Chapter 2.1

Functional characterization of missense variants in the creatine transporter gene (SLC6A8): improved diagnostic application.

*Human Mutation, 2007*
RESEARCH ARTICLE

Functional Characterization of Missense Variants in the Creatine Transporter Gene (SLC6A8): Improved Diagnostic Application

Efraim H. Rosenberg,1 Cristina Martínez Muñoz,1 Ofir T. Betsalel,1 Silvy J.M. van Dooren,1 Matilde Fernandez,1 Cornelis Jakobs,1 Ton J. deGrauw,2 Tjitske Kleefstra,2 Charles E. Schwartz,4 and Gajja S. Salomons1*

1Department of Clinical Chemistry, Metabolic Unit, VU University Medical Center, Amsterdam, The Netherlands; 2Division of Neurology, Cincinnati Children’s Hospital Medical Center, Cincinnati, Ohio; 3Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; 4J.C. Self Research Institute, Greenwood Genetic Center, Greenwood, South Carolina

Communicated by Ronald Wanders

Creatine transporter deficiency is an X-linked mental retardation disorder caused by mutations in the creatine transporter gene (SLC6A8). So far, 20 mutations in the SLC6A8 gene have been described. We have developed a diagnostic assay to test creatine uptake in fibroblasts. Additionally, we expanded the assay to characterize novel SLC6A8 missense variants. A total of 13 variants were introduced in the SLC6A8 cDNA by site-directed mutagenesis. All variants were transiently transfected in SLC6A8-deficient fibroblasts and tested for restoration of creatine uptake in deficient primary fibroblasts. Thus, we proved that nine variants (p.Gly87Arg, p.Phe107del, p.Tyr317X, p.Asn336del, p.Cys337Trp, p.Ile347del, p.Pro390Leu, p.Arg391Trp, and p.Pro554-Leu) are pathogenic mutations and four variants (p.Lys4Arg, p.Gly26Arg, p.Met560Val, and p.Val629Leu) are nonpathogenic. The present study provides an improved diagnostic tool to classify sequence variants of unknown significance. Hum Mutat 28(9), 890–896, 2007. © 2007 Wiley-Liss, Inc.

KEY WORDS: SLC6A8; creatine transporter; mental retardation; diagnostic; site directed mutagenesis; solute carrier

INTRODUCTION

Deficiency of the creatine transporter gene (SLC6A8/CRT/CT1; MIM# 300336) is an X-linked form of the cerebral creatine deficiency syndrome (CCDS). Common clinical features are mental retardation (MR) expressive speech and language delay, autistic behavior, and/or epilepsy. Creatine (Cr) and phosphocreatine play essential roles in the storage and transmission of phosphate-bound energy [Walker, 1979; Wyss and Kaddurah-Daouk, 2000]. Cellular creatine transport is of fundamental importance for Cr homeostasis in tissues void of robust creatine biosynthesis. The uptake of creatine is mediated by the creatine transporter protein. The SLC6A8 gene (GenetID 6535) has been mapped to Xq28 [Gregor et al., 1995]; it spans ~8.4 kb, consists of 13 exons, and encodes a protein of 635 amino acids with a predicted molecular weight of 70 kDa [Sandoval et al., 1996]. Since the description of the first patient [Salomons et al., 2001], 20 mutations have been described throughout the SLC6A8 gene [Anselm et al., 2006; Bizzz et al., 2002; Clark et al., 2006; Mancini et al., 2005; Poo-Arguelles et al., 2002; Rosenberg et al., 2004; Salomons et al., 2001, 2003; Schiaffino et al., 2005]. Previously, we have estimated that the prevalence of SLC6A8 deficiency to be 2.1% (95% confidence interval [CI]: 0.44–3.8%) in an European X-linked mental retardation (XLMR) panel [Rosenberg et al., 2004] (European Mental Retardation Consortium; www.euroxmrx.com). Two additional studies reported prevalences of approximately 0.8% and 2% in patients with idiopathic mental retardation [Clark et al., 2006; Newmeyer et al., 2005].

