Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: van Kuilenburg ABP, Tarailo-Graovac M, Richmond PA, et al. Glutaminase deficiency caused by short tandem repeat expansion in *GLS*. N Engl J Med 2019;380:1433-41. DOI: 10.1056/NEJMoa1806627

TABLE OF CONTENTS

Glutaminase deficiency Caused by Short Tandem Repeat Expansion in *GLS*

André B.P. van Kuilenburg[1,2*], Maja Tarailo-Graovac[3,4*], Phillip A. Richmond[5,6*], Britt I. Drögemöller[5,7], Mahmoud A. Pouladi[8,9], René Leen[1], Koroboshka Brand-Arzamendi[10], Doreen Dobritzsch[11], Egor Dolzhenko[12], Michael A. Eberle[12], Bruce Hayward[13], Meaghan J. Jones [5,6], Farhad Karbassi[10], Michael S. Kobor [5,6], Janet Koster[1], Daman Kumari[13], Meng Li[10], Julia MacIsaac[5,6], Cassandra McDonald[14], Judith Meijer[1], Charlotte Nguyen[15,16], Indhu-Shree Rajan-Babu[14], Stephen W. Scherer[15,16,17], Bernice Sim[9], Brett Trost[15], Laura A. Tseng[1], Marjolein Turkenburg[1], Joke J.F.A. van Vugt[18,19], Jan H. Veldink[18,19], Jagdeep S. Walia[20], Youdong Wang[10], Michel van Weeghel[1], Galen E.B. Wright[5,7], Xiaohong Xu[9], Ryan K.C. Yuen[15,16], Jinqiu Zhang[9], Colin J. Ross[5,7], Wyeth W. Wasserman[5,6,14], Michael T. Geraghty[20], Saikat Santra[21], Ronald J.A. Wanders[1,2***], Xiao-Yan Wen[10,22***], Hans R. Waterham[1,2***], Karen Usdin[13***], Clara D.M. van Karnebeek[1,2,5,6***]

* shared 1st authors (with equal contributions)

*** senior authors

1. Amsterdam UMC, University of Amsterdam, Departments of Clinical Chemistry, Pediatrics and Clinical Genetics, Amsterdam Gastroenterology & Metabolism, Meibergdreef 9, Amsterdam, Netherlands
2. On behalf of 'United for Metabolic Diseases', The Netherlands
3. Departments of Biochemistry, Molecular Biology and Medical Genetics, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
4. Alberta Children's Hospital Research Institute, University of Calgary, Calgary, AB, Canada
5. BC Children's Hospital Research Institute, Vancouver, BC, V5Z 4H4, Canada
6. Centre for Molecular Medicine and Therapeutics, BC Children's Hospital Research Institute, University of British Columbia, Vancouver, Canada

7. Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, V6T 1Z3, Canada

8. Department of Medicine, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 117597, Singapore

9. Translational Laboratory in Genetic Medicine (TLGM), Agency for Science, Technology and Research (A*STAR), 8A Biomedical Grove, Immunos, Level 5, Singapore 138648, Singapore

10. Zebrafish Centre for Advanced Drug Discovery, St. Michael's Hospital & University of Toronto

11. Uppsala University, Department of Chemistry-Biomedical Center, Uppsala, Sweden

12. Illumina Inc, San Diego, California 92122, USA

13. Gene Structure and Disease Section, Laboratory of Cell and Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK); National Institutes of Health (NIH), Bethesda MD, USA

14. Department of Medical Genetics, The University of British Columbia, Vancouver, BC V6H 3N1, Canada

15. The Centre for Applied Genomics, Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada

16. Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada

17. McLaughlin Centre, University of Toronto, Toronto, ON M5G 0A4, Canada

18. Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands

19. On behalf of ‘The Project MinE ALS Sequencing Consortium’

20. Department of Pediatrics, Children’s Hospital of eastern Ontario, University of Ottawa, Ottawa, Ontario, Canada

21. Department of Clinical Inherited Metabolic Disorders, Birmingham Children's Hospital, Birmingham, UK

22. Department of Medicine, Physiology, & Laboratory Medicine and Pathobiology; Institute of Medical Science, University of Toronto, Ontario, Canada

Corresponding author: Dr. Clara D.M. van Karnebeek, Department of Pediatrics, Emma Children's Hospital (Room H8-268), Amsterdam University Medical Centers, Meibergdreef 9, 1105 AZ Amsterdam, The Netherlands. Tel: +31 205662113; E-mail: c.d.vankarnebeek@amc.uva.nl

AUTHORS' CONTRIBUTIONS

André B.P. van Kuilenburg, Maja Tarailo-Graovac, Xiao-Yan Wen, Hans R. Waterham and Clara D.M. van Karnebeek designed the study.

Maja Tarailo-Graovac, Phillip Richmond, Britt I. Drögemöller, René Leen, Youdong Wang, Farhad Karbassi, Indhu-Shree Rajan-Babu, Galen EB Wright, Cassie McDonald, Rosalia K. Brand-Arzamendi, Meng Li, Laura A. Tseng, Jagdeep Walia, Michel van Weeghel, Doreen Dobritzsch, Egor Dolzhenko, Michael A. Eberle, Bruce Hayward, Meaghan Jones, Michael Kobor, Farhad Karbassi, Janet Koster, Daman Kumari, Meng Li, Julia MacIsaac, Cassie McDonald, Judith Meijer, Charlotte Nguyen, Indhu-Shree Rajan-Babu, Stephen W. Scherer, Bernice Sim, Brett Trost, Laura A. Tseng, Marjolein Turkenburg, Joke J.F.A. van Vugt, Michel van Weeghel, Jan Veldink, Jagdeep Walia, Youdong Wang, Galen Wright, Xiaohong Xu, Ryan K.C. Yuen, Jinqiu Zhang, The Project MinE ALS Sequencing Consortium, Colin J. Ross, Wyeth W. Wassermann, Michael Geraghty, Janet Koster, Judith Meijer, Marjolein Turkenburg, Saikat Santra, Clara D.M. van Karrnebeek gathered the data.

André B.P. van Kuilenburg, Maja Tarailo-Graovac, Phillip Richmond, Britt I. Drögemöller, René Leen, Youdong Wang, Farhad Karbassi, Indhu-Shree Rajan-Babu, Galen EB Wright, Cassie McDonald, Rosalia K. Brand-Arzamendi, Meng Li, Laura A. Tseng, Michel van Weeghel, Doreen Dobritzsch, Colin J. Ross, Wyeth W. Wassermann, Michael Geraghty, Janet Koster, Judith Meijer, Marjolein Turkenburg, Saikat Santra, Egor Dolzhenko, Michael A. Eberle, Bruce Hayward, Meaghan Jones, Michael Kobor, Farhad Karbassi, Janet Koster, Daman Kumari, Meng Li, Cassie McDonald, Judith Meijer, Charlotte Nguyen, Indhu-Shree Rajan-Babu, Stephen W. Scherer, Bernice Sim, Brett Trost, Laura A. Tseng, Marjolein Turkenburg, Joke J.F.A. van Vugt, Michel van Weeghel, Jan Veldink, Jagdeep Walia, Youdong Wang, Galen Wright, Xiaohong Xu, Ryan K.C. Yuen, Jinqiu Zhang, The Project MinE ALS Sequencing Consortium analyzed the data

Maja Tarailo-Graovac, André B.P. van Kuilenburg, Phillip Richmond, Britt Drögemöller, Karen Usdin, Mahmoud Pouladi, Xiao-Yan Wen, Hans R. Waterham and Clara D.M. van Karnebeek wrote the paper.

All authors reviewed the paper.

All authors vouch for the data and the analysis.

All authors agreed to publish the paper.

SUPPLEMENTARY METHODS

Exome sequencing

Duo-exome sequencing (mother-proband) and trio-exome sequencing (mother-father-proband) analysis for the Families 1 and 2 of the TIDEX study were performed using the TruSeq DNA PCR Free (350) kit and Illumina HiSeqX (Macrogen, Korea) through the TIDEX gene discovery project (UBC IRB approval H12-00067). For exome analysis of Family 1 and Family 2, we updated a previously described pipeline for the identification of causal variants from next-generation sequencing data (details below)[1]. Executable shell scripts for the analysis of the exome data are provided online (https://github.com/Phillip-a-richmond/AnnotateVariants/GLS_Manuscript/). The output candidate variant lists (CVLs) for both Family 1 and Family 2 were manually inspected for variants which fit the phenotype and inheritance pattern. Analysis of variants which fit the inheritance patterns did not identify any fully-penetrant candidates that match the phenotype. Therefore, we examined variants which are identified in the affected proband and meet filters for quality (overlapping Illumina Platinum Genome Confident Regions, not overlapping segmental duplications, DP[*] <= 300, AD[*:1] >= 10, genotype quality >= 30), pathogenicity (GEMINI HIGH/MED impact) and rarity (global gnomAD AF<0.01, global gnomAD Homozygotes <=10). Family 3 from Care4Rare was subjected to trio-ES sequencing. Exome target enrichment was performed with the Agilent SureSelect 50 Mb (V3) All Exon Kit; samples were sequenced on the Illumina HiSeq 2000 platform, multiplexing three samples per lane. After removal of duplicate reads, the mean coverage of coding sequence regions ranged from ×70 to ×200. Alignment and variant annotation were performed by the FORGE informatics team at each STIC, using comparable analytical pipelines with publicly available tools and custom scripts as described previously[2].

Detailed exome/genome analysis pipeline

An in-house, open-source bioinformatics pipeline was used to analyze the sequencing from families 1 and 2.

The updated pipeline is as follows (default options unless specified, citations in Table S2):

1) Map pair-end reads to GRCh37 using BWA mem (v0.7.12) with option –M
    a. (http://www.bcgsc.ca/downloads/genomes/9606/hg19/1000genomes/bwa_ind/genome/)
2) Convert SAM files to BAM, sort, and index using Samtools
3) Duplicate marking using Picard MarkDuplicates
4) Local indel realignment using GATK RealignerTargetCreator and IndelRealigner
5) GVCF calling using GATK HaplotypeCaller (with –emitRefConfidence GVCF option)
6) Multi-sample variant calling from GVCF files using GATK GenotypeGVCFs
7) Variant normalization and splitting using VT normalize and decompose
8) Variant annotation against genes using SNPEff
9) Compression and indexing of variants using bgzip and tabix
10) Variant filtering using BCFTools filter with soft filter: --include 'FORMAT/AD[*:1]>=10 && FORMAT/DP[*] < 300'
11) Variant annotation against various databases (Table S2) with VCFAnno
12) Conversion of annotated VCF to MySQL-lite database with VCF2DB
13) Queries against MySQL-lite database with GEMINI for inheritance patterns, rare and potentially damaging variants
14) Formatting and addition of gene-based databases using GeminiTable2CVL.py

Genome sequencing

Singleton-genome sequencing was performed for the proband of the Family 1 in search of a second variant in the *GLS* expected to be inherited from mother. For the genome sequencing, we also used the TruSeq DNA PCR Free (350) kit and Illumina's HiSeqX sequencer at Macrogen, Korea. The genome sequencing data was re-analyzed using the same approach described above and again the single missense variant in the *GLS* was considered the best candidate. Subsequent manual inspection of the cis-regulatory elements associated with *GLS* using Integrative Genomics Viewer (IGV v.2.4.10) revealed a potential GCA repeat within the 5' UTR. The recently published Expansion Hunter[3] software (details below) was used to analyze the locus in further detail and identified the GCA repeat expansion with at least 94 copies on the chromosome the proband had inherited from mother, while the copy inherited from father had 8 GCA copies as well as the original missense variant.

STR Genotyping from Genome Sequencing using Expansion Hunter

For Patient 1, Expansion Hunter (v.2.5.3) was run on the duplicate removed, realigned, genome sequencing mapped reads (described above). Since Expansion Hunter can detect larger repeat expansions[3] which can be disruptive when located within the UTRs of genes, we decided to use it for this case. Because expanded GCA/CAG/AGC repeats may be present in multiple loci, we ran Expansion Hunter in a conservative mode that only analyzes reads that can be confidently placed at the targeted locus (chromosome 2:191745599-191745646 build GRCh37). This is enabled by setting "CommonUnit" to "true" in the input repeat-specification file (available online at https://github.com/Phillip-a-richmond/STR_Analysis/blob/master/GenerateTargetLoci/GLS_NoOffTarget.json). In this mode, Expansion Hunter produces a size estimate bounded by the length of the sequenced

fragment. Thus, for repeats whose size exceeds fragment length, Expansion Hunter produces a lower bound size estimate.

