




## RESEARCH ARTICLE

# Quantitative determination of *SLC2A1* variant functional effects in GLUT1 deficiency syndrome

Naeimeh Tayebi<sup>1</sup> , Brian Leon-Ricardo<sup>1</sup>, Kevin McCall<sup>1</sup>, Elvira Mehinovic<sup>2</sup>, Kristin Engelstad<sup>3</sup>, Vincent Huynh<sup>3</sup>, Tychele N. Turner<sup>2</sup>, Judy Weisenberg<sup>1</sup>, Liu L. Thio<sup>1</sup>, Paul Hruz<sup>4</sup>, Robin S. B. Williams<sup>5</sup> , Darryl C. De Vivo<sup>3</sup>, Vincent Petit<sup>6</sup>, Gabe Haller<sup>1,2,7</sup> & Christina A. Gurnett<sup>1</sup> 

<sup>1</sup>Department of Neurology, Washington University in St Louis, St Louis, Missouri, USA

<sup>2</sup>Department of Genetics, Washington University in St Louis, St Louis, Missouri, USA

<sup>3</sup>Department of Neurology, Columbia University Irving Medical Center, New York, New York, USA

<sup>4</sup>Department of Pediatrics, Washington University in St Louis, St Louis, Missouri, USA

<sup>5</sup>Centre for Biomedical Sciences, Department of Biological Sciences, Royal Holloway University of London, Egham, UK

<sup>6</sup>Metafora Biosystems, Paris, France

<sup>7</sup>Department of Neurological Surgery, Washington University in St Louis, St Louis, Missouri, USA

## Correspondence

Christina A. Gurnett, Department of Neurology, Washington University in St Louis, St Louis, MO 63110, USA. Tel: 314-454-4225; E-mail: [gurnettc@wustl.edu](mailto:gurnettc@wustl.edu)

Received: 30 November 2022; Revised: 8 March 2023; Accepted: 12 March 2023

*Annals of Clinical and Translational Neurology* 2023; 10(5): 787–801

doi: 10.1002/acn3.51767

## Abstract

**Objective:** The goal of this study is to demonstrate the utility of a growth assay to quantify the functional impact of single nucleotide variants (SNVs) in *SLC2A1*, the gene responsible for Glut1DS. **Methods:** The functional impact of 40 SNVs in *SLC2A1* was quantitatively determined in HAP1 cells in which *SLC2A1* is required for growth. Donor libraries were introduced into the endogenous *SLC2A1* gene in HAP1-Lig4KO cells using CRISPR/Cas9. Cell populations were harvested and sequenced to quantify the effect of variants on growth and generate a functional score. Quantitative functional scores were compared to 3-OMG uptake, *SLC2A1* cell surface expression, CADD score, and clinical data, including CSF/blood glucose ratio. **Results:** Nonsense variants ( $N = 3$ ) were reduced in cell culture over time resulting in negative scores (mean score:  $-1.15 \pm 0.17$ ), whereas synonymous variants ( $N = 10$ ) were not depleted (mean score:  $0.25 \pm 0.12$ ) ( $P < 2e-16$ ). Missense variants ( $N = 27$ ) yielded a range of functional scores including slightly negative scores, supporting a partial function and intermediate phenotype. Several variants with normal results on either cell surface expression (p.N34S and p.W65R) or 3-OMG uptake (p.W65R) had negative functional scores. There is a moderate but significant correlation between our functional scores and CADD scores. **Interpretation:** Cell growth is useful to quantitatively determine the functional effects of *SLC2A1* variants. Nonsense variants were reliably distinguished from benign variants in this *in vitro* functional assay. For facilitating early diagnosis and therapeutic intervention, future work is needed to determine the functional effect of every possible variant in *SLC2A1*.

## Introduction

The genetic cause of many disorders is known but determining which variants are pathogenic is challenging. The accumulation of variants of uncertain significance (VUS) and the failure to accurately classify genetic variants presents a growing problem, particularly as sequencing before symptom onset becomes commonplace and for disorders in which treatments are available.

Computational tools for predicting variant pathogenicity are attractive because they scale to millions of variants,<sup>1</sup> but conflicting predictions and poor sensitivity and specificity reduces the impact of these tools.<sup>2–4</sup> This limitation is formalized in the guidelines for variant interpretation, which state that computational predictions should never be used as the only evidence of pathogenicity.<sup>5,6</sup> The American College of Medical Genetics, the Association for Molecular Pathology, and the NIH-supported Clinical

Genome Resource (ClinGen) all emphasize the importance of functional data,<sup>5</sup> yet variant-specific functional data remains sparse relative to the number of unique variants observed in the population. This limitation hinders efforts to use patient-specific variants for clinical decision-making or to improve clinical trial outcomes.

The Glucose Transporter Type 1 Deficiency Syndrome (Glut1DS; MIM#606777) was first described by De Vivo et al.,<sup>7</sup> and is caused by mutations in the *SLC2A1* gene, which encodes the main endothelial glucose transporter GLUT1. Diagnosis is essential because the ketogenic diet is effective for treating associated symptoms. GLUT1 is expressed in the blood–brain barrier and is responsible for transporting glucose into the brain<sup>7</sup> and other organs.<sup>8</sup> Glut1DS encompasses a spectrum of neurological disorders, including early onset seizures with acquired microcephaly and cognitive impairment (classic type), paroxysmal choreoathetosis and dyskinesia, atypical childhood absence epilepsy, and alternating hemiplegia, along with many related phenotypes.<sup>7</sup> Although diagnosis has historically relied upon detection of reduced cerebrospinal fluid (CSF) glucose, many patients are now diagnosed by sequencing.<sup>9</sup> However, the symptoms of Glut1DS are often nonspecific and there are now nearly 300 VUS in *SLC2A1* in ClinVar, highlighting the critical need for functional data.<sup>10,11</sup> Functional data may also provide important prognostic information about clinical severity.

Because *SLC2A1*, the gene responsible for Glut1DS, is required for growth of the haploid cell line (HAP1),<sup>12</sup> we hypothesized that introduction of single nucleotide variants into the single copy of the *SLC2A1* gene would allow us to determine which variants are detrimental to *SLC2A1* function. Similar methodologies were used to derive functional scores for the disease genes *BRCA1*,<sup>13</sup> *CARD11*,<sup>14</sup> and *DDX3X*<sup>15</sup> which had immediate impact on patient care.<sup>16,17</sup> Here, we quantified the change in variant abundance over time in cells containing 40 individual *SLC2A1* single nucleotide variants grown in culture and report the resultant quantitative functional scores. Experimentally derived functional scores were then compared to the previously reported erythrocyte 3-O-methyl-D-glucose (3-OMG) uptake, *SLC2A1* cell surface expression, CADD score, and clinical data, including CSF/blood glucose ratio.

## Methods

### Introduction of single nucleotide variants into the endogenous *SLC2A1* gene

All CRISPR gRNAs used in this project were cloned into pX459 vector that expresses the gRNA from a U6 promoter and has a Cas9-2A-puromycin resistance cassette. We selected gRNAs that: (1) have a high predicted on-

target activity and minimal predicted off-target activity in the Benchling website (<https://benchling.com>), (2) induce cleavage within *SLC2A1* coding sequence, and (3) target a genomic site permissive to synonymous substitution within the guanine dinucleotide of the PAM site. The oligonucleotides were obtained from Integrated DNA Technologies (IDT). The cloning was performed by site directed mutagenesis using NEBNext 2× polymerase master mix followed by DpnI to digest the parental plasmid template.<sup>18</sup> After purification of mutagenesis PCR product using AmpureXP (Beckman Coulter, Brea, CA, USA), it was transformed into *E. coli*, which were plated on LB-agar with ampicillin at 100 µg/mL. Colonies were cultured and confirmed for the correct gRNA sequences by Sanger sequencing.

As proof of principle, donor libraries consisting of 22 variants in exon 2 and 18 variants in exon 3 were designed to include pathogenic, benign, and VUS variants. Some variants were identified in patients who had clinical and diagnostic data available (Table 1), but some had limited clinical information or were not previously reported in patients (Table S1). Donor libraries consisted of WT sequence (~200 bp) (hg37) with one target single nucleotide substitution as well as 2–4 synonymous substitutions at the PAM site (labeled as markers) to reduce re-cleavage by the Cas9-gRNA complex and to distinguish introduced variants from sequencing errors. The DNA for each donor target including the markers (gBlock) was ordered from IDT. The homology arms (~650 bp upstream and downstream of the targeted region) were generated from HAP1 genomic DNA by PCR using NEBNext 2× polymerase master mix followed by purification with AmpureXP. Homology arms and gBlock were amplified with the primers to add the overlapping sequence, as described previously.<sup>19</sup> The PCR products were purified using a gel extraction kit (Qiagen, Venlo, the Netherlands) and cloned into a linearized Topo PCR 2.1 vector with HiFi assembly cloning (NEB). Cloning reactions were transformed into XL10 *E. coli* (Stratagene/Agilent, Wilmington, DE, USA) and selected with Kanamycin. Plasmid DNAs were isolated from colonies (Qiagen Mini-Prep kit) and Sanger sequenced for verification. Using a site-directed mutagenesis (SDM) protocol,<sup>18</sup> 40 variants across exon 2 and 3 were introduced into the donor gBlock of the targeted region with homology arms. All primers and gRNAs used in this study are summarized in Appendix S1.