The sequence variants that were identified in the SLC6A8 gene of mentally retarded patients include clear pathogenic DNA variants, such as single amino acid deletions, splice errors, and nonsense mutations, but also missense mutations, which are more difficult to interpret. The latter poses problems for adequate molecular diagnosis. In the first prevalence study [Rosenberg et al., 2004], six presumed pathogenic mutations were detected, and two unclassified missense variants. In the second study [Clark et al., 2006], six potentially pathogenic mutations were detected. Of these, four variants were considered pathogenic based on the absence of the variant in controls (n = 280), conservation of SLC6A8 between species and the human functionally related proteins (SLC6C), and segregation within the family.

In this work we describe an in vitro method for the diagnosis of SLC6A8 deficiency in fibroblasts using gas chromatography mass spectrometry and stable isotope-labeled creatine as internal standard (SID-GCMS). The test is validated and applied to cell

Received 22 September 2006; accepted revised manuscript 6 March 2007.

*Correspondence to: Gajja S. Salomons, Department of Clinical Chemistry, Metabolic Unit, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands.
E-mail: g.salomons@vumc.nl

Grant sponsor: Dutch Society for Scientific Research (ZonMW/NWO); Grant number: VIDI 917.56.349.
DOI: 10.1002/humu.20532
Published online 26 April 2007 in Wiley InterScience (www.interscience.wiley.com).
lines of 13 unaffected individuals and 12 affected patients. Difficulties in distinguishing disease-causing missense mutations from rare variants especially arise when no supplemental material (e.g., fibroblasts, proton magnetic resonance spectroscopy [H-MRS]) is available to substantiate the clinical diagnosis of SLC6A8 deficiency. Therefore, we expanded the creatine uptake assay to study in vitro the pathogenic nature of variants by introducing the mutation in an expression vector with the SLC6A8 coding sequence, followed by transient transfection in SLC6A8-deficient fibroblasts.

MATERIALS AND METHODS

Fibroblasts Cell Culture

A total of 13 fibroblast cell lines in which no inborn errors of metabolism were detected, were anonymized and used as control cell lines. Fibroblast cell lines of 12 patients, affected with creatine transporter deficiency, were reported previously [Mancini et al., 2005; Rosenberg et al., 2004; Salomons et al., 2001, 2003]. Fibroblasts were cultured in HAM-F10 supplemented with 1% penicillin/streptomycin and 10% fetal bovine serum (FBS; Gibco, Invitrogen, Breda, Netherlands) under a humidified atmosphere of 95% air/5% CO2 at 37°C. Culture medium in the absence or presence of 10% FBS was analyzed for Cr content by GCMS (see below) and contained 0 and 25 μM creatine, respectively.

Creatine Uptake Assay

Creatine monohydrate (50 mM; Fluka, Sigma-Aldrich, Zwijndrecht, Netherlands) was dissolved in Hank’s Balanced Salt Solution (HBSS). The stock solution was filter sterilized (0.22 μm filter) and stored at −20°C. The optimal testing conditions were determined by incubation with a concentration range of 25–1,000 μM Cr. After 24 hours, the cells were washed by HBSS twice, harvested by trypsination, followed by two washing steps in HBSS, pelleted (2 minutes; 4,000 g) and stored at −80°C until further use. The final washing solution of the cells was analyzed by GCMS and did not contain creatine.

A total of 12 SLC6A8-deficient patient fibroblast cell lines and 13 control fibroblasts were tested with 25 μM Cr and 500 μM Cr, respectively. Cells were plated into six-well culture plates (Greiner Bio-One, Frickenhausen, Germany) and cultured for at least 24 hours to obtain 50 to 80% confluent cell layers. Cells were incubated for 24 hours with 25 μM or 500 μM creatine. Cells were washed with HBSS twice, harvested by trypsination, followed by two washing steps in HBSS, pelleted (2 minutes; 4,000 g) and stored at −80°C before they were further processed for the quantification of creatine by GCMS, using stable isotope-labeled creatine as internal standard.

Cell pellets were re-suspended in 102 μl HBSS and 60 μl was used for analyzing intracellular creatine concentration; 2.5 nmol of [3H]-creatinine was added as internal standard. Creatine is converted to a pentfluorobenzyl derivate that can be quantitatively measured by the SID-GCMS method as described previously [Almeida et al., 2004; Struys et al., 1998]. Protein content was determined using the Bicinchoninic acid protein assay kit (Sigma-Aldrich), according to the instructions of the manufacturer. The creatine concentration is expressed as pmol creatine/μg total protein.