STR Genotyping in control populations

We further examined three independent series totaling 8,295 individuals (16,590 alleles) for the presence of the *GLS* repeat expansion (same methodology as for Patient 1). Series 1 included sequence data from 441 cell lines from the Coriell Institute for Medical Research (www.coriell.org). For these samples, DNA was prepared using PCR-free library preparation kits, and sequenced by the Illumina HiSeq X platform (2 x 150bp paired-end reads). Information about these samples and links for downloading the raw sequence data can be obtained at https://github.com/Illumina/Polaris. Series 2 included DNA samples from 1,658 unrelated individuals (unaffected parents of children diagnosed with autism spectrum disorder) from the MSSNG genomic database[4]. DNA was derived from whole-blood or lymphoblast-derived cell lines, prepared using PCR-free library preparation kits, and sequenced by the Illumina HiSeq X platform (2 x 150 bp paired-end reads). Further details about MSSNG and samples can be found at https://mssng-edge.dnastack.com. Series 3 included DNA samples from 6,196 unrelated individuals with and without ALS from Project MinE[5]. DNA was derived from whole-blood, prepared using PCR-free library preparation kits, and sequenced by the Illumina HiSeq 2000 platform (2 x 100 bp paired-end reads) and HiSeq X platform (2 x 150bp paired-end reads). Further details about Project MinE and the samples can be found at http://databrowser.projectmine.com/.

Sanger sequencing

Sanger sequencing confirmation of the missense and frameshift variant was performed in the patient and unaffected parents for Families 1 and 3 using BigDye® Terminator v3.1 Cycle

Sequencing chemistry (Life Technologies, Carlsbad, CA, USA). Subsequent capillary electrophoresis was performed by the CMMT/CFRI DNA Sequencing Core Facility using a Prism 3130xl 16-capillary automated genetic analyzer (Applied Biosystems, Foster City, CA, USA).

PCR and Sanger sequence analysis of the 5' UTR

PCR and subsequent Sanger sequencing analyses were performed to investigate variation within the 5' UTR. This region, which includes a GCA repeat, was amplified using the conditions and primers previously described by Mayer *et al*[6]. Examination of the sequence chromatograms in the index from Family 1 revealed discrepancies between the Sanger and genome sequencing data such that high quality heterozygous variants called from the genome sequencing data were observed to be homozygous in the Sanger sequencing data. It was hypothesized that these discrepancies may have arisen due to an expansion of the GCA repeat, which prevented the amplification of the allele carrying the GCA expansion. Attempts to amplify this region in the index from Family 2 were unsuccessful, suggesting that this individual may have inherited both copies with the GCA expansion. To confirm the presence of the GCA repeat expansion, a triplet repeat-primed PCR was performed in these two families, as well as an unrelated third family, as described below. Capillary electrophoresis was performed by the CMMT/BCCHRI DNA Sequencing Core Facility using a Prism 3130xl 16-capillary automated genetic analyzer (Applied Biosystems). Fragment analysis and scoring was performed with Fragman[7].

Triplet repeat-primed PCR

Triplet repeat-primed PCR was used to confirm the presence of the repeat expansions.

A triplet repeat-primed PCR was performed with the following primers (Supplementary Figure S12)

i. A locus-specific fluorescently labelled forward primer (5'- AGA GAG GTC GAG GAT TTG AGC CAA TCG -3') to allow for the sizing of the amplified fragments

ii. A repeat annealing reverse primer (5'- TGC TCT GGA CCC TGA AGT GTG CCG TTG ATA TGC TGC TGC TGC TGC TGC -3') to allow for the binding of the primer to different regions within the trinucleotide repeat. This primer also includes 30bp of unique sequence that do not match to the human genome

iii. A tailed reverse primer (5'- TGC TCT GGA CCC TGA AGT GTG CCG TTG ATA -3'), which contains the 30bp of unique sequence present in the repeat annealing primer to allow for the continued amplification of the fragment when the repeat annealing primer has been depleted, as described previously[8].

PCR amplifications were carried out in a total reaction volume of 15 µl, with each reaction containing 50 ng of gDNA, 0.6 µM of the locus-specific forward primer, 0.06 µM of the repeat annealing reverse primer, 0.6 µM of the tail reverse primer, 0.2 mM dNTPs, 2 mM $MgCl_2$, 1 M Betaine, 5U AmpliTaq Gold® DNA Polymerase and 1X AmpliTaq Gold 360 Buffer. The reaction cycle conditions consisted of an initial denaturation step at 95°C for 6 minutes, followed by 40 cycles of 45 seconds denaturation at 95°C, 45 seconds annealing at 68°C and 6 minutes extension at 72°C.

Generation of GLS-deficient Flp-In™ 293 cells by CRISPR-Cas9

The *GLS* gene in Flp-In™ 293 cells (ThermoFisher Scientific, Waltham, MA, USA) was disrupted using the CRISPR-Cas9 genome editing protocol described before[9]. To this end, two complementary oligonucleotides (sequence AGTGCTAAAAAGCAGTCTGG) were designed coding for a guide RNA upstream of a TGG PAM site in exon 4 (c.726_728) of the *GLS* gene (NM_014905.4) using the online CRISPR design tool (http://crispr.mit.edu/). The two oligo's were annealed and subsequently cloned into the pX458(-pSpCasq(BB)-2A-GFP) plasmid

followed by Sanger sequencing of the insert to confirm the correct sequence. Flp-In™ 293 cells were transfected with 2 µg plasmid, and single GFP-positive cells were sorted into wells of a 96 well-plate using FACS flow cytometry (S800H Cell Sorter, Sony). After 3-4 weeks, DNA was isolated from the expanded single colonies and exon 4 of the *GLS* gene was PCR-amplified using Phire animal tissue direct PCR kit (ThermoFisher Scientific) according to the manufacturer's instructions and subsequently Sanger sequenced. For further experiments we used Flp-In™ 293 cells that were heterozygous for two frame-shift mutations [c.722_725del (p.Gln241Leufs*27) and c.722_725dup (p.Gly243Valfs*57)] in *GLS* and confirmed by immunoblotting to completely lack GLS protein.

Transient expression of GLS

Wild-type Flp-In™ 293 cells and GLS-deficient Flp-In™ 293 cells were cultured in Dulbecco's modified eagle's medium with 4.5 g/l glucose, 25 mM Hepes and 584 mg/l L-glutamine (Dulbecco, Lonza, Basel, Switzerland), supplemented with 10% fetal bovine serum, 100 U/ml penicillin, 100 µg/ml streptomycin and 250 µg/ml fungizone at 37 °C in a humidified 5% CO2 incubator. For transient transfections, cell cultures were setup in six-well plates 24 h prior to transfection. GLS-deficient Flp-In™ 293 cells were transfected with pcDNA3.1Zeo-GLS (wild-type or variant) using X-treme GENE HP DNA Transfection reagent (Roche, Basel, Switzerland). 24 hours after transfection, cells were harvested and washed with phosphate-buffered saline (PBS). After centrifugation at 1000× g for 5 min at 4 °C, cell pellets were immediately frozen in liquid nitrogen and stored at −80 °C until use. Parental vector (pcDNA3.1Zeo) without insert was transfected as a negative control.

Immunoblot analysis of glutaminase

Protein lysates (20 µg) were separated on a 4-12% PAGEr™ EX polyacrylamide gel (Lonza, Breda, The Netherlands) and transferred to nitrocellulose (GE Healthcare Life Sciences, Marlborough, MA, USA). Membranes were blocked using 50% Odyssey blocking buffer (LI-COR, Lincoln, NE, USA) and 50% PBS. Subsequently, blots were incubated for one hour with a 1:1000 dilution of rabbit antibody against Glutaminase C (GeneTex, Irvine,CA, USA; GTX-131263) and a 1:2000 dilution of mouse antibody against α-tubulin (Sigma-Aldrich, Darmstadt, GER; T6199) in blocking buffer (50 % Odyssey blocking buffer, 50 % PBS and 0.1 % Tween20). Membranes were washed four times with PBS containing 0.1% (v/v) Tween20 and then incubated for one hour with a 1:5.000 dilution of IRDye®800 Goat Anti-Rabbit and IRDye®680 Donkey Anti-Mouse (both LI-COR) in the same blocking buffer, as used for the primary antibodies. After rinsing the membrane four times with PBS containing 0.1% (v/v) Tween20 the blots were scanned and band intensities analyzed using the LI-COR Odyssey infrared imaging system.

cDNA analysis and Quantitative RT-PCR

For cDNA analysis, RNA was isolated from primary skin fibroblasts of the patients using TRI Reagent (Sigma-Aldrich) extraction or from PAXgene Blood RNA Tubes using the Qiagen PAXgene Blood RNA system and reverse-transcribed into cDNA using the QuantiTect Reverse Transcription Kit (Qiagen, Hilden, GER) before Sanger-sequencing (primer sequences available on request).

For Quantitative RT-PCR, the LightCycler 480 SYBR Green I Master kit (Roche, Basel, CH) was used and the Lightcycler ® 480, Instrument II (Roche, Basel, CH). The N_0 values of target genes were normalized to the geometric mean of the reference genes RPS14 (in fibroblasts) and SDHA (in blood), which were determined most stable in expression and thus

best suitable for normalization out of nine assessed reference genes using the Normfinder algorithm [10]. For data analysis, Light Cycler 480 software release 1.5.0 and LinRegPCR version 2014.2 were used. Primer sequences used for *GLS* and reference genes are available on request.

Analysis of the structural context of the Pro313Leu and Tyr308* mutations

Analysis of available crystal structures of human glutaminase isoforms and manual introduction of the Pro313Leu amino acid exchange and the p.Tyr308* in the structure with PDB accession code 3unw were performed with WinCoot[11]. Figures S4 and S5 were generated using PyMol[12].

Glutaminase activity assay

Cell pellets of fibroblasts and peripheral blood mononuclear cells were suspended in phosphate-buffered saline to a final protein concentration of 3 mg/ml. All homogenates were sonicated once (40 Joule/ Watt·sec, output 7,5 Watt). The GLS activity was determined in a reaction mixture containing 100 mM Tris-HCl (pH 9.0), 100 mM potassium phosphate (pH 8.5), 20 mM glutamine, 0.2 mM EDTA. The reaction mixture and the sample were equilibrated separately at 37 °C in a stirring water bath for 2 min and the reaction was started by addition of the 10 µl sample resulting in a final volume of 100 µl. After 40 min of incubation, the samples were placed on ice and the reaction catalyzed by GLS was terminated by adding 10 µl 2 M HCl into the reaction tube. The reaction mixture was centrifuged in a microfuge (10,000 x g, 4 °C, for 5 min) to remove the protein. The supernatant (110 µl) was mixed with 10 µl 2M KOH/0.6M MOPS.

The amount of glutamate, formed by GLS, was determined in an assay mixture containing 100 mM CHES (pH 9.5), 200 mM hydrazine, 10 mM ADP, 50 mM NAD and 25 µl sample. The assay mixture was equilibrated at 37 °C for 7.5 min and the reaction was started by addition of 50 µl glutamate dehydrogenase (16 mg/ml, 16.4 U/mg, Roche 10197734001)

resulting in a final volume of 250 µl. The increase in absorbance at 340 nm was continuously monitored using a Cobas/Fara spectrophotometer. Quantification of the amounts of glutamate was performed by comparison with external glutamate standards (0-2 mM).

Fluxomics

Fibroblasts were routinely cultured in DMEM with 25 mM HEPES, 80 U/ml penicillin/streptomycin, 0.25 µg/ml fungizone and 10% (v/v) fetal calf serum (FCS). Cells were grown at 37°C, 5 % CO2 in 95 % humidified air. Cells were plated in 6-wells plates at a density of 200,000 cells per well in a total volume of 2 ml DMEM supplemented with 10% FCS and were allowed to adhere overnight. Subsequently, the medium was replaced by DMEM medium, deprived of glucose, glutamine, sodium pyruvate and phenol red (Gibco, A14430-0, Landsmeer, the Netherlands) and supplemented with 25 mM HEPES, 80U/ml penicillin/streptomycin/fungizone, 50 µM carnitine, 5 mM D-glucose - U-$^{13}C_6$ (Buchem, CLM-1396) or 1 mM L-glutamine -$^{13}C_5$, $^{15}N_2$ (Buchem, CNLM-1275) or 50 µM Oleic Acid-$^{13}C_{18}$ (Cambridge Isotope Laboratories, CLM-460-PK). The final concentration of added labeled or unlabeled glucose, glutamine and oleic acid to each well was 5 mM, 1 mM and 50 µM, respectively. After 18 hours of incubation, the medium was removed and the 6 well-plates were washed on ice three times with ice-cold 0.9% (w/v) NaCl. Subsequently, 250 µl methanol (-20 ^{o}C) was added to each well followed by 250 µl ice-cold Milli-Q water. The cell suspension was transferred to a 2 ml Eppendorf cup and each well of the 6 wells plate was washed with 500 µl 50% methanol (-20 ^{o}C) and added to the cell suspension in the 2 ml Eppendorf cup. Chloroform (1 ml) was added to each sample and the solution was sonicated (40 Joule/Watt·sec, 7.5 Watt output). After centrifugation [20,000 g (4 ^{o}C)], the "polar" top layer was transferred to a new 1.5 ml tube and dried to dryness in a vacuum concentrator. (60 ^{o}C, 1.5 h). The final

pellet was dissolved in 50 µl 60% methanol, centrifuged at 20,000 g (4 ºC) and 40 µl of the supernatant was saved for further analysis.