### HAP1 cell culture and growth selection

In this study, we used a LIG4-Knockout HAP1 cell line (Horizon Discovery, Waterbeach, UK) that has a higher knock-in efficiency,<sup>13</sup> and cultured them in Iscove's

Modified Dulbecco's medium (IMDM) with L-glutamine and 25 mM HEPES (GIBCO, Waltham, MA, USA) supplemented with 10% fetal bovine serum (Rocky Mountain Biologicals, Missoula, MT, USA) and 1% penicillin–streptomycin (GIBCO). Because HAP1 cells spontaneously convert to a diploid state in cell culture,<sup>20</sup> we sorted cells to generate a pure haploid (1n) cell population before transfection. To do so, cells were stained for DNA content with Hoechst 34580, and Fluorescence-Activated Cell Sorting (FACS) was performed to isolate the cells with the lowest intensity, corresponding to haploid cell population (Figure S1). These 1n cells were expanded for 2 weeks to obtain sufficient cells before transfection. HAP1 cells were transfected with Turbofectin 8.0 according to manufacturer's protocol. For each transfection, we performed three biological replicates with 20 million cells plated in a 10-cm dish per replicate. The day after plating the cells (D0) cells were co-transfected with 4.5 µg Cas9-gRNA plasmid and 4.5 µg of the donor pooled corresponding to a single exon (22 variants for exon 2, and 18 variants for exon 3). On day 1, the media was supplemented with puromycin (1 µg/mL), and the cells were treated for 48 h to select for successfully transfected cells. On day 5 post transfection, cells were washed once with 1× phosphate-buffered saline (PBS, GIBCO), trypsinized with 0.25% trypsin (GIBCO), and resuspended in medium. One-half cells were collected for sequencing at D5 (centrifuged for 5 min at 300 g), and the remainder were grown to D8. On day 8, half of the cells were pelleted for sequencing and the other half were passaged for harvesting at D11.

### gDNA preparation and sequencing

All genomic DNA (gDNA) was isolated from the cell population on days 5, 8, and 11 using the DNeasy kit (Qiagen). PCR primers for gDNA were designed outside of the homology arm sequence in order to select for amplicons derived from gDNA and not plasmid DNA. All gDNA were sampled by performing eight replicate PCR reactions, using 50 ng of gDNA per 10 µL reaction (NEB-Next 2× mastermix). After PCR, all replicates from the same sample were pooled and purified using Ampure XP. Next, a nested PCR was performed in duplicate with primers containing Illumina sequencing adapters followed by combined duplicates and then purified product. A final PCR was performed for the minimal number of cycles needed to complete amplification using purified products from the second reaction as template to add dual sample Illumina indexes and flow cell adapters. All sequencing libraries were purified with Ampure beads, diluted and denatured for sequencing for Illumina MiSeq and Novaseq machines.

### Sequencing alignment and quality control

Reads from Miseq and Novaseq were received in compressed Fastq format.<sup>21</sup> We used FastQC v0.11.7 to obtain quality control data including average base quality by position and by read.<sup>22</sup> The SeqPrep software with the default parameters and the added '-s' parameter merged the forward and reverse reads for each sample (<https://github.com/jstjohn>). We mapped the resulting merged reads using the BWA alignment software. The reference sequence for each alignment was obtained from the UCSC hg37 version of *SLC2A1* which consisted of the exonic sequence with a 20 bp intronic flanking sequence included.<sup>23</sup> The presence of the flanking sequences increased the accuracy of the alignment by preventing clipped sequences of mismatching base pairs in the first and last codon of the exons. Since the samples should only have the DNA sequence of the transfected donor, we calculated the percent of mapped reads as a quality control using the Flagstat module of Samtools v1.12. Experiments with >5% unmapped reads were rejected.

### Tabulation of mapped reads by read sequence variants

Each donor sequence contained at most one missense and two synonymous substitutions (markers) at PAM site in the same position for each exon. The synonymous substitutions were used as markers to distinguish SNVs caused by other experimental steps or sequencing error. A python script was created to classify each mapped read. The script was written using the PySam v0.17 module and executed using Python v3.5.3 (<https://github.com/pysam-developers/pysam>). Reads that were unmapped, including indels or those that did not cover the full exon sequence, were removed. The remaining reads were classified according to the nucleotide variants and the IUPAC nomenclature for the amino acid change when present. Reads without the synonymous markers or with more than one amino acid change were discarded. The script produced two tab-separated files for two technical PCR samples per timepoint for three biological replicates at days 5, 8, and 11: (1) a summary file with the count of unmapped, indel, discarded and knock-in reads, and (2) the count of each unique amino acid change and synonymous marker in the sequencing dataset from the experiments for each timepoint. The counts of each technical replicate were merged with an additional python script that compared the found variants in each timepoint. If a variant was not found in one of the sequencing dataset from the experiments, the count was zero. The merged tabulated data tables were used as the input for Enrich2<sup>24</sup> which used the change in frequency of the wild type and

**Table 1.** SLC2A1 variants with detailed clinical and/or diagnostic test data.

Variant/ Genomic coordinate	Mutation type	Exon Number	Sex	Clinical phenotype	CADD score	Glucose ratio in CSF/Blood (Absolute values, mg/ dL) <sup>b</sup>	Glucose uptake (Erythrocyte 3-OMG) % <sup>c</sup>	Surface expression (Metaglut1)	HAP1 growth assay functional score (Miseq data)	Clinical significance (ClinVar interpretation/ Inheritance mode)	References
p.N34S GRCh37: Chr1: 43408910	Missense	2	F	Atypical absence epilepsy with onset age 2 years, mild learning disabilities, paroxysmal nocturnal dyskinesia, mild cerebellar ataxia	26.5	0.39 (34/ 87)	ND	ND	-0.70	Pathogenic/AD	[32]
			M	Seizures started 8 weeks as weekly or monthly, microcephalic (-1 to 2 SDs), speaks in sentences at age 6 years with decreased vocabulary		ND	55	ND			[33]
			M	Seizure onset at age 3 months, mental retardation, hypotonia, normal head circumference	0.43 (32/ 74)		43	ND			[34]
			F	Early Onset Absence Epilepsy (EOAE) characterized by EEG in patient and two siblings with asymptomatic mother harboring the variant	ND	ND	67	+1 (NL)			Unpublished
p.N34K GRCh37: Chr1: 43408909	Missense	2	M	Few myoclonic seizures/absence seizures before age 5, lethargic episodes, delayed development, ataxia, dystonia, CNS score = 54 <sup>a</sup>	23.9	0.30 (28.8/ 96)	71	ND	-0.86	Likely pathogenic	[35]
p.N34I GRCh37: Chr1: 43408910	Missense	2	M	No epileptic seizures, normal EEG, delayed development, moderate mental retardation, dysarthric speech, ataxia	29.9	0.31 (28.8/ 92.9)	49	ND	-0.88	Pathogenic/de novo	[36]
p.Q46X GRCh37: Chr1: 43396856	Nonsense	3	Unknown	Encephalopathy due to Glut1DS	36	NA	NA	NA	-1.63	Pathogenic	<a href="https://www.ncbi.nlm.nih.gov/clinvar/variation/1027520">https://www.ncbi.nlm.nih.gov/clinvar/variation/1027520</a>
p.S66F GRCh37: Chr1: 43396794	Missense	3	M	CNS score = 58 <sup>a</sup>	26	0.35 (34/ 98)	45	ND	-0.79	NA	[35]

(Continued)

Table 1 Continued.