Interassay variability was determined by processing the same sample in five independent preparations on different days. Intraassay variability was established by processing one sample 10 times within the same experiment.

Mutation Nomenclature

The mutation nomenclature is based on the coding region of the solute carrier Family 6, Member 8 (SLC6A8) transcript, GenBank Accession NM_005629.1 (cDNA), according to recommendations of the Human Genome Variation Society (HGVS; www.hgvs.org/mutnomen). For cDNA numbering, +1 corresponds to the A of the ATG translation initiation codon in the reference sequence. For protein numbering, the initiation codon is codon 1.

Construction of SLC6A8 Expression Vector

RNA was isolated from control fibroblasts (Promega SV RNA isolation kit, Promega, Leiden, Netherlands) and cDNA was synthesized by reverse transcriptase (Qiagen, Venlo, Netherlands). The full-length SLC6A8 open reading frame (ORF) was amplified by PCR using specific primers, with BamHI or HindIII restriction site extensions. Forward primer 5′-CCCAAGCTTCCACCATTGCCAGAG-3′, reverse primer 5′-CCGGAATTCCTGAGCATCAGG-3′. This ampiclon was cloned into pCR2.1-TOPO (Invitrogen, Breda, Netherlands). The construct (pCR2.1-SLC6A8) was used for subcloning into pEGFP-N1 or as template for site-directed mutagenesis

Site-Directed Mutagenesis

Site-directed mutagenesis was introduced to the mutations/variants in pCR2.1-SLC6A8 using specific primers for each mutation/variant. The PCR reaction mix consisted of 50 ng plasmid DNA (1 μl), 100 ng forward and reverse primer (sequences available upon request), 0.8 U Phusion polymerase (Bioké, Leiden, The Netherlands), 10 μg dNTP and 1× PCR buffer in a total volume of 50 μl. PCR reaction was performed with 22 cycles of 98°C for 10 seconds, 56–68°C (depending on the primers) for 30 seconds and 72°C for 90 seconds. After amplification, the template pCR2.1-SLC6A8 was digested by incubation with DpnI (Roche, Almere, Netherlands) at 37°C for 2 hours. Subsequently, the mutant SLC6A8-ORF was excised with BamHI and HindIII, and ligated into a predigested pEGFP-N1 expression vector (Clontech, Mountain View, CA).

After cloning and subcloning, the SLC6A8 cDNA of all the different constructs was completely sequenced to confirm the presence of the desired mutation, and absence of PCR artifacts. Concentration and purity of the constructs was determined with the Nanodrop (Wilmington, DE) N-1000 according to the manufacturers directions. The A260/A280 ratio of all constructs that were used for transfection was between 1.8 and 2.0.

Transfection

The SLC6A8-deficient primary fibroblasts (hemizygous for the p.Arg514X mutation; [Salomons et al., 2001]) were used for transfection. Cells were grown to ~70% confluence. The pEGFP-SLC6A8 constructs were incubated at room temperature for 10 minutes with polyethyleneimine (PEI, Polysciences, Eppelheim, Germany) in a 1:3 ratio in serum-free medium and subsequently applied to the cells. The medium was refreshed after 4 hours. Incubation with creatine was started 48 hours after transfection. Transfections and the creatine uptake assay of each variant was performed in triplicate (500 μM Cr), and measured by SID-GCMS. The creatine concentration results were corrected for the protein content.

Interassay variability was determined by processing five independent transfectants with the same construct on different days. Intraassay variability was established by processing 10 transfectants with the same construct within the same experiment.
RT-PCR

To prove expression of the constructs, RT-PCRs were performed. RNA isolation and cDNA synthesis from the transfected cells was performed as described above. A SLC6A8-specific forward primer (5'-CTGGTGTGACTCAGACGCGC-3') and enhanced green fluorescence protein (EGFP) specific reverse primer (5'-CTGGCGCCCTCACGTGACGAG-3') were used to amplify amplicons of 376 bp by PCR. To rule out genomic and plasmid DNA contamination, RT-PCR was performed without reverse transcriptase for each transfected cell line.