Mass Spectrometry (MS)

For the analysis, we used a Thermo Scientific ultra-high pressure liquid chromatography system (ThermoFisher Scientific) coupled to Thermo Q Exactive (Plus) Orbitrap mass spectrometer (ThermoFisher Scientific). The autosampler was kept at 10°C during the runs and 5 µl of sample was injected on the analytical column. The chromatographic separation was established using a SeQuant ZIC-cHILIC column (PEEK 100 x 2.1mm, 3.0µm particle size, Merck, Darmstadt, Germany) and kept at 15°C. The flow rate was 0.250 mL/min. The mobile phase was composed of (A) 9/1 acetonitrile/water with 5 mM ammonium acetate; pH 6.8 and (B) 1/9 acetonitrile/water with 5 mM ammonium acetate; pH 6.8, respectively. The LC gradient program was: beginning with 100% (A) hold 0-3 min; ramping 3-24 min to 20% (A); hold from 24-27 min at 20% (A); ramping from 27-28 min to 100% (A); and re-equilibrate from 28-35 min with 100% (A). The MS data were acquired in negative mode at full scan range at 140.000 resolution. Interpretation of the data was performed in the Xcalibur software (ThermoFisher Scientific). Statistical analysis and visualization of the acquired data were performed using ggplot, ropls and mixOmics packages in an R environment. We were able to identify over 10.000 compounds untargeted and are able to annotate >100 metabolites.

Zebrafish gene knockdown models

All zebrafish (*Danio rerio*) experiments were conducted under approved animal protocol ACC660 from St. Michael's Hospital (Toronto, Ontario, Canada). The zebrafish were housed in the Li Ka Shing Knowledge Institute (St. Michael's Hospital, Toronto, Ontario, Canada) research vivarium and maintained and staged as previously described[13]. In short, the fish were

housed under a 14 h light:10 h dark cycle at 28°C. Embryos were produced by pair mating and raised in 1× E3 embryo medium (5 mM NaCl, 0.17 mM KCl, 0.33 mM $CaCl_2$, 0.33 mM $MgSO_4$. Wild type TU strain was used in this study. The collection of fertilized eggs was obtained through pair-wise breeding according to the standard method previously described[13].

To knock down gene expression, we injected morpholino (MO) oligonucleotides targeting the first ATG codon for *glsa, gls-like (glsl)* genes and a standard control MO (ctrl MO) was also used. The MO sequences were as follows:

glsa-atg MO 5'-CCCTCGAAAGCCTAAAGTGTAACAT-3',

glsl-atg MO 5'-AAGCTGTTCTCCATGATAAAGCCGT-3',

and standard control MO 5'-CCTCTTACCTCAGTTACAATTTATA-3'.

MOs were designed by Gene Tools, LLC. Both MOs were injected individually or in combination into 1-cell-stage zebrafish embryos. We injected individually 4 nl of *glsa-atg* (8 ng/µl) and 4 nl of *glsl-atg* (8 ng/µl) MOs and in combination 4 nl of *glsa-atg* (4 ng/µl) / *glsl-atg* (4 ng/µl).

GLS repeat PCR assay

GLS repeat PCR assays were used to confirm the presence of the repeat expansions.

The Repeat PCR assay was carried out as described previously for the Fragile X repeats[14] with minor modifications.

The primers used in this case were:

GLS-F: 5'-AGTTCAGCGGCCGCGAGAAATTTGACTGCCCTAATCGCCAATC-3'

GLS-R: 5'-CAAGTCGCGGCCGCGGGCGGTGAGATGACAGAAGAGGAAAG-3'

Briefly, 600 ng of genomic DNA was digested in 40 µl of 1 x PCR buffer (50 mM Tris.Cl pH 9.0, 1.5mM $MgCl_2$, 22mM $(NH_4)_2SO_4$, 0.2% Triton) with 1 µl of FastDigest-Pst I (Life Technologies) overnight at 37°C. Ten microliters of the digested DNA was then combined with

30 µl of a reaction mix containing 0.67 µM each of GLS-F and GLS-R, 3.3 M betaine (Sigma, St. Louis, MO), 2.67% DMSO (Sigma), 0.67 mM each dNTP (New England Biolabs, Ipswich, MA) and 0.8 U Q5 Hot Start polymerase (New England Biolabs) in 1 x PCR buffer. The samples were loaded into a preheated PCR block (> 70ºC) and the DNA amplified using the following cycling protocol: 1 x (98ºC for 3 min), 30 x (98ºC 30 sec, 50ºC 30 sec, 72ºC 210 sec) followed by a final incubation at 72ºC for 10 min. The DNA was then precipitated using a standard ammonium acetate/ethanol procedure and resuspended in 10 µl of TE buffer. Five microliters of this DNA was used for analysis on agarose gels that were then post-stained with ethidium bromide or SYBR® Gold (ThermoFisher Scientific) according to standard procedures.

Chromatin analysis

Chromatin immunoprecipitation (ChIP) assays were performed as described before[15] using a ChIP assay kit from EMD-Millipore. The following antibodies were used in ChIP assays, normal rabbit IgG (EMD-Millipore, 12-370, 5 µg/IP), acetyl H3 (Abcam, ab47915, 10 µl/IP), H3K9me3 (Abcam, ab8898, 4 µg/IP) and H3K4me3 (Millipore, 04-745, 4 µl/IP). To prepare chromatin for immunoprecipitation, cells were fixed with 1% formaldehyde for 10 minutes at room temperature and lysed as per the kit manufacturer's instructions. The chromatin was sonicated into <500 bp fragments using a Bioruptor® sonicator (Diagenode, Denville, NJ) and 20 µg chromatin was used for each IP. Real-time PCRs on the immunoprecipitated DNAs were carried out in triplicate in 20 µl final volume using the Power SYBR™ Green PCR master mix (ThermoFisher Scientific) and 2 µl of DNA using StepOne Plus™ Real-time PCR system (Applied Biosystems). For amplification of *GLS* promoter region upstream of the repeat, 100 nM each of primer GLS-F2 (5'-GATTTGAGCCAATCGCAGC-3') and GLS-R1 (5'-GGCTAGAGACCTTCAGCGCT-3') were used. *GAPDH* was amplified with 200 nM each of

primer hsGAPDH exon1F1 (5′-TCGACAGTCAGCCGCATCT-3′) and hsGAPDH intron1R1 (5′-CTAGCCTCCCGGGTTTCTCT-3′). For quantitation, the comparative threshold (Ct) method was used. Enrichment over 100% of input was calculated and normalized to *GAPDH*.

Cloning *GLS* promoter into luciferase vector

To amplify the *GLS* promoter with [GCA] expansion, 200 ng of purified gDNA from patient fibroblasts was used as template DNA with KOD Xtreme (Novagen, #71975) supplemented with 8% DMSO (final v/v). Cycling conditions were as follows: initial denaturation at 96°C for 5 min, followed by 7 cycles of 96°C for 45s, 70°C for 30s, 72°C for 2 min; and 33 cycles of 96°C for 45s, 58°C for 30s and 72°C for 2mins and a final elongation at 72°C for 10 min. Primers used for amplification were GLS-L and GLS-R (**Table S3**), amplifying 1022nt upstream and 164nt downstream of [GCA] repeats. Amplicons spanning 1.5 – 2kb (100-300 GCA repeats) were cloned into pCR-Blunt (Invitrogen, K2750-20) and screened for repeat length with EcoRI and Sanger sequenced with primers T7F and M13R (**Table S3**) by Axil Scientific Pte Ltd (Singapore). Promoters with 104, and >201 GCA repeats were then ligated into pGL3 and sized with EcoRI and BglII. Genomic DNA from H9 hESCs was used as a WT control (13, 14 GCA). Plasmids were also sequenced with primers F1ori, LucR and GLS-L2 for verification (**Table S3**). All amplicons and digested DNA were visualized on the Geldoc XR system (Bio-Rad) (Supplementary figure S11).

Transient transfection and luciferase reporter assays

The HEK293 cells in 24-well plate were co-transfected with 0.72 μg Firefly plasmid-constructed *GLS*-promoter and 0.08 μg Renilla plasmid phRL-TK using Lipofectamine 2000 (Life Technologies) according to the manufacturer's instructions. After 48 h post-transfection, the cells were harvested and lysed in 100 μl passive lysis buffer (Promega)

for 20 min at room temperature. 10 µl of the lysate was subjected to Dual-luciferase Reporter Assay (Promega, E1910). The firefly luciferase activity and Renilla luciferase activity were measured sequentially using a Spectra Max M5 reader (Molecular Devices). The promoter activities were expressed as ratios of Firefly luminescence signal versus Renilla luminescence signal.

Pyrosequencing

Bilsulfite PCR pyrosequencing was performed using the PyroMark Q96 Vacuum Prep Workstation (Qiagen) and the PyroMark Q96 ID pyrosequencer (Qiagen), as previously described.[16] In short, bisulfite-treated gDNA (derived from fibroblasts in the patients from family 1 and 2 and blood from the patient in family 3) was amplified using the following primers:

F1: 5'-AAGGAGAGTTGGATAGGTT-3'

R1: 5'-biot-CCAAATAACTACCCTTTCCTCTT-3'

S1: 5'-GTTTAATAGGGGAGGAT-3'

F2: 5'-biot-AGGTTAAAGAGTAGGGGAGGGTTAGA-3'

R2: 5'-AACCCAAAACCCCTAATACC-3'

S2: 5'-CAAAAAAAAAATCCAAAATTTAAACCAAT-3'

DNA methylation levels of CpGs located at chr2 (GRCh37): 191745753, 191745757, 191745763, 191745767, 191745477, 191745482, 191745486 and 191745495 were examined in all three patients, as well as a control individual, and compared to measurements of DNA with known methylation levels (0, 25, 50, 75, 100%) using the Pyromark CpG 1.0.11 software (Qiagen).

SUPPLEMENTARY RESULTS

Patients' clinical and metabolic phenotype

Patient 1 is a female child, born at term to non-consanguineous Caucasian parents, who had evidence of early gross motor and speech delay. She sat unsupported only at 11 months and crawled by 18 months. When first assessed at 3 years of age, she had a broad-based, ataxic gait with hands held. Patient 2 is a male child, born at term to non-consanguineous Caucasian parents, who had evidence of early gross motor and speech delay. He sat at 13 months and crawled at 17 months. When first assessed at just under 3 years of age he could only walk with support; recently at age 6 years, his gait has become ataxic. Patient 3 is a female child born to non-consanguineous parents, followed from 2 years of age due to gross motor delay. She sat and crawled at usual time but did not stand until 3 years. She had a febrile seizure at 3 years of age. Brain MR imaging revealed no diagnostic abnormalities in all three patients. Each child showed progressive neurological deterioration: on most recent examination, patient 1 at age 7 years relies on a wheelchair due to an ataxic gait, patient's ataxia worsened, while patient 3 at age 12 years requires a walker, has hypotonia, tremor, and dysarthria with severe ataxia. The repeated MRI-brain shows cerebellar atrophy (Fig. S1).

Metabolic investigations revealed normal urinary organic acids (including orotic acid) and acylcarnitines but a persistently and significantly elevated plasma glutamine of 1800 μM (Patient 1), 2000 μM (patient 2) and 2209 μM (patient 3; reference range: 405-781 μmol/L). There was no observed hyperammonaemia nor any history of hyperammonaemic decompensations in the three patients. Urea cycle disorders gene panel sequencing was undertaken and revealed a single, novel and likely pathogenic *CPS1* variant in patient 1 only. Subsequent liver biopsy demonstrated normal carbamoylphosphate synthase activity in patient 1, ruling the *CPS1* out carbamoylphosphate synthase deficiency as the cause for the child's hyperglutaminaemia. Treatment of the patients with oral sodium phenylbutyrate at 250

mg/kg/day did not result in a sustained reduced plasma glutamine level nor to any observed clinical improvement.

Zebrafish phenotype

Zebrafish has two *GLS* genes *glsa* and *glsb*. It also has a *GLS*-like gene (*glsl*). Both *glsa* and *glsl* are strongly expressed in the zebrafish brain. However, *glsb* is not expressed in the brain but only expressed in the pectoral fin bud and the *in situ* expression data of the zebrafish *gls*-genes are already published[17]. Because our patients have neurological phenotypes (walking and speech delay, and encephalopathy), we evaluated the gene knockdown effect of *glsa* and *glsl* as well as *glsa*/*glsl* double knockdown.