Variant/ Genomic coordinate	Mutation type	Exon Number	Sex	Clinical phenotype	CADD score	Glucose ratio in CSF/Blood (Absolute values, mg/ dL) <sup>b</sup>	Glucose uptake (Erythrocyte 3-OMG) % <sup>c</sup>	Surface expression (Metaglut1)	HAP1 growth assay functional score (Miseq data)	Clinical significance (ClinVar interpretation/ Inheritance mode)	References
p.W65R GRCh37: Chr1: 43,396,799	Missense	3	M	Developmental delay. Discomfort of unknown origin. Dystonia; Eye twitching at 3 months, Dysarthria; Dystonia; Tremor	26.1	0.45 (34/ 75)	112	+13 (NL)	-0.81	NA	Unpublished
p.S73F GRCh37: Chr1: 3396774	Missense	3	Unknown	Childhood onset Glut1DS	26.6	NA	NA	NA	-0.61	VUS	<a href="https://www.ncbi.nlm.nih.gov/clinvar/variation/989292">https:// www.ncbi. nlm.nih. gov/clinvar/ variation/ 989292</a> Unpublished
p.G76S GRCh37: Chr1: 43396765	Missense	3	F	Generalized tonic-clonic epilepsy, axial hypotonia and dystonia induced by effort	28.4	0.5 (45/90)	52	-42 (ABNL)	-0.64	NA	Unpublished
p.V87I GRCh37: Chr1: 43396733	Missense	3	F	Complex-partial seizures occurred at age of 3.5 years and stopped after meal, mild cognitive impairment and no movement disorders. Three other family members have this variant, including the mother with migraine and paroxysmal hemiparesis during childhood; later, she presented with PED. Twin siblings also have the variant; one has mild absence seizures and eyelid myoclonia along with mild developmental delay and cognitive impairment. The other twin sibling is unaffected but has the variant.	22.6	0.45 (NA)	NA	ND	0.23	VUS/Possibly AD with incomplete penetrance	[27]
G91D	Missense	3	F	Complex partial seizure started at age of 3 years, moderate mental retardation, ataxia, nystagmus and pyramidal signs	28.2	0.29 (34/ 118)	42	ND	-0.14	Pathogenic/AD	[37,38]

(Continued)

Table 1 Continued.

Variant/ Genomic coordinate	Mutation type	Exon Number	Sex	Clinical phenotype	CADD score	Glucose ratio in CSF/Blood (Absolute values, mg/ dL) <sup>b</sup>	Glucose uptake (Erythrocyte 3-OMG) % <sup>c</sup>	Surface expression (Metaglut1)	HAP1 growth assay functional score (Miseq data)	Clinical significance (ClinVar interpretation/ Inheritance mode)	References
GRCh37: Chr1: 43396720			F (Sister)	Generalized tonic clonic seizures, moderate mental retardation, ataxia and pyramidal signs		0.35 (34/ 97)	57				
			M (father)	Generalized tonic-clonic, myoclonic, mild mental retardation		ND	57				

3-OMG, 3-O-methyl-D-glucose; ABNL, abnormal; AD, autosomal dominant; CSF, cerebrospinal fluid; F, female; M, male; ND, not done; NA, not available; NL, normal; SD, standard deviation; VUS, variant of uncertain significance.

<sup>a</sup>CNS score provided by Dr. De Vivo; Severe phenotype (40–49), moderate phenotype (50–59), mild phenotype (60–69) and minimal or healthy (70–76).

<sup>b</sup>Normal range of glucose ratio CSF/blood: 0.6–0.7, and normal range of CSF glucose: 50–80 mg/dL.<sup>39</sup>

<sup>c</sup>Normal erythrocyte 3-OMG uptake %: 100 ± 22%.<sup>40</sup>

variant reads at each timepoint to calculate a linear regression. The slope of the linear regression for each variant is the functional score from one replicate, which is then combined using a meta-analysis model to calculate a final functional score that takes into account all replicate scores. The final scores and individual replicate scores are used to create figures. The Enrich2 software ran using the “identifiers only” mode and filtering for a minimum of one read of a variant in all sequencing dataset from the experiments. For software input, each replicate was identified as a selection event under the same condition and each selection event had days 5, 8 and 11 time points using day 5 as the initial condition for the calculations. The scores were computed using Enrich2’s weighted least squares method and normalized using the wild type reads in the each timepoint.

### Visualization of scores for each exon

In addition to the figures generated by Enrich2, we generated our own correlation plots using R’s “cor” function and visualized the resulting Spearman scores for all replicates. We also used the HDF5 formatted output files from Enrich2 to regenerate the linear regression plots for every replicate and variant. The individual replicate scores were also visualized with ggplot2 (<https://ggplot2.tidyverse.org>) as a scatter plot with error bars where each dot was color-coded by its functional effect.

### Statistical analysis

To measure if the difference between functional scores of the nonsense and synonymous variants were significant, we performed an ANOVA using the R function `av()` with the formula: `enrichScore ~ consequence + variantName`.

### Classification of SNV data

The functional scores were correlated with clinical phenotypic data using published cases and data shared by collaborators at Washington University (Drs. Weisenberg and Thio) and Columbia University (Dr. De Vivo), as well as data from ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar>). We also compared the performance of our functional scores with the computational algorithms of CADD score.<sup>25</sup>

### Results

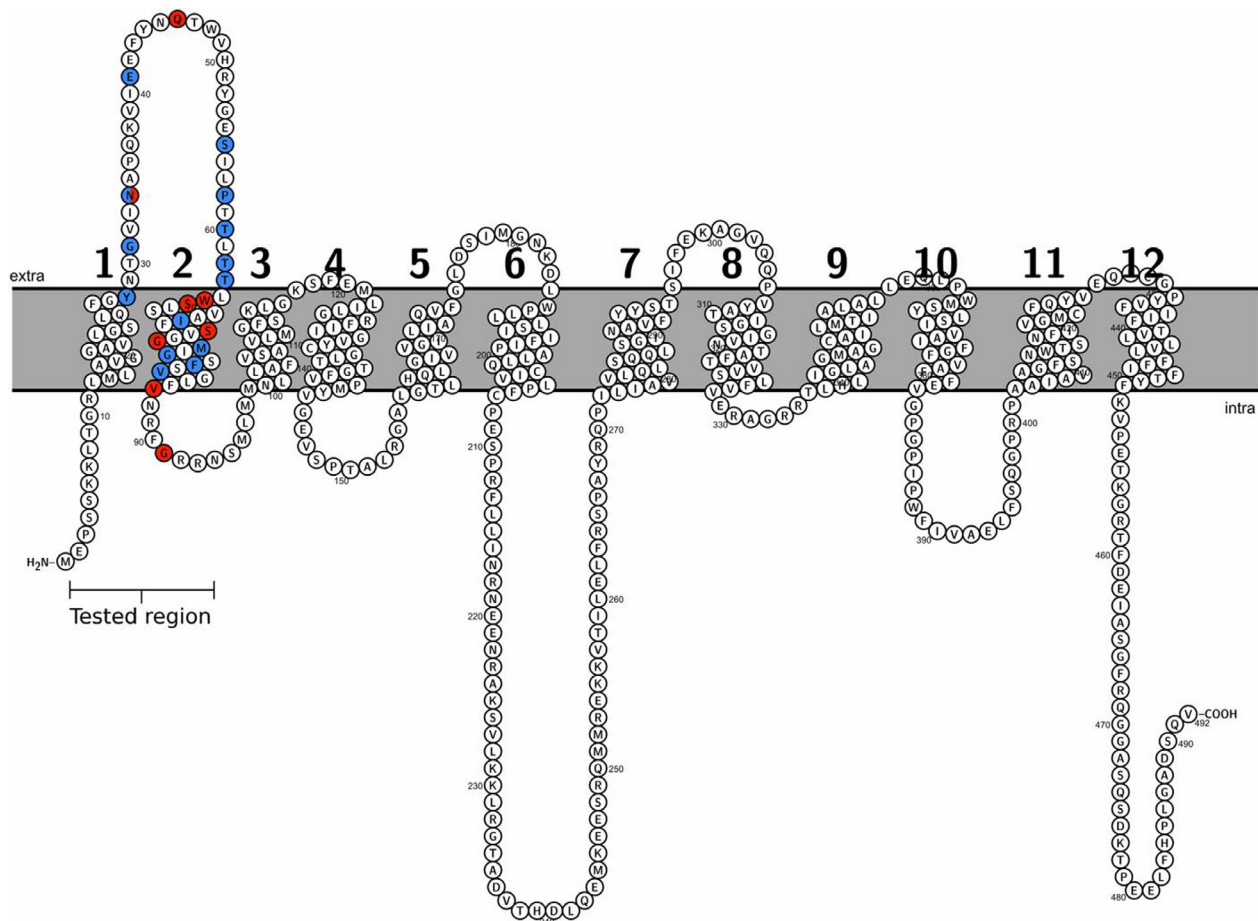
We used a growth assay in HAP1-Lig4KO cells to analyze the functional effects of 40 single nucleotide variants in *SLC2A1* (exon 2 and 3), including 6 reported in ClinVar as pathogenic/likely pathogenic, 7 benign/likely benign,