RESULTS

Creatine Uptake Assay

Interassay and intraassay values. Inter- and intraassay variability was determined for all used conditions, i.e., control, SLC6A8-deficient fibroblasts after incubation with 25 μM Cr and 500 μM Cr, and SLC6A8-deficient fibroblasts transfected with mock vector or pEGFP-SLC6A8 after incubation with 500 μM Cr. All results are listed in Table 1.

Reference Values

Cr uptake was measured in control fibroblasts (n = 13) and SLC6A8-deficient fibroblasts (n = 12). The reference values obtained for both groups are listed in Table 1. In all 12 patient fibroblast cell lines with a pathogenic mutation in the SLC6A8 gene, a significant difference in the creatine uptake profile was found compared to the 13 control cell lines. When SLC6A8-deficient fibroblasts were grown in the presence of a physiological Cr concentration (25 μM Cr), the average creatine uptake (0.58 ± 1.03 pmol Cr/μg protein) was just above the detection limit, whereas in the control fibroblasts it was significantly higher (27.8 ± 5.6 pmol Cr/μg protein). When fibroblasts of SLC6A8-deficient patients were incubated with 500 μM up to 1000 μM creatine, some intracellular Cr (6–17 pmol Cr/μg protein) could be detected. In the presence of 500 μM Cr, SLC6A8-deficient fibroblasts showed an average creatine uptake of 6.22 ± 2.14 pmol Cr/μg protein, which is approximately seven times lower than that in control fibroblasts (41.3 ± 7.0 pmol Cr/μg protein).

<table>
<thead>
<tr>
<th>Table 1. Reference Values for the Creatine Uptake Assay in Fibroblasts*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Average (n = 13)</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Intraassay (n = 5)</td>
</tr>
<tr>
<td>Interassay (n = 10)</td>
</tr>
<tr>
<td>Deficient</td>
</tr>
<tr>
<td>Average (n = 12)</td>
</tr>
<tr>
<td>Intraassay (n = 5)</td>
</tr>
<tr>
<td>Interassay (n = 10)</td>
</tr>
<tr>
<td>Transfected wild-type</td>
</tr>
<tr>
<td>Transfected mock</td>
</tr>
</tbody>
</table>

*Control, transfected, or nontransfected SLC6A8-deficient fibroblasts were incubated for 24 hours in presence of 25 μM or 500 μM creatine. Intracellular Cr was measured by SID-GCMS and expressed in pmol Cr/μg total protein.

Transfected Transfection of pEGFP-SLC6A8 Variants

Previously, we reported nine likely pathogenic variants, three unclassified variants and one nonpathogenic variant in the ORF of SLC6A8 (Table 2) [Clark et al., 2006; Rosenberg et al., 2004]. To obtain final proof for the nature of these variants, each of these variants was introduced in the SLC6A8 ORF by site-directed mutagenesis, subcloned in the pEGFP-N1 vector (pEGFP-Lys4Arg, pEGFP-Met560Val, and pEGFP-Val629Ile), expressed in SLC6A8-deficient primary fibroblasts, and tested for Cr uptake capacity when incubated with 500 μM Cr, as described above. RT-PCR confirmed mRNA expression, indicating successful transfection. In contrast to the mock and nontransfected cells, expression of SLC6A8-EGFP mRNA was observed in all cells transfected with wild-type or variant pEGFP-SLC6A8 constructs.

There was no significant difference in creatine uptake between the nontransfected cells (6.2–10.3 pmol Cr/μg protein) vs. mock transfectants (pEGFP-N1; 5.1 pmol Cr/μg protein; Fig. 1) in the same assay. Transient transfection with wild-type pEGFP-SLC6A8 resulted in a six-fold increase of Cr uptake (29.3 pmol Cr/μg protein) (Fig. 1) compared to the mock transfectants (3.1 pmol Cr/μg protein). Overexpression of nine variants (p.Gly87Arg, p.Phe107del, p.Tyr317X, p.Asn336del, p.Cys337Stop, p.Leu347del, p.Pro390Leu, p.Arg391Stop, and p.Pro554Leu) did not result in a significant increase of Cr uptake (range 4.9–8.6 pmol Cr/μg protein; Fig. 1). Overexpression of pEGFP-Lys4Arg, pEGFP-Met560Val, and pEGFP-Val629Ile (23.5–34.3 pmol Cr/μg protein) resulted in a Cr uptake profile similar to the wild-type pEGFP-SLC6A8 (29.3 pmol Cr/μg protein; Fig. 1).