The translation start codons of the zebrafish *glsa* and *glsl* genes are located in exon-1 or exon-2, respectively. We designed morpholinos (MOs) to block the translation of the *glsa* or *glsl* protein. After individual MO microinjection or co-injection of *glsa* and *glsl* (half dose of individual MO injection), we observed a consistent phenotype, including smaller body size, curved body and cardiac edema. The phenotype could be classified according to three categories: severe (20-28% of embryos), medium (22-34%) and mild (43-66%) (Figure S7).

SUPPLEMENTARY FIGURES

Figure S1

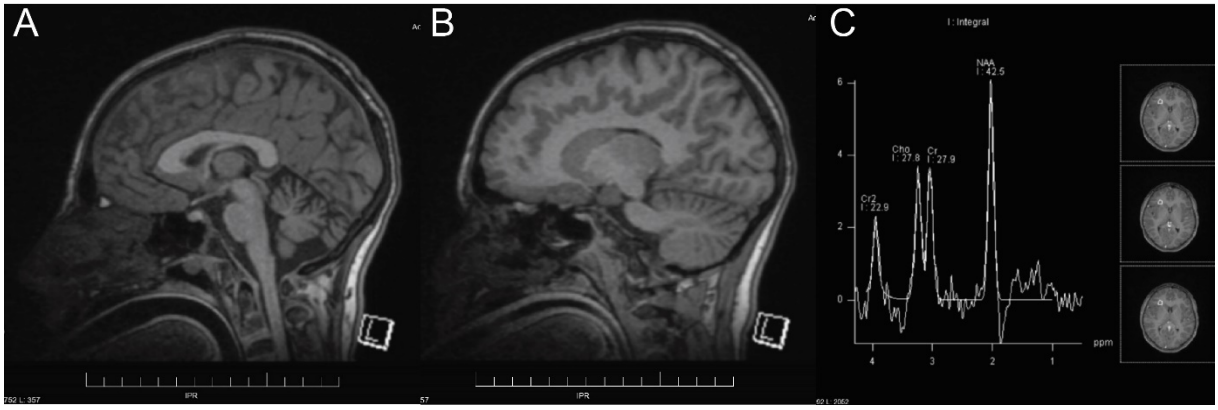

Fig. S1. **MRI-brain scan of patient 3 at age 11 years.** Brain MRI sagittal views (A&B) shows moderate decreased cortical volume affecting the lateral cerebellar hemispheres and cerebellar vermis, with ex-vacuo dilatation of the 4th ventricle. Cerebral volume is normal. Both high (135) and low (60) TE MR spectroscopy (C) were normal with no increase in Glx (glutamate and glutamine).

Figure S2

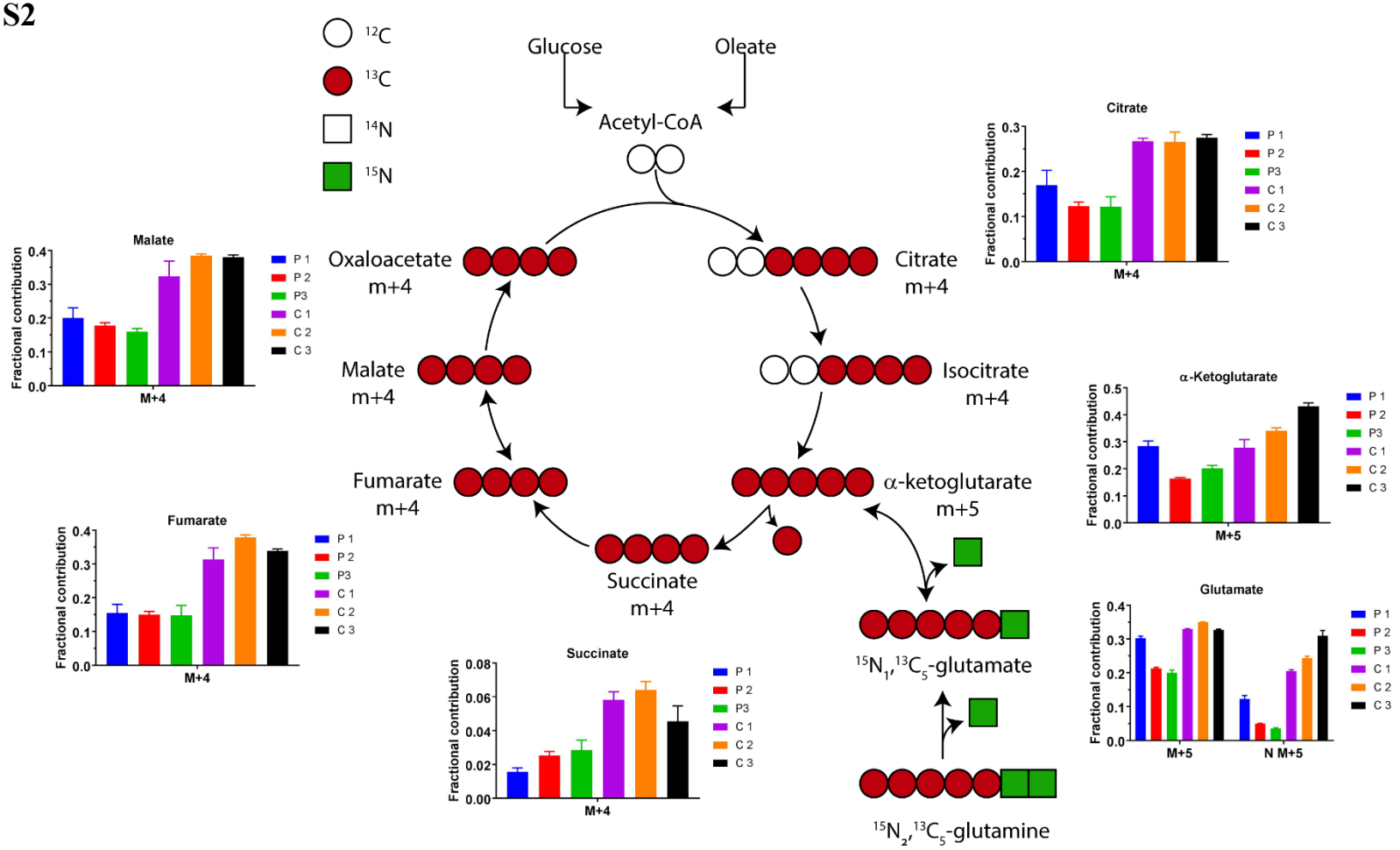

Fig. S2. Flux analysis of stable-isotope labeled glutamine. Mass isotopomer distributions of metabolites of the citric acid cycle after introduction of fibroblasts with $^{15}N_2$,$^{13}C_5$-glutamine. The mass isotopomer distributions were corrected for natural isotope abundances using MIDA

Figure S3

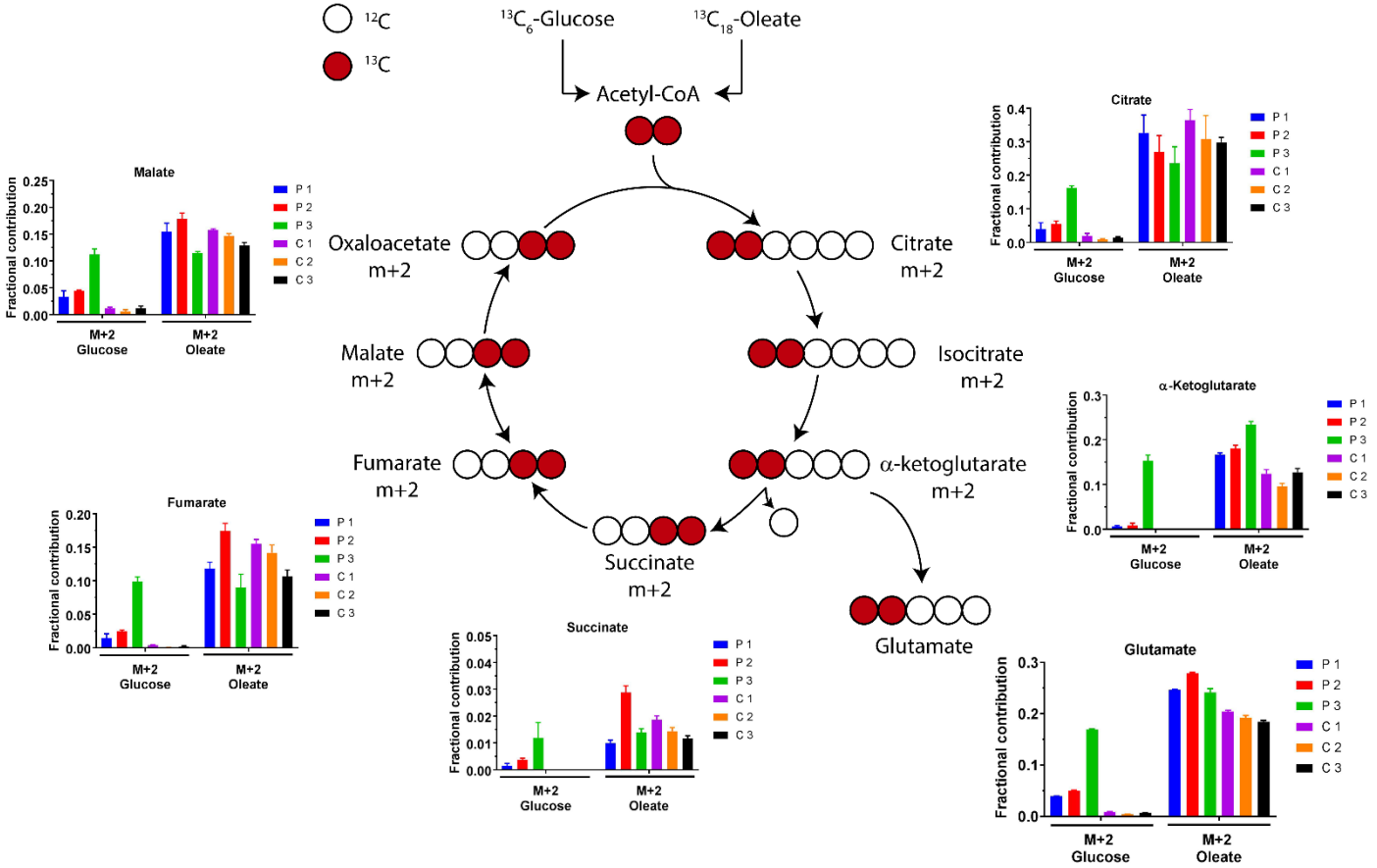

Fig. S3. Flux analysis of stable-isotope labeled glucose and oleic acid. Mass isotopomer distributions of metabolites of the citric acid cycle after introduction of fibroblasts with $^{13}C_6$-glucose and oleic acid or $^{13}C_{18}$-oleic acid and glucose. The mass isotopomer distributions were corrected for natural isotope abundances using mass isotopomer distribution analysis.

Figure S4

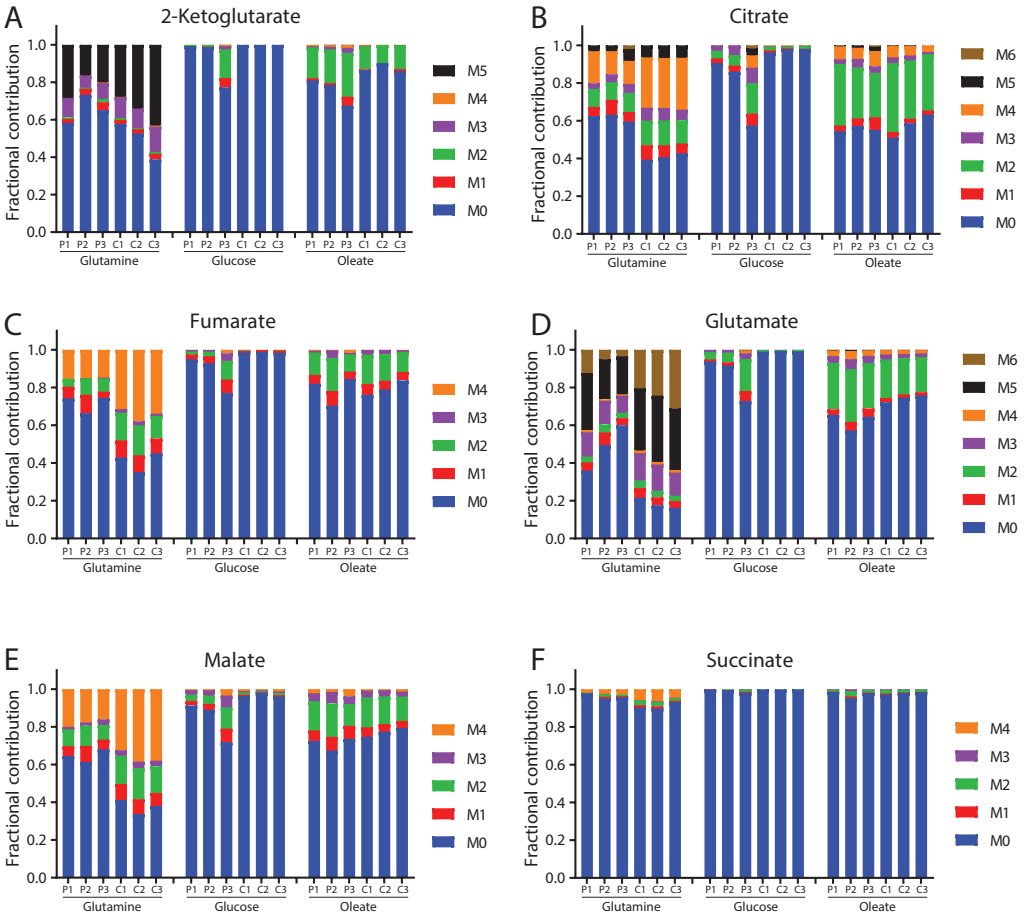

Fig. S4. Fractional contribution of stable-isotope labeled glutamine, glucose and oleic acid in TCA cycle metabolites. Fractional contribution was calculated with mass isotopomer distribution analysis (MIDA) of different stable-isotope-labeled metabolites from the TCA cycle after incubation of fibroblasts from three GLS deficient patients and three controls with either with $^{15}N_2$,$^{13}C_5$-glutamine, $^{13}C_6$-glucose or $^{13}C_{18}$-oleic acid.