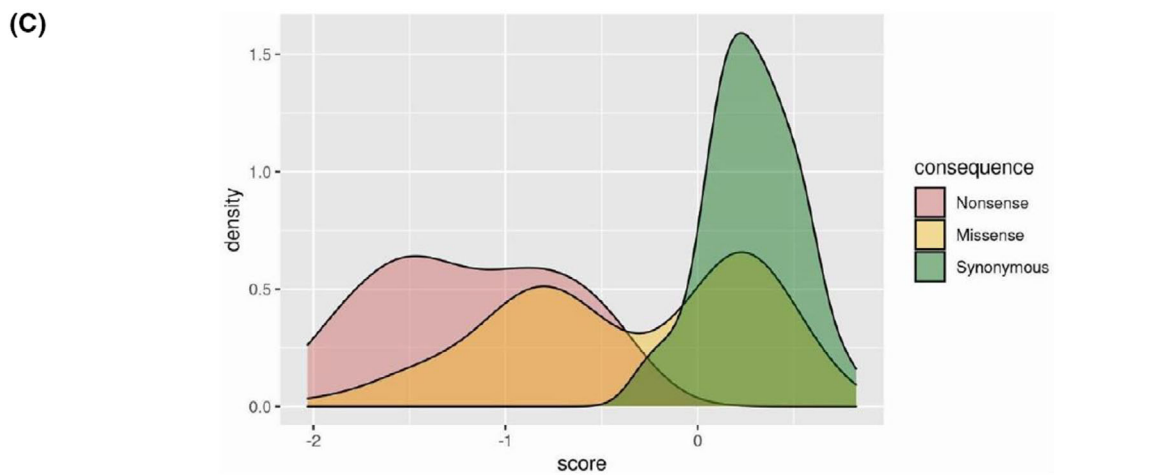
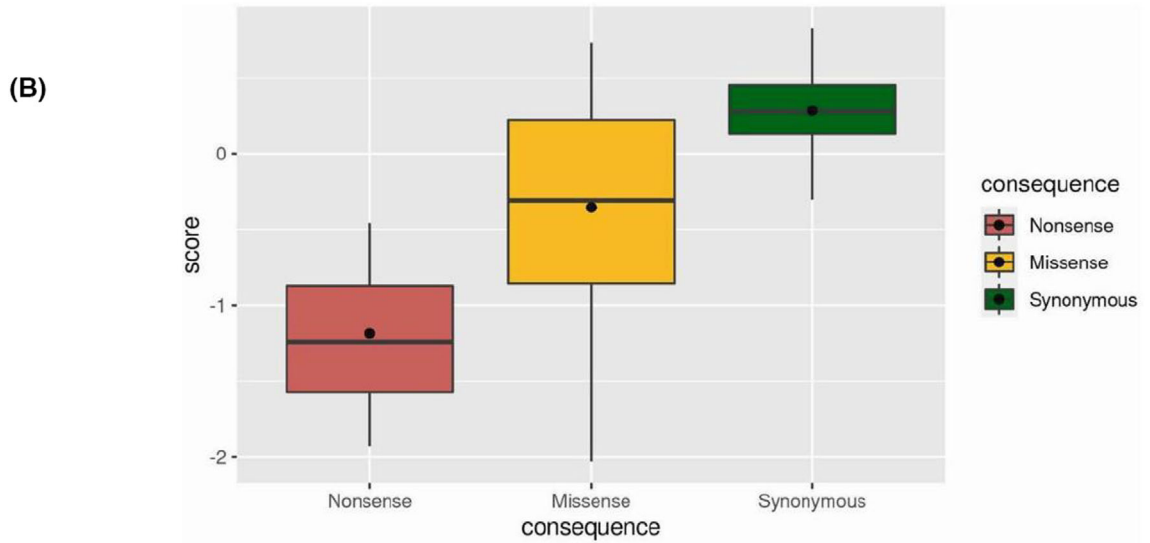
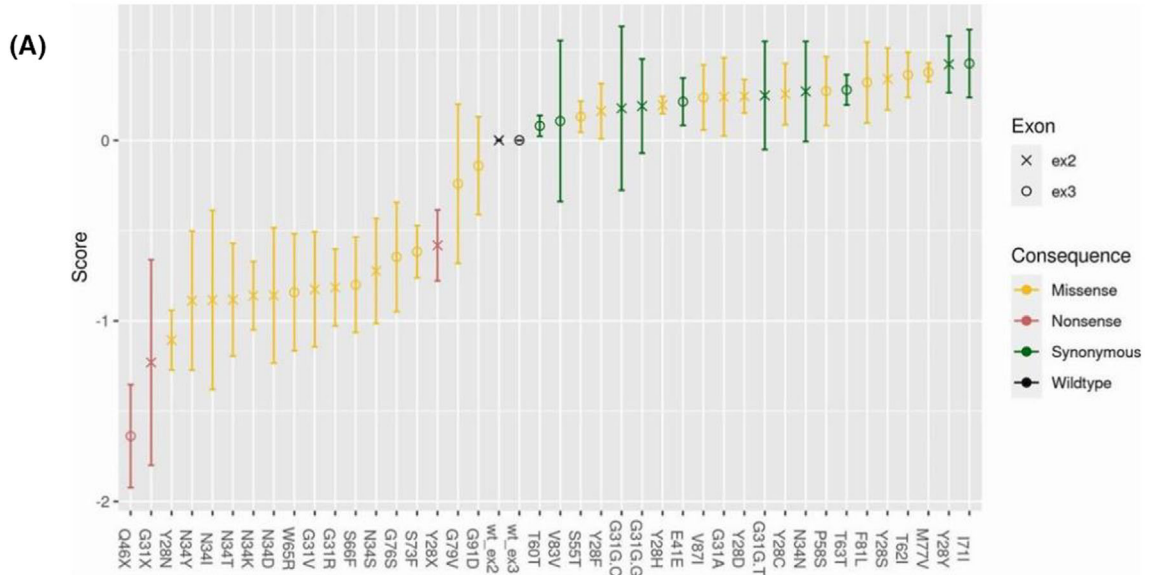
and 7 VUS (Table 1, Table S1 and Figure 1). Some variants in exon 2 were not previously reported but were designed to study the effects of nonsense variants or multiple amino acid substitutions at the same position. Clinical and/or *in vitro* functional data from cell surface expression or erythrocyte 3-OMG uptake assay were available for 10 variants (Table 1). Variants in either exon 2 (designed to study multiple amino acid changes at the same residue) or exon 3 were introduced, as a group, into the endogenous *SLC2A1* locus in a pool of cells using a single gRNA and Cas9 for each exon in three biological replicates. For each exon, <1% of reads contained frame-shift indels, and the homology directed repair (HDR) rate (defined by the presence of a single introduced nucleotide variant and associated nearby markers) ranged from 20% to 60% at different timepoints in each biological replicate (Figure S2A,B). As expected, HDR efficiency correlated with the proximity of the introduced variants to the gRNA cut site, as demonstrated by the lower HDR rate of exon 3 variants (mean HDR rate was 34.15%, and mean

distance of the 18 variants to the cut site was  $34 \pm 5.7$  bp) compared to exon 2 variants (the mean HDR rate was 50.2%, and the mean distance of the 22 variants to the cut site was  $16 \pm 1.5$  bp).

To determine the reproducibility of technical and biological replicates and the impact of sequencing method, we sequenced the PCR products (two technical PCR samples per timepoint for three biological replicates) at days 5, 8, and 11 using both Illumina Miseq and Novaseq sequencers. The sequencing depth was much greater on the Novaseq (the mean read depth of Miseq data for exons 2 and 3 was  $107 \pm 92$ , and  $313 \pm 228$ , respectively; the mean read depth of Novaseq data for exon 2 and 3 was  $7056 \pm 5736$  and  $30,269 \pm 21,738$ , respectively). We calculated functional scores via Enrich2 software which compares reads containing the variant at each of the three timepoints after normalization to wild type reads for each of the 40 *SLC2A1* variants. Using the Enrich2 software, the fitted line of the weighted linear regression for all 40 variants in each replicate at three timepoints was



**Figure 1.** GLUT1 protein diagram showing the locations of variants from exon 2 and 3 using DOG plotter. Variants identified in patients with clinical data are highlighted in red (Table 1), all others are blue (Table S1).