DISCUSSION

The broad spectrum of sequence variants in the SLC6A8 gene in (XL)MR patients poses problems with regard to classification of these changes as pathogenic mutations or nonpathogenic variants/polymorphisms. Deletions in SLC6A8, ranging from single nucleotides to contiguous gene deletions [Corzo et al., 2002] or splice errors, and (other) truncating mutations can easily be classified as pathogenic mutations. However, in order to gain insight in the significance of missense variants, additional workup is needed. This is particularly true for the missense variants that we have identified in the patient panels studied for the prevalence of SLC6A8 mutations [Clark et al., 2006; Rosenberg et al., 2004]. In most cases, it was impossible to prove the clinical diagnosis of SLC6A8 deficiency, since further investigations, such as urinary Cr/Cr(creatine) ratio assessment and/or brain H-MRS imaging could not be performed. We now have developed and validated a functional diagnostic assay, in which we can determine Cr uptake activity in cultured primary fibroblasts, which is reliable and reproducible.

Previously, we have shown that SLC6A8-deficient primary fibroblasts can be complemented by stable transfection with the wild-type ORF of SLC6A8 [Rosenberg et al., 2006]. Since fibroblasts are difficult to transfected stably and are slowly proliferating, we have adapted the protocol for transient transfections. The mutations that were classified as pathogenic mutations or unclassified variants in our previous reports, have been further investigated. As positive controls, four pathogenic mutations (p.Phe107del, p.Tyr317X, p.Asn336del, and p.Leu347del), based on their deleterious character (deletion, nonsense mutation) were transiently transfected. Indeed, overexpression of any of these variants did not restore Cr uptake in SLC6A8-deficient primary fibroblasts (range 7.0–7.9 pmol Cr/μg protein; Fig. 1). Conversely,
transient transfection of the wild-type SLC6A8 ORF restored Cr uptake to normal levels (29.3 pmol Cr/mg protein; Fig. 1). In addition, the p.Met560Val variant, originally identified in a patient from the European XLMR panel [Rosenberg et al., 2004], but more recently found in a healthy control male [Clark et al., 2006], could be used to validate our diagnostic approach, since overexpression of
this variant resulted in similar uptake activity as observed in the wild-type transfectants (31.6 pmol Cr/g protein; Fig. 1).

Overexpression of five missense variants (p.Gly87Arg, p.Cys337Trp, p.Pro390Leu, p.Arg391Trp, and p.Pro354Leu) did not result in a significant increase of Cr uptake (4.9–8.6 pmol Cr/g protein; Fig. 1). Overexpression of the p.Val629Ile allele results in a six-fold increase of Cr uptake (29.8 pmol Cr/g protein), which is similar to the Cr uptake detected in wild-type and p.Met560Val transfectants, confirming that this variant is not pathogenic. Identical conclusions could be drawn for the p.Lys4Arg and p.Gly26Arg alleles (23.5 and 34.1 pmol Cr/g protein, respectively).

The results of the overexpression studies are in good agreement with the previous classification based on theoretical arguments [Rosenberg et al., 2004]. Therefore, for other genes, in which functional analysis is not achievable, the theoretical arguments may prove valuable [LoveLock et al., 2006; Yntema et al., 2002].