Figure S5

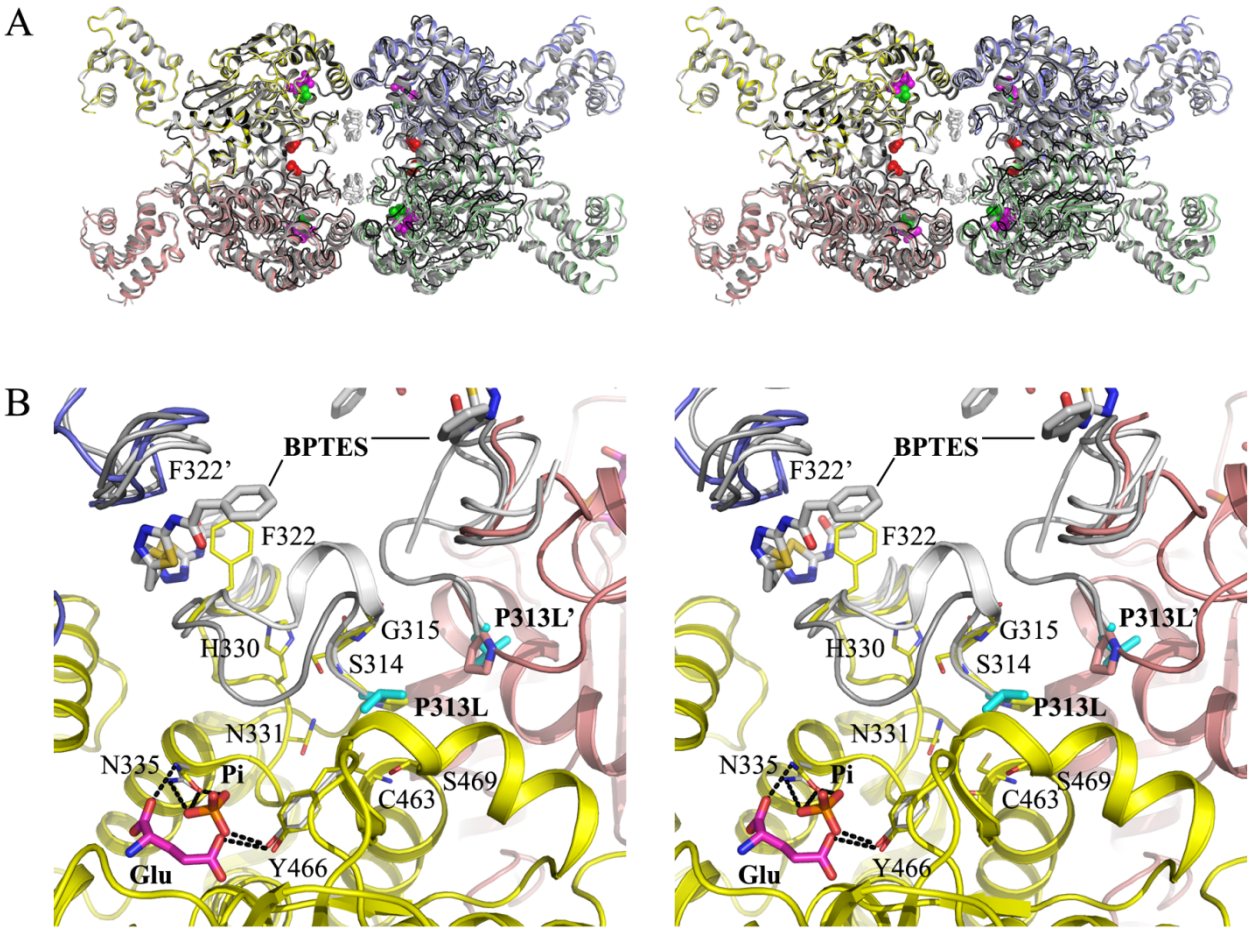

Fig. S5: The structural context of the P313L mutation site. (**A**) Stereo view of the superimposed crystal structures of GLS isoforms in complex with activators and inhibitors that highlight distinct conformations of the gating loops and location of Pro313. For full-length GLS (GAC isoform) in complex with the GLS reaction product glutamate (PDB accession code 3unw[18]), the distinct subunits of the homotetrameric enzyme are colored yellow, salmon, blue and green, respectively. The four glutamate molecules shown as spheres in magenta indicate the location of the active sites. The location of the P313L mutation site near the dimerization and tetramerization interfaces is highlighted by the stick model of Pro313 in red. It directly precedes the gating loop that is found disordered or adopting alternate conformations in

available crystal structures of glutaminase isoforms[18-24] (e.g. the structure of GLS (GAC isoform) in complex with glutamate and the allosteric inhibitor BPTES (3uo9)[18] shown in light-grey (with the BPTES molecules depicted as stick models in white), the structure of GLS (GAC isoform) in complex with the activator inorganic phosphate (3ss4)[19] shown in grey (with phosphate molecules in green), and the structure of GLS (KGA isoform) in complex with the inhibitor 6-diazo-5-oxo-L-norleucine (DON, not shown) (4o7d)[20] depicted in black. In the figure, the gating loops of the four subunits are found to the left and right of the BPTES molecules. Engagement of gating loop residue Phe322 at the dimer-dimer interface upon tetramerization is thought to stabilize the open conformation of this loop providing substrate access to the active site, which may represent the mechanism by which enzyme activation is linked to oligomerization[19]. The gating loop is also involved in binding bis-2-(5-phenylacetimido-1,2,4,thiadiazol-2-yl) ethyl sulfide (BPTES), which selectively and allosterically inhibits GLS by stabilizing an inactive enzyme conformation[18]. Thus, the identified Pro313Leu mutation affects a GLS region that is sensitive to structural changes and crucial for enzyme function. (**B**) Close-up stereo view of the P313L mutation site of subunit A in the full-length GLS (GAC) structure (3unw). The subunits shown in cartoon representation are colored as in (**A**). Pro313 is shown with thick sticks and carbon atoms in yellow and salmon for subunit A and C, respectively. The leucine side introduced by the mutation is depicted as sticks in cyan. For subunit A, the side chain conformer causing the least steric clashes has been modelled, and for subunit C the most preferred conformer, which also causes the most steric clashes. In the crystal structure of the full-length GLS-glutamate complex, the gating loops were disordered. The conformations of this loop as observed in the GLS-BPTES (light-grey) and GLS-phosphate (grey) complexes are indicated. BPTES (white carbon atoms), glutamate (magenta carbon atoms) and phosphate molecules are shown as sticks. All depicted residues and ligands are labeled, with an apostrophe indicating residue origin in a subunit other than A.

Hydrogen bonds are depicted as black dashed lines. Pro313 is well conserved between glutaminases from diverse species. Its backbone conformational flexibility is restricted due to its cyclic side chain involving the peptidic amide group, which potentially is of functional or structural importance. Mutation of Pro313 to leucine increases backbone conformational flexibility and causes steric clashes with surrounding residues for all frequently observed rotamers of its side chain. Accommodation of the leucine side chain is thus likely to require repositioning of the gating loop-preceding sequence. In absence of a crystal structure for a catalytically active GLS conformation it is however difficult to predict what impact this may have on the flexibility of the gating loop and its functionally crucial capability to adopt specific conformations. Amongst the residues within a 5 Å radius to Leu313 likely to be affected by local conformational changes are His330 and Asn331, as well as Ser469, which closely precede and follow, respectively, active site residues Asn335 and Tyr466 (all shown as sticks with carbon atoms colored as the corresponding subunit). Direct influence of the mutation on active site architecture can therefore not be excluded. Furthermore, sites 313 of the two subunits forming a dimer are in immediate vicinity to each other, and Pro313Leu exchange in both subunits prone to occur in case of mutation homozygosity is therefore expected to have cumulative detrimental effects, including compromised GLS capability to properly dimerize and consequently also to form tetramers. As the latter is essential for enzyme function, this should render corresponding GLS variants inactive.

Figure S6

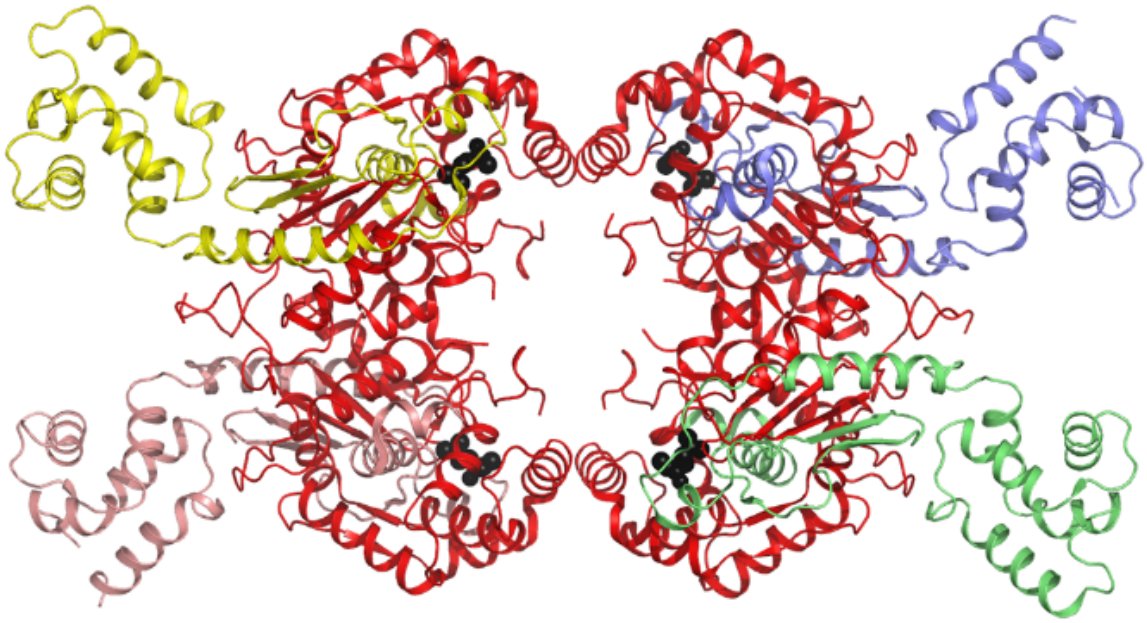

Figure S6: Effect of the c.923dupA (p.Tyr308*) mutation on the expressed GLS protein. The crystals structure of the GLS homotetramer (GAC isoform) in complex with glutamate (PDB accession code 3unw[18]) is shown, with the parts of the distinct subunits that should still be present in the expressed protein colored yellow, salmon, blue and green, respectively. Sequences absent due to the deletion mutation are colored red. The latter include large portions of the α/β domain (residues 224-275, 421-526) and most residues of the extended α-helical domain (276-420). Thus, neither the active site, whose location is indicated by the space fill model of the reaction product glutamate in black, nor any of the subunit interface are (fully) formed, rendering the truncated protein completely inactive.