**Figure 2.** Functional scores of 40 *SLC2A1* single nucleotide variants derived from HAP1 growth assay. (A) Scores for individual variants (with 95th percentile error bars derived from Miseq data) are shown sorted from lowest score (most negative) to highest. (B) Functional scores of groups of nonsense ( $N = 3$ ), missense ( $N = 27$ ), and synonymous ( $N = 10$ ) variants. The black dot is the mean of the variants. (C) Density plot showing the distribution of individual score for each replicate of variant data within each group.

generated along with the calculation of slope (Figure S3A, B). The Spearman's correlations for biological replicate were between 0.46 and 0.72, and for technical replicates were between 0.45 and 0.95 for experiments with pooled variants in both exons 2 and 3. (Figure S4A,B). When comparing both methods of sequencing, the correlation was 0.92 (Figure S5B).

After confirming the reproducibility of the assay, we calculated functional scores for all variants (Figure 2A and Figure S5A). A group of variants remained stable in abundance from days 5 to 11, resulting in a score higher than 0 as a consequence of these cells becoming more predominant in the population as cells in the same pool as damaging variants dropped out. The synonymous variants (functional variants) were all in this group of non-damaging variants and had a mean score of  $0.25 \pm 0.12$  (Figure 2B). A second group of variants showed strongly negative scores ( $<-1$ ), with nonsense variants (non-functional variant) having the most negative scores of all variants tested. The mean score was  $-1.15 \pm 0.17$  for the three nonsense variants tested, and there were no significant differences between the nonsense variants (ANOVA test,  $P = 0.1025$ ) (Figure 2B). However, there is a significant difference between the synonymous and nonsense variants based on the ANOVA test ( $P < 2e-16$ ) (Figure 2B). Missense variants appeared in a bimodal distribution, with some having positive scores similar to the synonymous group, and others having more intermediate negative scores (Figure 2C).

Several missense variants scored intermediate between the nonsense and the synonymous variant groups (Figure 2A and Figure S5A). These include p.G79V (functional score =  $-0.24$ ), p.G91D (functional score =  $-0.14$ ), and p.S73F (functional score =  $-0.61$ ). Notably, p.G91D is an autosomal dominant inherited variant from a family with multiple affected individuals whose clinical phenotype included mild to moderate cognitive impairment, with one patient having later onset epilepsy (age 3 years) (Table 1). The patient with the p.S73F variant has childhood onset Glut1DS (<https://www.ncbi.nlm.nih.gov/clinvar/variation/989292>), an autosomal dominant disorder characterized by childhood onset exercise-induced dyskinesia<sup>26</sup> (Table 1). The G79V variant has not been reported clinically but was experimentally designed and tested (Table S1). Interestingly, we found a positive functional score suggesting that it is a benign variant for one missense variant (p.V87I) that had

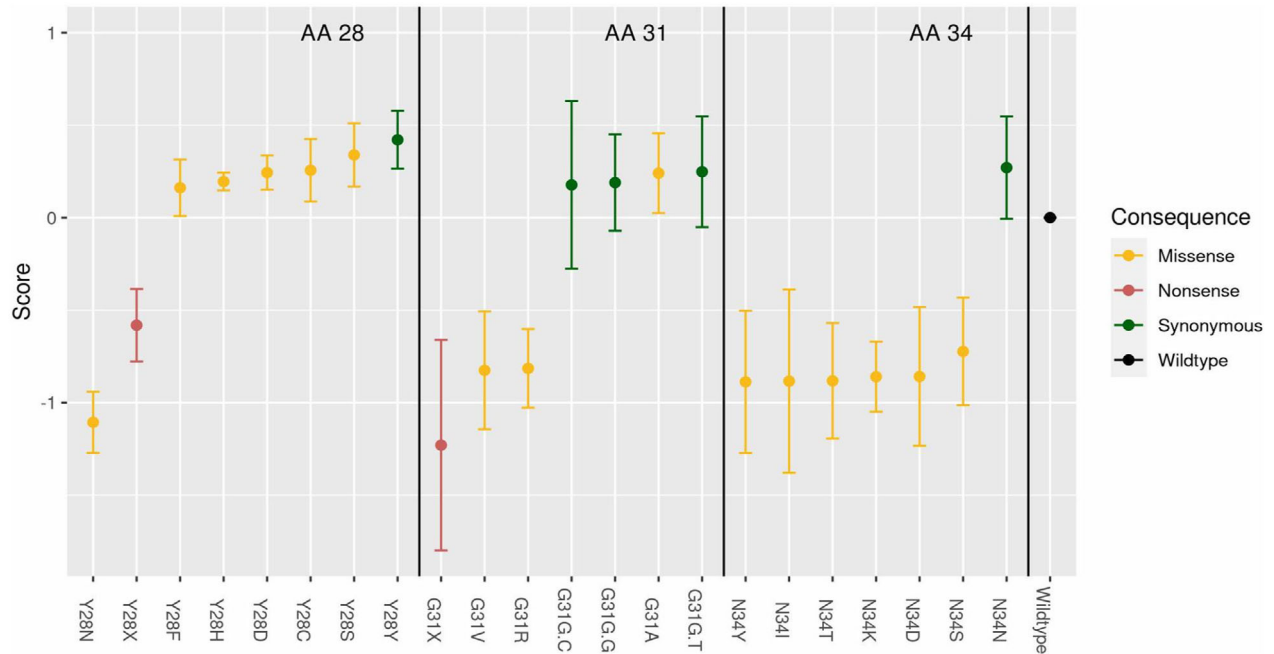
previously been identified in a child with a mild phenotype, including focal epilepsy with onset at 3.5 years old and mild cognitive impairment.<sup>27</sup> The variant was also identified in the child's mother who had migraine, paroxysmal hemiparesis, and late onset PED, and one of a pair of twin siblings who had mild absence epilepsy, eyelid myoclonia, mild developmental delay and cognitive impairment. However, the other unaffected twin also had the variant, making this VUS difficult to interpret and less likely to be pathogenic.

We specifically designed variants in exon 2 to allow us to study the effects of multiple amino acid substitutions at the same position (Figure 3). Interestingly, all missense variants resulting in an amino acid substitution at position 34 amino acid had negative functional scores. In contrast, only 2 of the 7 variants tested at position 28 were damaging. Of note, while Y28X, which has not been clinically reported, resulted in a negative functional score, it was less deleterious compared to other nonsense variants (Figure 2A and Figure S5A). Notably, p.Y28X is the most proximal of the three nonsense variants that we tested.

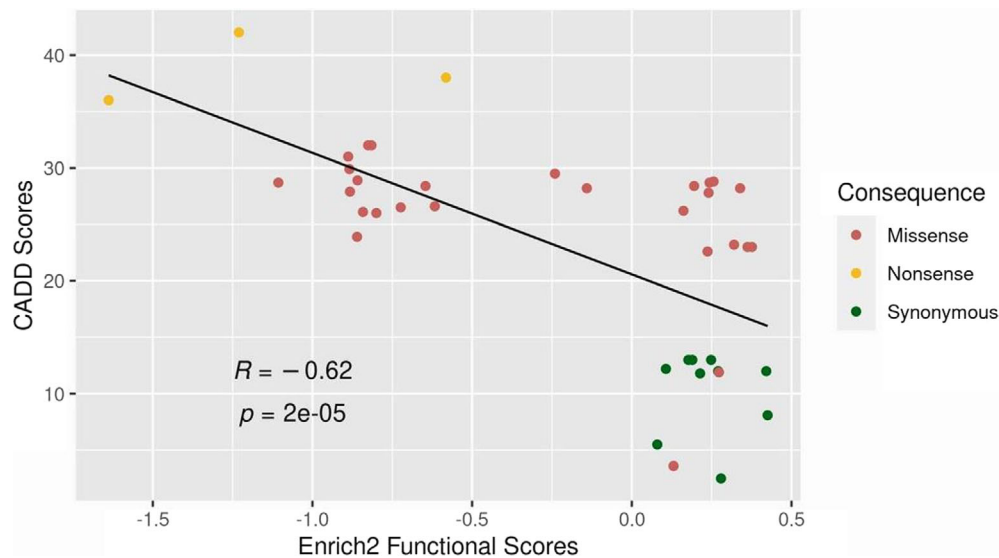
We then compared our functional assay performance with several computational metrics that are currently used to investigate deleteriousness of variants, including CADD score.<sup>25</sup> There is a moderate but significant correlation between our experimentally derived functional scores and the CADD scores (Figure 4).