We and others have shown previously that the prevalence of SLC6A8 deficiency is relatively high in patient panels with mental retardation of unknown etiology as well as in a panel of XLMR males [Clark et al., 2006; Newmeyer et al., 2005; Rosenberg et al., 2004]. This justifies inclusion of SLC6A8 deficiency in the routine testing in these male patients. The diagnostic approach depends on the available facilities and the cooperation of the patients and their guardians. For initial testing, three tests are available: 1) brain MRS, 2) urine metabolite testing, and 3) DNA analysis (Fig. 2). It should be noted that the uptake assay in fibroblasts is not feasible as primary screening because of its invasive and laborious nature. For accurate diagnosis, a combination of these three tests is preferred, as all have their own respective advantages. 1) In brain H-MRS, a highly decreased Cr signal is a strong indication for a primary cerebral creatine deficiency syndrome. In the majority of the cases, this is caused by SLC6A8 deficiency. However, this may also be the result of a Cr biosynthesis defect. Thus, metabolite and/or functional assays and/or DNA analysis is warranted for a definitive diagnosis. The advantage of brain MRI/MRS is that other disorders can be detected simultaneously [Newmeyer et al., 2005]. However, MRS is only available in specialized institutes and usually sedation is needed for patients with mental retardation. 2) An elevated urinary Cr:Crn ratio is a strong indication for SLC6A8 deficiency [Almeida et al., 2004]. In that case, DNA analysis is warranted to confirm the diagnosis.

![FIGURE 2. Schematic diagram for screening for SLC6A8 deficiency in patients with mental retardation of unknown etiology. Brain H-MRS: A highly decreased Cr signal in brain H-MRS is a strong indication for primary cerebral creatine deficiency syndrome. Metabolite and/or DNA analysis is warranted for a definitive diagnosis. Urine Cr:Crn: An elevated urinary Cr:Crn ratio is a strong indication for SLC6A8 deficiency [Almeida et al., 2004]. In that case, DNA analysis is warranted to confirm the diagnosis. DNA sequence analysis: In many cases the result of DNA analysis can provide the definitive diagnosis of SLC6A8 deficiency. The DNA analysis can result in three variant options: a pathogenic mutation is found, a novel missense variant is found, or a novel neutral variant or intronic variant is detected. *Brain MRI/H-MRS may result in detection of an other disorder. †Usually, in females the Cr:Crn ratio is not informative. ‡In approximately 10% a false-positive Cr:Crn ratio is found (see Discussion). §There is no information on the frequency of false-negative results in metabolite and DNA analysis. Mutations outside the analyzed regions will be missed. ¶A mutation is considered to be pathogenic if it: 1) is a nonsense mutation, 2) is a deletion, 3) causes a frameshift, or 4) has been proven to be pathogenic. ¶¶In case the Cr signal is reduced on brain H-MRS, AGAT and GAMT-deficiency need to be investigated. §§If brain H-MRS and/or uptake in fibroblasts indicate SLC6A8 deficiency, but site-directed mutagenesis classifies the missense variant as a polymorphism, additional molecular analysis should be performed to identify the pathogenic mutation.](Image)
Furthermore, urine metabolite analysis may also reveal a Cr biosynthesis defect especially if guanidoacetic acid, the precursor of creatine, is included [Almeida et al., 2004]. Our experience is that in about 10% of the patients with an elevated urinary Cr/Crnm ratio, this normalizes in repeat samples. In those patients we did not detect a DNA variant. For the case in which the urinary Cr/Crnm ratio is consistently increased in the absence of a mutation, follow-up investigations (e.g., uptake in cultured fibroblasts or H-MRS of the brain) are needed to exclude that a mutation has been missed. 3) DNA is available from most patients for routine testing of common causes of mental retardation. Therefore, genomic testing for SLC6A8 deficiency will not put additional burden on the patient. Furthermore, in many cases the result of DNA analysis can provide the definitive diagnosis of SLC6A8 deficiency.

The DNA analysis can result in three findings (Fig. 2): 1) no DNA alteration, 2) a pathogenic mutation, or 3) a novel DNA variant (missense, neutral, or intronic).

1. If no DNA alteration is found, there is no indication for creatine transporter defect, although formally mutations outside the analyzed regions (i.e., promoter or deep intronic mutations) will be missed. In case the Cr signal is reduced on brain H-MRS, l-arginine: glycine amidinotransferase (AGAT) and guanidinoacetate methyltransferase (GAMT) deficiency need to be investigated.

2. A mutation is considered pathogenic if it is a nonsense mutation, a deletion, a frameshift mutation, a splice error, or previously published and well-described as pathogenic. In such a case, SLC6A8-deficiency is sufficiently proven and further workup is not necessary.