Figure S7

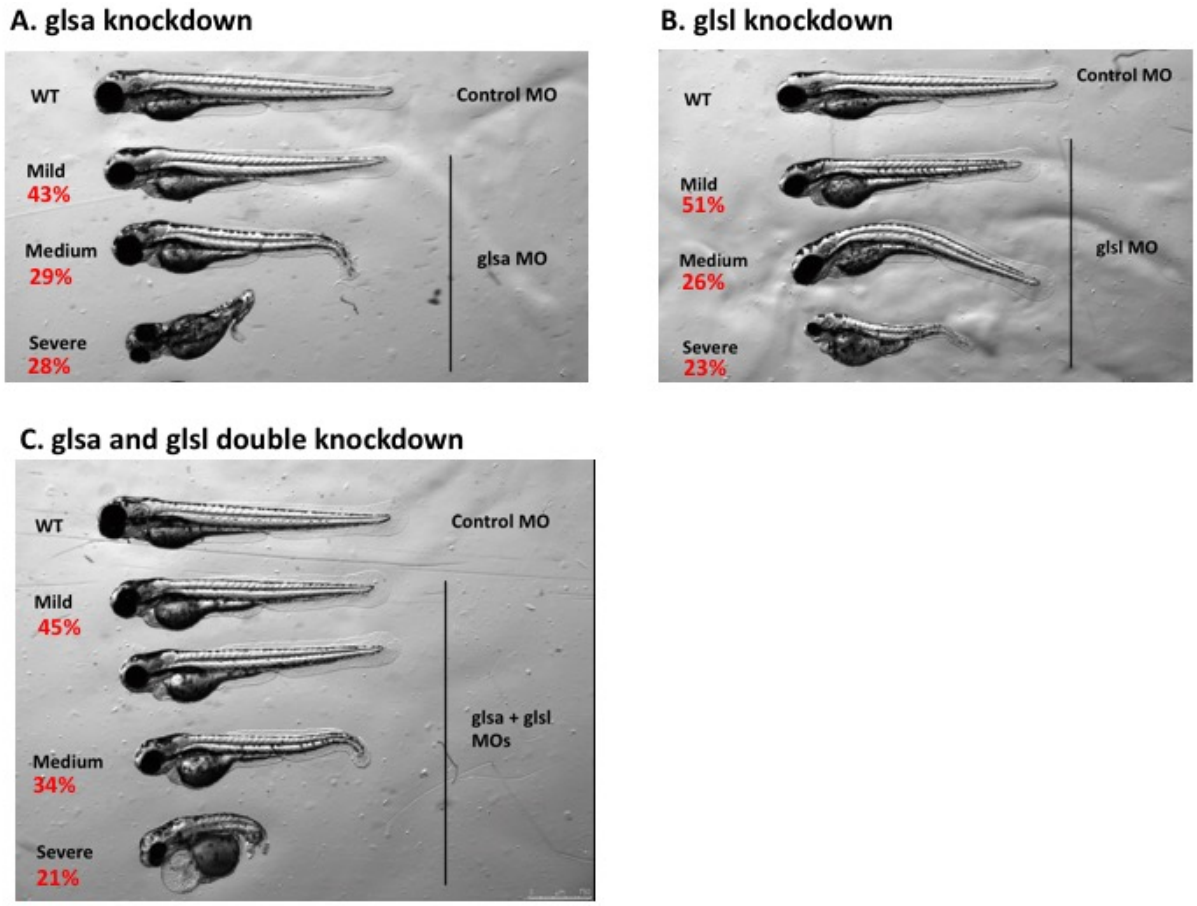

Fig. S7. Glutaminase deficient zebrafish. Zebrafish *glsa* and *glsl* gene knockdown by ATG translation blocking morpholinos. *glsa-atg* and/or *glsl-atg* MOs as well as a standard control MO were microinjected into newly fertilized WT zebrafish embryos and the phenotype were scored at 3 day post fertilization (dpf). Three categories of phenotypes are recorded: mild, medium and severe, showing various degrees of smaller body size, curved body and cardiac edema. The percentage of Zebrafish with a particular phenotype is indicated in red.

Figure S8

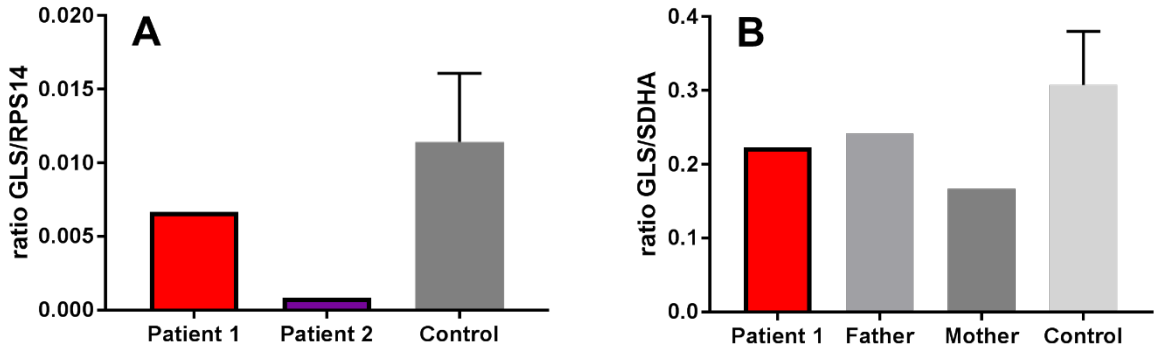

Fig. S8 Quantitative RT-PCR analysis in families 1 and 2. (**A**) Quantitative RT-PCR analysis of *GLS* cDNA levels prepared from skin fibroblasts of patients 1 and 2 reveals ~50% lower expression in patient 1 and ~90% lower expression in patient 2 compared to average levels in skin fibroblasts from healthy controls. (**B**) Quantitative PCR analysis of *GLS* cDNA levels prepared from blood samples of family 1 reveals somewhat lower expression in the samples of the patient and mother.

Figure S9

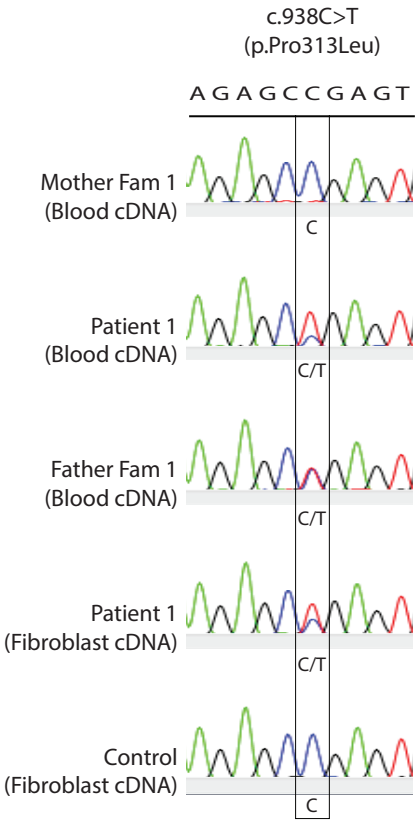

Fig. S9. Allelic expression imbalance in family 1. *GLS* cDNA sequencing of members of Family 1 showing a relative deficit of the mRNA from the expanded allele relative to *GLS*-c.938T allele in the patient.

Figure S10

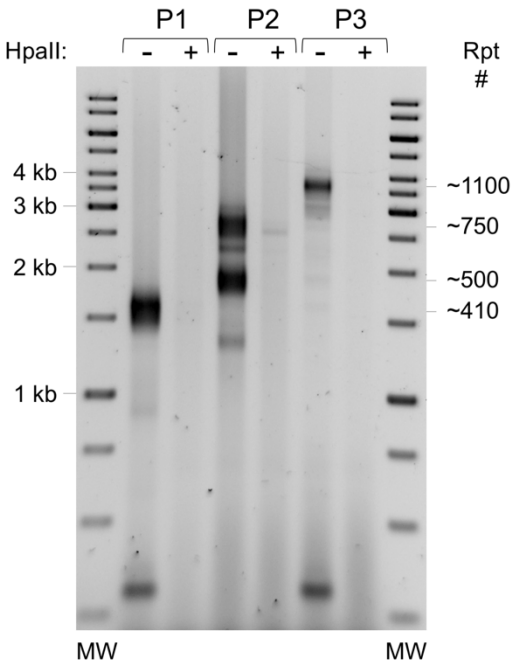

Fig. S10. Repeat PCR of the GCA expansion in fibroblasts for all 3 patients with and without predigestion of the PCR template with HpaII. The DNA from all patients was isolated from the fibroblasts that were used for the ChIP-qPCR analysis shown in Fig. 3. The ~300 bp PCR product seen in the P1 and P3 samples without HpaII predigestion represents the non-expanded allele in the patients heterozygous for the expansion.

Figure S11

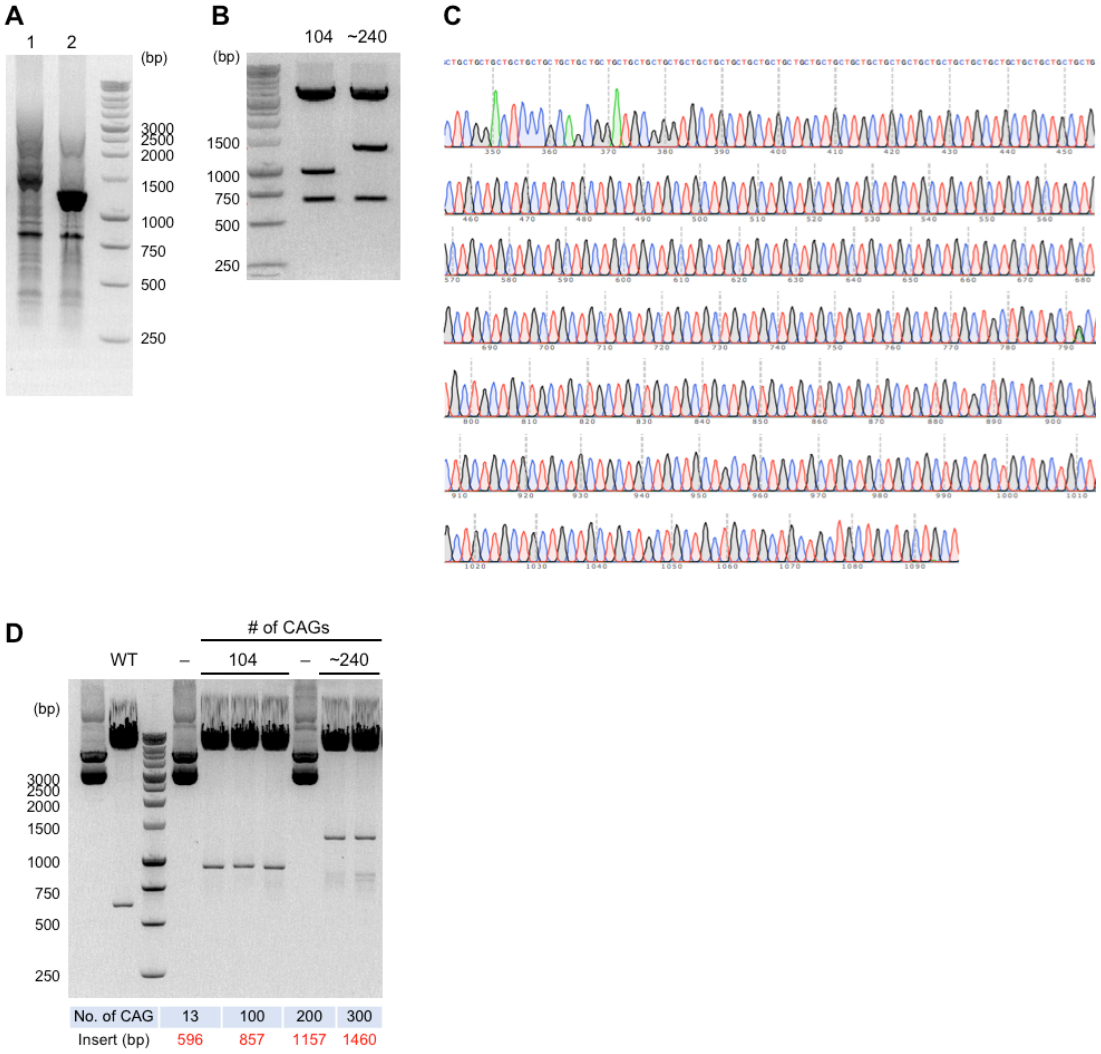

Fig. S11. Cloning of *GLS* promoters with expanded GCA repeats. (**A**) Amplicons carrying an expanded [GCA] in *GLS* promoter amplified from patients' fibroblasts - lane 1: *exp/exp*, lane 2: WT/*exp*. Excluding the GCA tract, amplicons are expected to be 1186 bp. Patient 1 (*exp/exp*) is predicted to have >100-300 GCA, and patient 2 is heterozygous with an expanded copy with ~270 [GCA]; (**B**) EcoRI digestion of pCR-blunt clones with expanded [GCA] in *GLS* promoter shows variable GCA length (104, and ~240); (**C**) Sanger sequencing trace of pCR-blunt clone shows good quality peaks for >201 GCA repeats for a pCR-blunt clone with ~240 repeats (sequence shown is the reverse strand); (**D**) EcoRI and BglII digestion of empty vector pGL3 (-), WT (H9), 104 (<120), and >201 (~240) [GCA] with expected insert size highlighted in red.