To compare our functional assay to other *in vitro* and structural studies of GLUT1, we evaluated variants that had normal results on either erythrocyte 3-OMG uptake or GLUT1 surface expression assays. Two variants were reported in Glut1DS patients who had abnormal CSF/blood glucose ratios; one was reported in ClinVar (p.N34S), but the other is reported for the first time here (p.W65R) (Table 1). Both variants had negative functional scores on our growth assay. These included p.N34S, which had normal surface expression and a functional score of  $-0.70$ , and p.W65R, which had normal surface expression as well as normal erythrocyte 3-OMG uptake and a functional score of  $-0.81$  (Figure S3A, B).

We also generated an area under the receiver operating characteristic (ROC) curve (AUC) based on 18 variants including 10 synonymous and 8 pathogenic/likely pathogenic (3 nonsense and 5 missense) to illustrate the sensitivity and specificity of our growth assay. Our assay has an AUC of 1 (Figure 5).



**Figure 3.** Functional scores for variants altering three amino acid residues (Y28, G31, and N34) in exon 2. Data presented is the % CI as calculated from Miseq sequencing data.

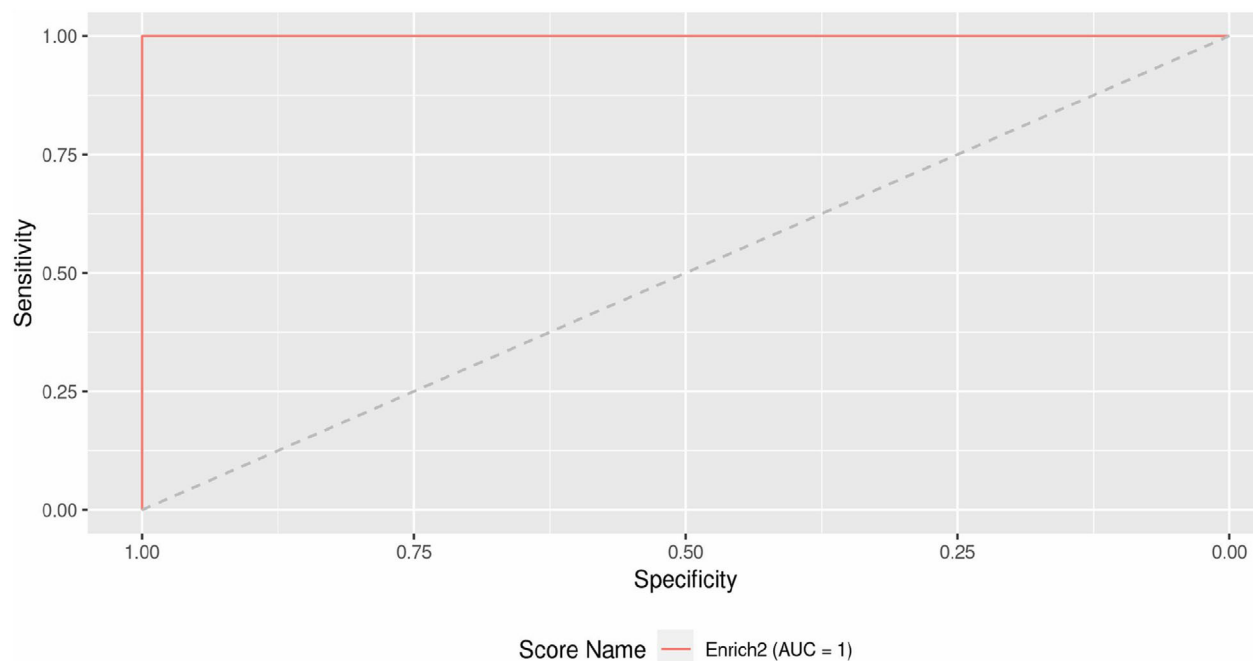


**Figure 4.** The functional scores of 40 *SLC2A1* variants are inversely correlated with CADD scores (Spearman’s correlation:  $-0.62$ ). Scores are derived from Miseq data.

## Discussion

The American College of Medical Genetics guidelines emphasize the importance of functional evidence for interpretation of gene variants.<sup>41</sup> Here, we describe the

reproducibility and utility of a growth assay to quantify the functional effects of 40 variants in *SLC2A1*. Our results demonstrate that this functional assay is highly reproducible between replicates and different sequencing methods, and yields negative scores for nonsense variants while



**Figure 5.** Receiver operating characteristic curve for GLUT1 growth assay in HAP1 cells yields an AUC of 1.

synonymous variants are all non-negative, and has significant advantages over existing *in vitro* functional assays which have inherent limitations.

There are multiple advantages of this *in vitro* HAP1 growth assay for determining the functional effects of *SLC2A1* variants. First, the assay does not require patient biospecimens. Currently, the gold standard diagnostic test for Glut1DS is hypoglycorrachia, or reduced CSF glucose concentration. However, CSF requires an invasive spinal tap, which may require sedation for infants and young children, and adds both risk and expense. Second, the HAP1 growth assay can be scaled to study a large number of variants. While glucose uptake can be quantified *in vitro* using a patient's own erythrocytes or by exogenous expression in *Xenopus* oocytes,<sup>42</sup> these assays were not designed for the high throughput screening and are performed by only a handful of labs worldwide.

A major advantage of the HAP1 growth assay is that it integrates the multiple ways that a gene variant can disrupt function which makes it highly valuable for assessing the effects of gene variants. For *SLC2A1* and the other genes required for growth in HAP1 cells, such as *BRCA1*, variants that result in nonsense-mediated decay of the transcript, protein instability, or impaired function (i.e., glucose transport) of a protein that otherwise traffic appropriately to the cell surface, all result in impaired growth. The growth assay is advantageous over those that quantify surface expression (Metaglut1 test),<sup>43</sup> which

detects as abnormal only those variants that impair protein synthesis, stability or trafficking. For example, the autosomal dominant pathogenic *SLC2A1* p.G91D variant<sup>32,34</sup> demonstrated impaired cell growth in our assay despite its normal erythrocyte surface expression.<sup>37,38</sup> In addition, there are reports of normal erythrocyte 3-OMG uptake assays in patients with Glut1DS phenotypes, presumably due to normal uptake at the non-physiologic lower temperature (4°C) at which the assay is performed.<sup>44–46</sup> Indeed, our growth assay demonstrated impaired function of the *SLC2A1* p.W65R variant in an individual with a Glut1DS phenotype in which erythrocyte glucose uptake and surface expression was normal, highlighting the advantages of our growth assay over existing functional tests.

Functional data requires calibration with known benign and pathogenic variants to allow it to be useful for informing variant classification. Because variants in *SLC2A1* cause a spectrum of neurological disease, functional scores would ideally correlate with clinical outcome. Previous functional studies of the breast cancer gene *BRCA1* proposed a strictly binary (pathogenic/benign) mode.<sup>11</sup> The quantitative scores we obtained for our positive controls (nonsense variants) and negative controls (synonymous variants) were consistent with expected values and did not overlap, resulting in a highly predictive ROC curve. Because several *SLC2A1* missense variants identified in patients with autosomal dominant

inheritance and less severe phenotypes resulted in scores that were intermediate between the positive and negative controls, we are optimistic that our quantitative functional scores may prove to be predictive of neurological outcome in larger studies, as has recently been shown for *DDX3X*.<sup>15</sup> Although the genotype–phenotype correlation in Glut1DS is complex and clinical variability is present even within families,<sup>47,48</sup> patients with missense variants often have milder phenotypes. However, individual outcomes are likely influenced by genetic modifiers and environmental factors such as seizure control. Clearly, a much larger dataset that includes quantitative scores, such as IQ, are needed to prove correlations between quantitative functional scores on the growth assay and clinical outcome.

Although the Glut1DS mechanism is haploinsufficiency and the growth assay appears sufficiently sensitive to detect deleterious variants, we expect that the HAP1 growth assay may, by nature of it being in a haploid cell line, be less effective for evaluating variants with dominant-negative effect or gain of function mechanisms. However, to date, there is no evidence for dominant-negative effect or gain of function mechanisms for Glut1DS. We also do not know whether our assay is sufficiently sensitive for variants that show incomplete clinical penetrance. Prior work on *DDX3X*<sup>15</sup> variants, studied using the same assay, suggested that mildly damaging variants may be slowly depleting and required modifications of the growth assay. Our functional data for *SLC2A1* p.V87I, which is a VUS reported in three affected and one unaffected individual in a small family,<sup>27</sup> suggest that this variant is benign. The clinical phenotype in this family, while consistent with mild Glut1DS, is not specific for this syndrome given the many causes of mild cognitive impairment and migraine. Even hypoglycorrhachia is nonspecific, as some patients with hypoglycorrhachia do not carry mutations in *SLC2A1*.<sup>49,50</sup> Therefore, we cannot exclude the possibility that this variant occurred in a family with a condition that phenocopies Glut1DS, or that this variant has an effect on *SLC2A1* that is too small to detect with our assay. Another possible limitation of our assay is that some of the effects that we are seeing on cell growth are due to off-target effects, though we selected the gRNAs based on high on-target and low off-target scores and have shown similar findings across three biological replicates.

