Res* 32:D493–D496
- Lek M, Karczewski K, Minike E et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536:285–291
- Leonard H, Wen X (2002) The epidemiology of mental retardation: Challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* 8:117–134
- McEntee WJ, Crook TH (1993) Glutamate: its role in learning, memory, and the aging brain. *Psychopharmacology (Berl)* 111:391–401
- McLaren J, Bryson SE (1987) Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. *Am J Ment Retard* 92:243–254
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81:559–575
- Rutledge J, Andersen J, Fink CW, Cook J, Strickland A (1985) Persistent hypertransaminasemia as the presenting finding of childhood muscle disease. *Clin Pediatr* 24:500–503
- Sayle RA, Milner-White EJ (1995) RASMOL: biomolecular graphics for all. *Trends Biochem Sci* 20:374
- Sherman KE (1991) Alanine aminotransferase in clinical practice. A review. *Arch Intern Med* 151:260–265
- Strauss RS, Barlow SE, Dietz WH (2000) Prevalence of abnormal serum aminotransferase values in overweight and obese adolescents. *J Pediatr* 136:727–733
- Thiele H, Nürnberg P (2005) HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics* 21:1730–1732
- Yang R, Blaileanu G, Hansen BC, Shuldiner AR, Gong DW (2002) cDNA cloning, genomic structure, chromosomal mapping, and functional expression of a novel human alanine aminotransferase. *Genomics* 79:445–450