Figure S12

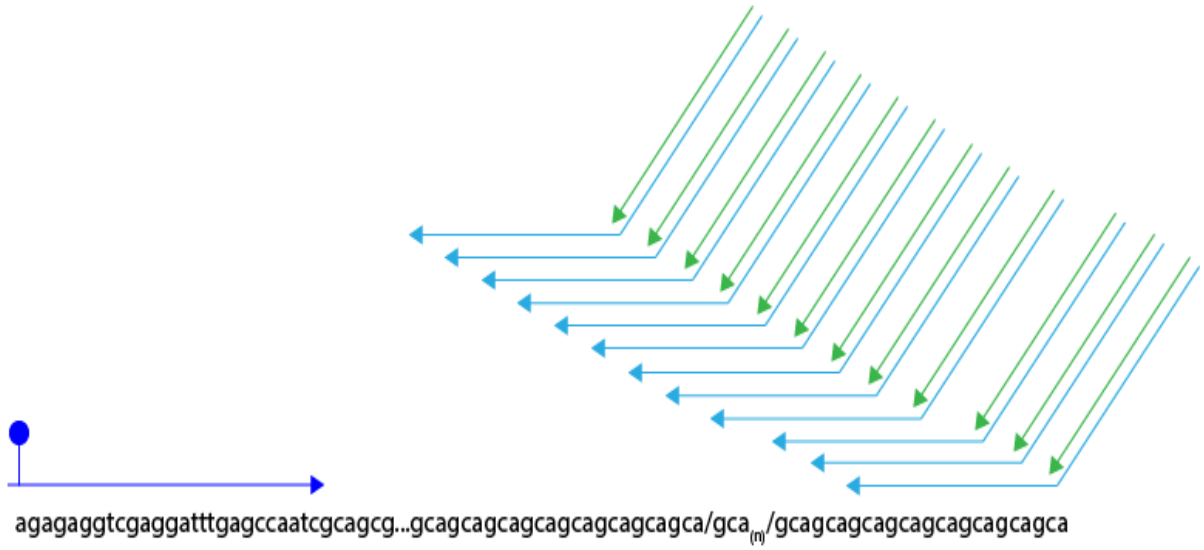

Fig. S12. Primers used for the triplet repeat-primed PCR. A PCR was performed with the three primers indicated above, with the dark blue arrow representing the locus-specific fluorescently labelled forward primer, the light blue arrow representing the repeat annealing reverse primer and the green arrow representing the tail reverse primer.

SUPPLEMENTARY TABLES

Table S1 Overview of the variants identified in the trio-WES data for family 1 and duo-WES data in family 2, and the corresponding gene lists for each category. Details about categories in Supplemental Methods.

Family	*De novo*	Recessive (homozygous)	Recessive (compound het)	General Damaging	General Damaging Intellectual Disability / Encephalopathy
Family 1	4	0	6	345	24
Family 2	0	0	0	320	29

Family 1 (exome)						
De novo	Recessive (compound het)	General Damaging				General Damaging Intellectual Disability / Encephalopathy
PUSL1	GPA33	ABCC11	FMNL1	PCDHAC1	TEX10	ANK3
SKA3	GPA33	ABI3BP	FMO6P	PCNXL4	TGFBI	ATR
SYT16	ABCC11	AC005076.5	FOXQ1	PDE12	TGFBR3	CDH23
IGLV5-45	ABCC11	AC005609.1	FZD6	PDE5A	THBS3	CEP152
	ABCC11	AC006486.1	GABRE	PDK2	TIMM44	DYM
	ABCC11	AC104809.3	**GLS**	PDZD2	TJP3	EP300
	ICAM1	ACADL	GMIP	PDZD4	TM9SF4	EVX1
	ICAM1	ADM2	GPA33	PGM3	TMEM131	**GLS**
	TTN	AGPAT4	GPBP1	PHF23	TNFAIP2	GTF2I
	TTN	AL078585.1	GPR107	PHKA1	TNFAIP8L2	IFT172
	TTN	ALDH9A1	GPR137B	PIAS2	TNK2	LAMP2
	TTN	ALG14	GPRC5C	PIEZO1	TRAK2	MCCC2
	ANK2	ALS2CL	GPS1	PIK3R5	TRDN	MYO7A
	ANK2	AMN	GTF2I	PITPNM3	TRMT44	PAX8
	LRPAP1	AMN1	HARS	PKD1L1	TSHZ1	PGM3
	LRPAP1	ANK2	HARS2	PLCG2	TTC40	PLEKHG2
		ANK3	HBS1L	PLEKHG1	TTN	RREB1
		ANKRD1	HMGB4	PLEKHG2	TTYH3	SCO1
		ANKUB1	ICAM1	PLEKHG3	TULP4	SLX4
		ANPEP	IFT172	PLEKHG4B	UBR2	SMARCA2
		ANXA6	IGFN1	PLXNB2	UNC5B	SPG11
		ARHGAP11A	IGLV5-45	POLD4	UNC5CL	TBC1D24

Family 1 (exome)						
De novo	Recessive (compound het)	General Damaging				General Damaging Intellectual Disability / Encephalopathy
		ARL6IP5	IGSF1	POLL	URB1	USH2A
		ATP6V1F	IKZF2	POU6F2	USH2A	VPS11
		ATP9B	IL22RA1	PRKG2	USP50	
		ATR	ITFG3	PTK7	VPS11	
		ATXN1	ITPKB	PTPN14	VPS45	
		ATXN7	JMJD7	PUS7L	VWA3A	
		B4GALNT2	JMJD8	PUSL1	WDR17	
		BRAT1	KANK2	PWP1	WDR19	
		BRF2	KANK4	PYGB	WFDC1	
		C10orf90	KCNA10	PYGM	WRNIP1	
		C12orf57	KCNC1	RAB11FIP1	XAB2	
		C17orf78	KCNH2	RASL12	XIRP2	
		C2orf16	KIAA0319L	RBM45	XPO5	
		C3orf49	KIAA1033	RCN3	YBX3	
		C4orf6	KIAA1429	RFX1	ZBTB7C	
		CABYR	KIAA1468	RHBDD1	ZFAT	
		CACNG1	KLB	RHBDL1	ZFYVE28	
		CAPN1	KLC4	RIN2	ZNF142	
		CAPN15	KLHL36	RIPK3	ZNF205	
		CAPNS2	LAMP2	RLF	ZNF460	
		CBFB	LARP4B	RNF128	ZNF462	
		CBX8	LPAR4	RPL29	ZNF627	
		CCDC183	LRIG1	RPTOR	ZNF766	
		CCDC86	LRPAP1	RREB1		
		CCNO	LRRC16A	RTL1		
		CCR6	LRRC31	SAMD12		
		CCSER1	LRRC41	SCLT1		
		CCT5	LRRN4	SCO1		
		CDC42BPG	MAR1	SEC31B		
		CDC42EP4	MATN2	SERINC2		
		CDH13	MC1R	SETD1B		
		CDH23	MCCC2	SGIP1		
		CEACAM5	MDN1	SH2D4B		
		CEP152	MED11	SH3D21		
		CHCHD5	MGA	SHANK2		
		CLEC12A	MIPEP	SIRPG		
		CLEC4A	MLK4	SKA3		
		CLGN	MMP9	SLAIN1		
		CNKSR1	MOK	SLC18A1		
		CNOT10	MRPL44	SLC18B1		
		CNTN2	MSLN	SLC22A25		
		COL7A1	MTHFD1	SLC30A6		
		CPNE7	MUC16	SLC35A1		
		CPPED1	MUC17	SLC38A6		
		CSF2RB	MUC3A	SLC45A4		
		CSNK1G1	MUC5B	SLC4A1		
		CTB-133G6.1	MUTYH	SLC4A3		

Family 1 (exome)						
De novo	Recessive (compound het)	General Damaging				General Damaging Intellectual Disability / Encephalopath
		CTD-2207O23.3	MYH9	SLC9A5		
		CYP26C1	MYO15A	SLCO5A1		
		DAZL	MYO19	SLX4		
		DIDO1	MYO7A	SMARCA2		
		DNAH10	MYO7B	SMG5		
		DNAJB8	NADK	SOGA2		
		DNASE1	NBR1	SPEN		
		DTNB	NCK1	SPG11		
		DYM	NEBL	SPG20		
		DZIP3	NET1	SPNS1		
		EBLN2	NID1	SPRY4		
		EDA	NISCH	SPTBN1		
		EDA2R	NKPD1	SPTLC3		
		EDAR	NLRP4	ST8SIA1		
		EFCAB4A	NMNAT1	STEAP4		
		ENTPD4	NMRK1	STK32A		
		EP300	NOL11	SUCO		
		EPPK1	NOS1	SUPT16H		
		ERBB2	NOX4	SYT15		
		ERMP1	NRP1	SYT16		
		ESPNL	NUBP2	SYT6		
		ETS1	OR10S1	SYTL2		
		EVX1	OR11H4	SYTL4		
		FAM110B	OR2L2	TACC2		
		FAM149A	PABPC1	TAF4B		
		FAM178A	PABPC4L	TAP1		
		FAM20C	PAN2	TBC1D24		
		FAN1	PAX8	TCF3		
		FASN	PCDHA1	TDRD15		
		FAT2	PCDHA12	TDRD9		
		FLG	PCDHA13	TEKT5		

Family 2 (exome)				
General Damaging				General Damaging Intellectual Disability / Encephalopathy
ABCA13	FAM133A	PDYN	USH2A	ABCD4
ABCA3	FAM170A	PELI3	VWA3A	BACE2
ABCA9	FAM193B	PELP1	VWA5B1	CACNA1D
ABCD4	FAM81B	PITRM1	WDR52	CC2D2A
ABL1	FAP	PIWIL3	WDR87	CHD7
ABLIM1	FBN3	PKD1L1	WRN	CNTNAP2
AC005008.2	FBXL15	PLCG2	WWOX	COG5
AC079354.1	FBXO33	PLEC	ZAP70	COG8
ACACB	FBXO38	POLB	ZDHHC13	CYP24A1
ACAD8	FIGN	POLR3A	ZDHHC2	DVL1
ACAN	FLT3	POMGNT1	ZNF18	ELN
ADAMTS6	FMN2	PPEF1	ZNF451	FMN2
ADAMTSL1	FMNL2	PPRC1	ZNF462	FOS
AFF1	FOCAD	PRMT7	ZNF510	GNPAT
AGRN	FOS	PTGER4	ZNF608	GRIFIN
AGTPBP1	GAK	PTPN14	ZNF623	IFIH1
AKAP9	GALR2	PTPN7	ZNF687	KCNQ2
AL354898.1	GAPDHS	QPCT	ZNF831	MCCC2
ANKRD6	GBP7	RAB21	ZSCAN29	NDUFS2
ANLN	GDAP1L1	RAD54L2	ZSCAN5C	POMGNT1
ANO5	GLB1L3	RASAL2		PRMT7
AP000889.3	GNAT1	RBP3		RBP3
AQP5	GNPAT	RC3H1		SMARCA4
ARAP3	GPBAR1	RCN3		TG
ARHGAP19-SLIT1	GPR123	RHBDL1		TOPORS
ASPH	GPR152	RHOT2		TYMP
ATAD2B	GRIFIN	RIF1		UMODL1
ATG4B	GRK7	RIN3		USH2A
ATXN7L1	GULP1	RNF103		WWOX
AZI1	Gene	RNF135		
BACE2	HIF1A	RNF25		
BCAM	HIP1R	ROBO1		
BOK	HMCN1	RP11-113D6.6		
C12orf55	IFIH1	RP11-156E8.1		

Family 2 (exome)				
General Damaging				General Damaging Intellectual Disability / Encephalopathy
C16orf71	IGFLR1	RP11-355N15.1		
C17orf64	INSC	RP11-598P20.5		
C1QA	IRF2	RPA3-AS1		
C1QC	KCNN3	RPS10		
C1orf189	KCNQ2	SAMD15		
C6orf123	KCTD19	SAMD9		
CABIN1	KHDRBS2	SBF1		
CABP4	KIAA0947	SCIN		
CACNA1D	KIAA2026	SDCBP2		
CAND2	KRT34	SEC31B		
CAPN8	LAMA1	SEMA4B		
CATSPERB	LAMA5	SENP6		
CC2D2A	LAPTM4B	SH3GL2		
CCDC33	LCTL	SH3KBP1		
CD248	LIMCH1	SH3PXD2B		
CD300LB	LIMD1	SHANK2		
CD3EAP	LIN9	SHPRH		
CD8A	LMF2	SLAIN1		
CDT1	LOXHD1	SLC10A2		
CHD7	LOXL2	SLC17A6		
CLASP2	LPPR4	SLC23A3		
CLCN7	LRRC27	SLC26A10		
CLEC1B	LRRC39	SLC35A1		
CNTNAP2	MACF1	SLC4A9		
COG5	MAN2A1	SLCO5A1		
COG8	MCCC2	SLTM		
COL16A1	MGAT3	SMAD7		
COL17A1	MKRN1	SMARCA4		
COL20A1	MTMR12	SMG7		
COL6A2	MUC16	SNIP1		
COL6A5	MUC19	SORL1		
COQ5	MUC2	SPAG17		
CORO2A	MUC5B	SPATA20		
CORO7-PAM16	MUSK	SPDL1		
CPED1	MYBPC2	SRSF2		