The results of our functional assay also provide insight into protein function and disease mechanisms. Multiple mechanisms may account for the effects of missense variants on GLUT1 activity, including direct effect of steric interference with substrate access to the glucose permeation pathway (likely important only for a small number of transporter amino acids), indirect effects through

changes in helix packing in the membrane, altered folding, decreased stability/increased degradation and changes in conformational dynamics. For instance, we demonstrated that all possible missense variants at N34 have negative functional scores consistent with this being a hot spot for disease due to the importance of this amino acid in hydrogen bonding with glucose at the exofacial sugar binding site.<sup>51,52</sup> In contrast, only one of the six possible amino acids substitutions at Y28, located within the transmembrane domain, was functionally damaging in our assay. Although prior work on cysteine-scanning mutagenesis showed this residue to be functionally important,<sup>53</sup> our results on individual amino acid substitutions would not have been predicted based on CADD scores. In addition, we cannot fully explain why the functional score for p.Y28X was consistently less negative than the other two nonsense variants we tested, though they were statistically not significantly different. We note that p.Y28X is the most proximal of the three nonsense variants, leading us to consider the possibility that in an alternative downstream start site may yield a protein with a possible partial function.<sup>28–31</sup>

The simplicity of the HAP1 growth assay will enable its use to evaluate all possible variants in *SLC2A1* in future studies through deep mutational scanning (DMS). DMS refers to the approach where all possible variants in a sequence are created and functionally assayed.<sup>54,55</sup> Powered by rapid decreases in the cost of DNA sequencing and DNA synthesis, the highly parallel design of DMS reduces the cost per variant per assay by orders of magnitude.<sup>56</sup> With DMS, the effects of individual variants are directly comparable since all variants are assayed at the same time and in the same system, eliminating problems associated with comparing results across labs.<sup>57</sup> In contrast to one-at-a-time assays, DMS generates a “look-up table” of all possible variants, which is immediately useful to clinicians. Generation of a look-up table with functional scores for all possible variants will improve genetic diagnosis and may reduce the need for additional invasive confirmatory testing, such as CSF evaluation, or experimental generation of *in vitro* functional data. Accordingly, DMS is beginning to be used to assist clinical interpretations, as demonstrated by proof-of-principle papers involving the RING domain of *BRCA1*, and *PPARG* genes.<sup>13,58,59</sup> The performance of DMS is consistently superior to computational prediction programs or assessments of pathogenicity based on minor allele frequency. DMS provides a quantitative measure of variant effect, and the large scale of the resulting data provides the statistical power to assess the sensitivity and specificity of each assay. The results from DMS are easily integrated with ongoing efforts to build probabilistic models for variant interpretation that combine existing information

sources, including allele frequency data, family history, physical and chemical properties, and structural data, among others.<sup>60</sup>

Access to quantitative functional data at the time of diagnostic sequencing is essential for reducing delays in diagnosis, improving access to preventive treatment strategies, and optimizing enrollment in clinical trials. The growth assay described here for *SLC2A1*, and the quantitative functional scores that can be now generated for all variants across the gene, will markedly improve the clinical care of patients with Glut1DS.

## Acknowledgements

We thank the patients and their families for their role in this work. We thank the genome Technology Access Center at the McDonnell Genome Institute at Washington University for providing Illumina MiSeq and NovaSeq sequencer support.

## Funding Information

This work was supported by a research grant from the University of Pennsylvania Orphan Disease Center in partnership with the Team Glut1, Miles for Millie, Mission for Macie, and the Glut1 Deficiency Foundation. Research reported in this publication was also supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number P50 HD103525 to the Intellectual and Developmental Disabilities Research Center at Washington University, the National Institute of Arthritis and Musculoskeletal and Skin Disease R01AR067715, and by the Washington University Institute of Clinical and Translational Sciences grant UL1TR002345 from the National Center for Advancing Translational Sciences (NCATS) of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This work was also supported by funds provided by the McDonnell Center for Cellular and Molecular Neurobiology at Washington University in St. Louis.

## Author Contributions

NT, BLR, KM, GH, and CAG designed the experiment, performed data analysis and interpretation, wrote the manuscript; KE, VH, JW, LLT, DCDV, VP, and CAG provided the clinical information from patients for this research; EM, TNT, PH, and RSBW provided additional advice on the manuscript; all authors edited the manuscript.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## Ethics approval statement

The study was approved by the Washington University in Saint Louis Institutional Review Board, IRB #201102406.

## References

1. Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet.* 2014;46:310-315.
2. Grimm DG, Azencott CA, Aicheler F, et al. The evaluation of tools used to predict the impact of missense variants is hindered by two types of circularity. *Hum Mutat.* 2015;36:513-523.
3. van der Velde KJ, de Boer EN, van Diemen CC, et al. GAVIN: gene-aware variant interpretation for medical sequencing. *Genome Biol.* 2017;18:6.
4. Itan Y, Shang L, Boisson B, et al. The mutation significance cutoff: gene-level thresholds for variant predictions. *Nat Methods.* 2016;13:109-110.
5. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405-424.
6. Li MM, Datto M, Duncavage EJ, et al. Standards and guidelines for the interpretation and reporting of sequence variants in cancer: a joint consensus recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists. *J Mol Diagn.* 2017;19:4-23.
7. De Vivo DC, Trifiletti RR, Jacobson RI, Ronen GM, Behmand RA, Harik SI. Defective glucose transport across the blood-brain barrier as a cause of persistent hypoglycorrhachia, seizures, and developmental delay. *N Engl J Med.* 1991;325:703-709.
8. Maher F, Vannucci SJ, Simpson IA. Glucose transporter proteins in brain. *FASEB J.* 1994;8:1003-1011.
9. Verrotti A, D’Egidio C, Agostinelli S, Gobbi G. Glut1 deficiency: when to suspect and how to diagnose? *Eur J Paediatr Neurol.* 2012;16:3-9.
10. Kramer J, Smith L. Ketogenic diet in Glut 1 deficiency through the life cycle: pregnancy to neonate to preschooler. *Child Neurol Open.* 2021;8:2329048X211034655.
11. Kolic I, Radic Nisevic J, Vlasic Cicvaric I, et al. GLUT1 deficiency syndrome-early treatment maintains cognitive development? (Literature Review and Case Report). *Genes (Basel).* 2021;12:1379.



12. Blomen VA, Majek P, Jae LT, et al. Gene essentiality and synthetic lethality in haploid human cells. *Science*. 2015;350:1092-1096.
13. Findlay GM, Daza RM, Martin B, et al. Accurate classification of BRCA1 variants with saturation genome editing. *Nature*. 2018;562:217-222.
14. Meitlis I, Allenspach EJ, Bauman BM, et al. Multiplexed functional assessment of genetic variants in CARD11. *Am J Hum Genet*. 2020;107:1029-1043.
15. Radford EJ, Tan HK, Andersson MHL, et al. Saturation genome editing of DDX3X clarifies pathogenicity of germline and somatic variation. *medRxiv*; 2022. doi:10.1101/2022.06.10.22276179
16. Fayer S, Horton C, Dines JN, et al. Closing the gap: systematic integration of multiplexed functional data resolves variants of uncertain significance in BRCA1, TP53, and PTEN. *Am J Hum Genet*. 2021;108:2248-2258.
17. Kim HK, Lee EJ, Lee YJ, et al. Impact of proactive high-throughput functional assay data on BRCA1 variant interpretation in 3684 patients with breast or ovarian cancer. *J Hum Genet*. 2020;65:209-220.
18. Liu H, Naismith JH. An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol*. 2008;8:91.
19. Heckman KL, Pease LR. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc*. 2007;2:924-932.
20. Beigl TB, Kjosas I, Seljeseth E, Glomnes N, Aksnes H. Efficient and crucial quality control of HAP1 cell ploidy status. *Biol Open*. 2020;9:bio057174.
21. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25:2078-2079.
22. Leggett RM, Ramirez-Gonzalez RH, Clavijo BJ, Waite D, Davey RP. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. *Front Genet*. 2013;4:288.
23. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*. 2009;25:1754-1760.
24. Rubin AF, Gelman H, Lucas N, et al. A statistical framework for analyzing deep mutational scanning data. *Genome Biol*. 2017;18:150.
25. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. 2019;47:D886-D894.
26. Wang D, Pascual JM, De Vivo D. Glucose transporter type 1 deficiency syndrome. In: Adam MP, Mirzaa GM, Pagon RA, et al., eds. *GeneReviews*(®). National Library of Medicine, National Center for Biotechnology Information, US; 1993.
27. Wolking S, Becker F, Bast T, et al. Focal epilepsy in glucose transporter type 1 (Glut1) defects: case reports and a review of literature. *J Neurol*. 2014;261:1881-1886.
28. Bowling A, Eastman A, Merlo C, et al. Downstream alternate start site allows N-terminal nonsense variants to escape NMD and results in functional recovery by readthrough and modulator combination. *J Pers Med*. 2022;12:1448.
29. Zhang J, Maquat LE. Evidence that translation reinitiation abrogates nonsense-mediated mRNA decay in mammalian cells. *EMBO J*. 1997;16:826-833.
30. Santagata S, Gomez CA, Sobacchi C, et al. N-terminal RAG1 frameshift mutations in Omenn's syndrome: internal methionine usage leads to partial V(D)J recombination activity and reveals a fundamental role in vivo for the N-terminal domains. *Proc Natl Acad Sci USA*. 2000;97:14572-14577.
31. Rinne T, Clements SE, Lamme E, et al. A novel translation re-initiation mechanism for the p63 gene revealed by amino-terminal truncating mutations in Rapp-Hodgkin/Hay-Wells-like syndromes. *Hum Mol Genet*. 2008;17:1968-1977.
32. Tzadok M, Nissenkorn A, Porper K, et al. The many faces of Glut1 deficiency syndrome. *J Child Neurol*. 2014;29:349-359.
33. Wang D, Pascual JM, Yang H, et al. Glut-1 deficiency syndrome: clinical, genetic, and therapeutic aspects. *Ann Neurol*. 2005;57:111-118.
34. Klepper J, Scheffer H, Leiendecker B, et al. Seizure control and acceptance of the ketogenic diet in GLUT1 deficiency syndrome: a 2- to 5-year follow-up of 15 children enrolled prospectively. *Neuropediatrics*. 2005;36:302-308.
35. Pascual JM, Wang D, Yang R, Shi L, Yang H, De Vivo DC. Structural signatures and membrane helix 4 in GLUT1: inferences from human blood-brain glucose transport mutants. *J Biol Chem*. 2008;283:16732-16742.
36. Overweg-Plandsoen WC, Groener JE, Wang D, et al. GLUT-1 deficiency without epilepsy--an exceptional case. *J Inher Metab Dis*. 2003;26:559-563.
37. Klepper J, Willemsen M, Verrips A, et al. Autosomal dominant transmission of GLUT1 deficiency. *Hum Mol Genet*. 2001;10:63-68.
38. Klepper J, Monden I, Guertsen E, Voit T, Willemsen M, Keller K. Functional consequences of the autosomal dominant G272A mutation in the human GLUT1 gene. *FEBS Lett*. 2001;498:104-109.
39. Nakamura S, Osaka H, Muramatsu S, Aoki S, Jimbo EF, Yamagata T. Mutational and functional analysis of glucose transporter I deficiency syndrome. *Mol Genet Metab*. 2015;116:157-162.
40. Klepper J, Garcia-Alvarez M, O'Driscoll KR, et al. Erythrocyte 3-O-methyl-D-glucose uptake assay for diagnosis of glucose-transporter-protein syndrome. *J Clin Lab Anal*. 1999;13:116-121.
41. Brnich SE, Abou Tayoun AN, Couch FJ, et al. Clinical genome resource sequence variant interpretation working G. recommendations for application of the functional



- evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Med.* 2019;12:3.
42. Zaman SM, Mullen SA, Petrovski S, et al. Development of a rapid functional assay that predicts GLUT1 disease severity. *Neurol Genet.* 2018;4:e297.
  43. Gras D, Cousin C, Kappeler C, et al. A simple blood test expedites the diagnosis of glucose transporter type 1 deficiency syndrome. *Ann Neurol.* 2017;82:133-138.
  44. Wang D, Yang H, Shi L, et al. Functional studies of the T295M mutation causing Glut1 deficiency: glucose efflux preferentially affected by T295M. *Pediatr Res.* 2008;64:538-543.
  45. Cunningham P, Naftalin RJ. Implications of aberrant temperature-sensitive glucose transport via the glucose transporter deficiency mutant (GLUT1DS) T295M for the alternate-access and fixed-site transport models. *J Membr Biol.* 2013;246:495-511.
  46. Fujii T, Morimoto M, Yoshioka H, et al. T295M-associated Glut1 deficiency syndrome with normal erythrocyte 3-OMG uptake. *Brain Dev.* 2011;33:316-320.
  47. Mullen SA, Suls A, De Jonghe P, Berkovic SF, Scheffer IE. Absence epilepsies with widely variable onset are a key feature of familial GLUT1 deficiency. *Neurology.* 2010;75:432-440.
  48. Pearson TS, Akman C, Hinton VJ, Engelstad K, De Vivo DC. Phenotypic spectrum of glucose transporter type 1 deficiency syndrome (Glut1 DS). *Curr Neurol Neurosci Rep.* 2013;13:342.
  49. Klepper J, Leiendecker B. GLUT1 deficiency syndrome--2007 update. *Dev Med Child Neurol.* 2007;49:707-716.
  50. Leen WG, de Wit CJ, Wevers RA, et al. Child neurology: differential diagnosis of a low CSF glucose in children and young adults. *Neurology.* 2013;81:e178-e181.
  51. Galochkina T, Ng Fuk Chong M, Challali L, Abbar S, Etchebest C. New insights into GluT1 mechanics during glucose transfer. *Sci Rep.* 2019;9:998.
  52. Fu X, Zhang G, Liu R, et al. Mechanistic study of human glucose transport mediated by GLUT1. *J Chem Inf Model.* 2016;56:517-526.
  53. Heinze M, Monden I, Keller K. Cysteine-scanning mutagenesis of transmembrane segment 1 of glucose transporter GLUT1: extracellular accessibility of helix positions. *Biochemistry.* 2004;43:931-936.
  54. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. *Nat Methods.* 2014;11:801-807.
  55. Haller G, Alvarado D, McCall K, Mitra RD, Dobbs MB, Gurnett CA. Massively parallel single-nucleotide mutagenesis using reversibly terminated inosine. *Nat Methods.* 2016;13:923-924.
  56. Gasperini M, Starita L, Shendure J. The power of multiplexed functional analysis of genetic variants. *Nat Protoc.* 2016;11:1782-1787.
  57. Harrison SM, Dolinsky JS, Knight Johnson AE, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med.* 2017;19:1096-1104.
  58. Starita LM, Young DL, Islam M, et al. Massively parallel functional analysis of BRCA1 RING domain variants. *Genetics.* 2015;200:413-422.
  59. Majithia AR, Tsuda B, Agostini M, et al. Prospective functional classification of all possible missense variants in PPAR $\gamma$ . *Nat Genet.* 2016;48:1570-1575.
  60. Tavtigian SV, Greenblatt MS, Harrison SM, et al. Modeling the ACMG/AMP variant classification guidelines as a Bayesian classification framework. *Genet Med.* 2018;20:1054-1060.

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1.** Selection of haploid cells for HAP1 growth assay. To purify HAP1-Lig4KO cells with haploid chromosomal number, live cells were stained for DNA content with Hoechst 34580 and sorted to isolate cells with the lowest DNA content, corresponding to 1n cells (20%) in G1.

**Figure S2.** Knock-in efficiency of HAP1-LIG4KO cells. The average HDR % for three biological replicates in each timepoint for (A) 22 variants in exon 2 and (B) 18 variants in exon 3 of *SLC2A1* gene. The bar indicates standard error.

**Figure S3.** Quantification of effects of variants on growth over time. Representative plots of Enrich2's fitted line of the weighted linear least square regression for each variant in (A) exon 2 and (B) exon 3.

**Figure S4.** Correlation between biological and technical replicates. Spearman's correlation of the Enrich2 functional scores for all three timepoints (D5, D8 and D11) between three biological replicates (1–3, and) and two technical replicates (A and B) for exon 2 (A) and exon 3 (B) of *SLC2A1* gene sequenced with Miseq.

**Figure S5.** Functional scores and correlation using Nova-seq data. (A) Scores for individual variants (with 95th percentile error bars derived from Novaseq data) are shown sorted from lowest score (most negative) to highest. (B) Correlation of the Enrich2 functional scores between read counts from Miseq and Novaseq sequencing data.

**Table S1.** *SLC2A1* variants with no clinical information or that were not previously reported in patients.

**Appendix S1.** All primers and gRNAs used in this study are summarized in this file.