Family 2 (exome)				
General Damaging				General Damaging Intellectual Disability / Encephalopathy
CR2	MYH2	SSRP1		
CRTAP	MYL6	ST5		
CSGALNACT1	MYO10	ST8SIA3		
CTBP2	NAE1	STAT6		
CTC-432M15.3	NAV1	STK33		
CTSC	NAV2	SULF1		
CYP24A1	NCF4	SUSD4		
CYP4F11	NDOR1	SYT16		
CYP4F2	NDUFS2	T		
DAB2IP	NEDD4	TAX1BP1		
DENND5A	NEO1	TG		
DHX38	NFE2L3	TM4SF5		
DHX57	NOTCH2	TMEM102		
DLGAP5	NPHS1	TMEM82		
DNAJC10	NPY4R	TMEM8A		
DOCK11	NTN5	TMPRSS3		
DR1	NUDT5	TNFRSF10C		
DVL1	NXN	TNFRSF1A		
DYSF	NXPE4	TOPORS		
EDEM2	OC90	TOX3		
EEF1E1	ODF2	TRPM7		
EEPD1	OGFOD2	TRPV3		
ELN	OR10A4	TSPAN32		
EML4	OR2T8	TSR3		
ENPP3	OR51S1	TTC24		
EP400	OR5K1	TTLL10		
EPHX2	OSBP2	TTN		
EPPK1	OTOG	TYMP		
ESCO1	PABPC1	UBE2L6		
ETV4	PACRG	UMODL1		
EXOSC5	PCM1	UNC5CL		

Table S2 Information on the bioinformatics pipeline, including versions, acquisition dates, PubMed identifiers (PMIDs), and arXiv citations where applicable.

Tool	Version || {Date Acquired}	PMID (or bioRxiv/arXiv)	Source
BWA mem	0.7.12	arXiv:1303.3997v1	https://sourceforge.net/projects/bio-bwa/files/
Samtools	1.2	19505943	http://www.htslib.org/download/
Picard	1.139	.	https://broadinstitute.github.io/picard/
GATK	3.4-46	27250555	https://software.broadinstitute.org/gatk/download/
VT	0.5772	.	https://genome.sph.umich.edu/wiki/Vt
SNPEff	4.11 (gene: GRCh37.75)	22728672	http://snpeff.sourceforge.net/download.html
BCFTools	1.8	.	http://www.htslib.org/download/
bgzip	0.2.5	.	http://www.htslib.org/download/
tabix	0.2.5	.	http://www.htslib.org/download/
VCFAnno	0.2.8	27250555	https://github.com/brentp/vcfanno
VCF2DB	{January 22nd 2018}	.	https://github.com/quinlan-lab/vcf2db
GEMINI	0.19.1	23874191	https://gemini.readthedocs.io/en/latest/
Database	**Version || {Date Acquired}**	**PMID (or bioRxiv/arXiv)**	**Source**
GRCh37	{August 2017}	.	http://www.bcgsc.ca/downloads/genomes/9606/hg19/1000genomes/bwa_ind/genome/
CADD	1.3	24487276	https://cadd.gs.washington.edu
PolyPhen2	2.2.2	23315928	ftp://genetics.bwh.harvard.edu/pph2/whess/polyphen-2.2.2-whess-2011_12.tab.tar.bz2
PrimateAI	0.2	30038395	https://github.com/Illumina/PrimateAI
gnomAD (exome)	2.0.2	27535533	http://gnomad.broadinstitute.org/downloads
gnomAD (genome)	2.0.2	27535533	http://gnomad.broadinstitute.org/downloads
GEMINI Databases*	0.19.1	23874191	https://gemini.readthedocs.io/en/latest/
RLCRs	{April 2018}	29304372	http://tcag.ca/documents/projects/RLCRs_no_Repeat_Masker.zip
ConfidentRegions	{April 2018}	27903644	ftp://ussd-ftp.illumina.com/2017-1.0/hg19/small_variants/
OMIM	{August 2018}	.	http://data.omim.org/downloads/
RefSeq Gene Summary	{October 2016}	.	ftp://ftp.ncbi.nlm.nih.gov/refseq/H_sapiens/RefSeqGene/
RVIS	Release 3, March 2016	23990802	http://genic-intolerance.org
pLI	Release 0.3.1, March 2016	27535533	https://decipher.sanger.ac.uk/info/loss-intolerance
HPO	{May 2018}	27899602	https://hpo.jax.org/app/download/annotation
MeSHOPs	Release 1, July 2010	23017167	http://www.meshops.oicr.on.ca/meshop/index.html
GeneHancer	{October 2017}	28605766	https://genecards.weizmann.ac.il/geneloc/index.shtml
TrAP	{January 2018}	28794409	http://trap-score.org
*GEMINI Databases includes all databases installed with the gemini package			

Table S3 – GLS Primer Sequences	
GLS-L	[SacI/KpnI]ATGTAGCCTCAGGGAATAACTCAG
GLS-R	[XhoI]GCCGCCGGGTCCGTCAGCGCCC
T7F	TAATACGACTCACTATAGGGGGG
M13R	CAGGAAACAGCTATGAC
F1ori	TATTACGCCAGCCCAAGCTA
LucR	CATCGACTGAAATCCCTGGT
GLS-L2	GCAGCACCCGCATCCGCTGCGGGA

Table S3. Sequences of primers used for cloning and verification of *GLS* promoter luciferase reporter vector (all primers are listed (5′→3′)

SUPPLEMENTARY REFERENCES

1. Tarailo-Graovac M, Shyr C, Ross CJ, et al. Exome Sequencing and the Management of Neurometabolic Disorders. N Engl J Med 2016;374:2246-55.

2. Beaulieu CL, Majewski J, Schwartzentruber J, et al. FORGE Canada Consortium: outcomes of a 2-year national rare-disease gene-discovery project. Am J Hum Genet 2014;94:809-17.

3. Dolzhenko E, van Vugt J, Shaw RJ, et al. Detection of long repeat expansions from PCR-free whole-genome sequence data. Genome Res 2017;27:1895-903.

4. RK CY, Merico D, Bookman M, et al. Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci 2017;20:602-11.

5. Project MinE: study design and pilot analyses of a large-scale whole-genome sequencing study in amyotrophic lateral sclerosis. Eur J Hum Genet 2018;26:1537-46.

6. Mayer LB, Krawczyk M, Grunhage F, Lammert F, Stokes CS. A genetic variant in the promoter of phosphate-activated glutaminase is associated with hepatic encephalopathy. J Intern Med 2015;278:313-22.

7. Covarrubias-Pazaran G, Diaz-Garcia L, Schlautman B, Salazar W, Zalapa J. Fragman: an R package for fragment analysis. BMC Genet 2016;17:62.

8. Rajan-Babu IS, Law HY, Yoon CS, Lee CG, Chong SS. Simplified strategy for rapid first-line screening of fragile X syndrome: closed-tube triplet-primed PCR and amplicon melt peak analysis. Expert Rev Mol Med 2015;17:e7.

9. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 2013;8:2281-308.

10. Andersen CL, Jensen JL, Orntoft TF. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited

for normalization, applied to bladder and colon cancer data sets. Cancer Res 2004;64:5245-50.

11. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66:486-501.

12. Schrodinger LLC. The PyMOL Molecular Graphics System, Version 1.3r1. 2010.

13. Avdesh A, Chen M, Martin-Iverson MT, et al. Regular care and maintenance of a zebrafish (Danio rerio) laboratory: an introduction. J Vis Exp 2012:e4196.

14. Hayward BE, Zhou Y, Kumari D, Usdin K. A Set of Assays for the Comprehensive Analysis of FMR1 Alleles in the Fragile X-Related Disorders. J Mol Diagn 2016;18:762-74.

15. Kumari D, Usdin K. Polycomb group complexes are recruited to reactivated FMR1 alleles in Fragile X syndrome in response to FMR1 transcription. Hum Mol Genet 2014;23:6575-83.

16. Gurnot C, Martin-Subero I, Mah SM, et al. Prenatal antidepressant exposure associated with CYP2E1 DNA methylation change in neonates. Epigenetics 2015;10:361-72.

17. Furukawa F, Tseng YC, Liu ST, et al. Induction of Phosphoenolpyruvate Carboxykinase (PEPCK) during Acute Acidosis and Its Role in Acid Secretion by V-ATPase-Expressing Ionocytes. Int J Biol Sci 2015;11:712-25.

18. DeLaBarre B, Gross S, Fang C, et al. Full-length human glutaminase in complex with an allosteric inhibitor. Biochemistry 2011;50:10764-70.

19. Cassago A, Ferreira AP, Ferreira IM, et al. Mitochondrial localization and structure-based phosphate activation mechanism of Glutaminase C with implications for cancer metabolism. Proc Natl Acad Sci U S A 2012;109:1092-7.

20. Thangavelu K, Pan CQ, Karlberg T, et al. Structural basis for the allosteric inhibitory mechanism of human kidney-type glutaminase (KGA) and its regulation by Raf-Mek-Erk signaling in cancer cell metabolism. Proc Natl Acad Sci U S A 2012;109:7705-10.

21. Thangavelu K, Chong QY, Low BC, Sivaraman J. Structural basis for the active site inhibition mechanism of human kidney-type glutaminase (KGA). Sci Rep 2014;4:3827.

22. McDermott LA, Iyer P, Vernetti L, et al. Design and evaluation of novel glutaminase inhibitors. Bioorg Med Chem 2016;24:1819-39.

23. Ramachandran S, Pan CQ, Zimmermann SC, et al. Structural basis for exploring the allosteric inhibition of human kidney type glutaminase. Oncotarget 2016;7:57943-54.

24. Huang Q, Stalnecker C, Zhang C, et al. Characterization of the interactions of potent allosteric inhibitors with glutaminase C, a key enzyme in cancer cell glutamine metabolism. J Biol Chem 2018;293:3535-45.

ACKNOWLEDGMENTS

This work was supported by funding from: the B.C. Children's Hospital Foundation (1st Collaborative Area of Innovation), Alberta Children's Hospital Research Institute Foundation (MTG), the Canadian Institutes of Health Research (#301221 grant; CvK), Rare Diseases Foundation (MTG, CvK), Canadian Rare Disease Models & Mechanisms Network (BD, XYW, CvK), Brain Canada Foundation/Health Canada Platform Support Grant (#PSG14-3505, XYW), Natural Sciences and Engineering Research Council of Canada (#RGPIN 05389-14, XYW), National Ataxia Foundation, National Institutes of Health (#DK057808, KU) and Stichting Metakids (CvK). Informatics infrastructure was supported by Genome BC and Genome Canada (ABC4DE Project). CvK is recipient of the Michael Smith Foundation for Health Research Scholar Award and a Stichting Metakids salary stipend. BID received stipends from CIHR-DSECT, CIHR and the Michael Smith Foundation for Health Research during the period of this study.

‘PROJECT MINE ALS SEQUENCING CONSORTIUM’ MEMBERS

Surname	First name	Afilliation
Al-Chalabi	Ammar	Maurice Wohl Clinical Neuroscience Institute, King's College London, Department of Basic and Clinical Neuroscience, London, UK.
Basak	A. Nazli	Neurodegeneration Research Laboratory, Bogazici University, Istanbul, Turkey.
Glass	Jonathan D.	Department Neurology, Emory University School of Medicine, Atlanta, GA, USA.
Hardiman	Orla	Academic Unit of Neurology, Trinity College Dublin, Trinity Biomedical Sciences Institute, Dublin, Republic of Ireland.
Landers	John E.	Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA.
Mora	Jesus S.	Hospital San Rafael, Madrid, Spain
Morrison	Karen E.	Faculty of Medicine, University of Southampton, Southampton, UK.
Panades	Monica P.	Neurology Department, Hospital Universitari deBellvitge, Barcelona, Spain
Robberecht	Wim	KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), B-3000 Leuven, Belgium
Shaw	Pamela J.	Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK.
Shaw	Christopher E.	Maurice Wohl Clinical Neuroscience Institute, King's College London, Department of Basic and Clinical Neuroscience, London, UK.
Van Damme	Philip	KU Leuven - University of Leuven, Department of Neurosciences, Experimental Neurology and Leuven Research Institute for Neuroscience and Disease (LIND), B-3000 Leuven, Belgium
Veldink	Jan	Department of Neurology, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, The Netherlands