3. A novel DNA variant may be found that was not reported before, nor was it found in male control DNA samples. Then, Cr uptake assay in patient fibroblasts should be performed. A normal Cr uptake profile in fibroblasts excludes SLC6A8 deficiency, and the patient should be tested for GAMT and AGAT deficiency in case the Cr signal in brain H-MRS was reduced. If, however, a Cr uptake assay could not be performed, or the fibroblasts prove to be deficient in Cr uptake, further studies are needed to prove that the variant is pathogenic, especially in case of prenatal diagnosis. Hereeto, overexpression of the missense variant, as described in the present study, should be performed. Although in our experiments 500 μM Cr concentrations were used, in future experiments 25 μM Cr concentrations would be preferable, since this represents the physiological condition. If overexpression of the variant does not restore Cr uptake in the SLC6A8-deficient host cells, it is a pathogenic mutation. If, however, Cr uptake is restored, then the variant represents a nonpathogenic variant and molecular analysis should be refined by SLC6A8 mRNA analysis. If a novel intronic (IVS) or neutral variant is detected that was not reported before, nor was it found in male control DNA samples, SLC6A8 mRNA analysis is needed. In the case that alternative splice variants are detected, SLC6A8 deficiency is confirmed.

The guidance of the flowchart, facilitates in most cases a conclusive diagnosis (or exclusion) of either SLC6A8 deficiency or a Cr biosynthesis defect.

In summary, the Cr uptake assay in primary fibroblasts is a valuable tool for proper diagnosis (or ruling out) of SLC6A8 deficiency, whereas overexpression of SLC6A8 alleles containing missense variants, to be investigated in SLC6A8-deficient cells, is a valid tool for unequivocal classification of these variants. By applying these studies, we have shown that nine alleles proved not to result in Cr uptake activity in SLC6A8-deficient fibroblasts upon transient transfection, and thus could be classified as pathogenic mutations, and four alleles were proven to be rare variants that have Cr uptake activity similar to the wild-type SLC6A8. In case prenatal diagnosis is required, it is essential to prove pathogenicity of the missense variant by overexpression studies (e.g., transient transfection into SLC6A8-deficient cells).

ACKNOWLEDGMENTS

We thank Lígia Almeida, Patricia Darmin, Ulbe Holwerda, and Lorna Landegge Pope for their profound technical support and intellectual contribution. The work of E.H. Rosenberg, C. Martínez-Muñoz, O.T. Betsalel, and G.S. Salomons was supported by the Dutch Society for Scientific Research (ZonMW/NWO), VIDI grant number 917.56.349.

REFERENCES


Rosenberg EH, Almeida LS, Kleefstra T, deGrauw RS, Yntema HG, Bath N, Moraine C, Rogers HH, Fryns JP, deGrauw TJ, Jakobs C, Salomons GS.
Chapter 2.2

Detection of variants in SLC6A8 and functional analysis of unclassified missense variants.

Molecular Genetics and Metabolism, 2012
Detection of variants in SLC6A8 and functional analysis of unclassified missense variants

Ofir T. Betsalel a, Ana Pop a, Efraim H. Rosenberg b, Matilde Fernandez-Ojeda a, Creatine Transporter Research, Group 1, Cornelis Jakobs a, Gajja S. Salomons a,b,*

a Metabolic Unit, Department of Clinical Chemistry, VU University Medical Center, Amsterdam, The Netherlands
b Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands

ABSTRACT

Creatine transporter deficiency is an X-linked disorder caused by mutations in the SLC6A8 gene. Currently, 38 pathogenic, including 15 missense variants, are reported. In this study, we report 33 novel, including 6 missense variants. To classify all known missense variants, we transfected creatine deficient fibroblasts with the SLC6A8 ORF containing one of the unique variants and tested their ability to restore creatine uptake. This resulted in the definitive classification of 2 non-disease associated and 19 pathogenic variants of which 3 have residual activity. Furthermore, we report the development and validation of a novel DHPLC method for the detection of heterozygous SLC6A8 variants. The method was validated by analysis of DNAs that in total contained 67 unique variants of which 66 could be detected. Therefore, this rapid screening method may prove valuable for the analysis of large cohorts of females with mild intellectual disability of unknown etiology, since in this group heterozygous SLC6A8 mutations may be detected. DIHPLC proved also to be important for the detection of somatic mosaicism in mothers of patients who have a pathogenic mutation in SLC6A8. All variants reported in the present and previous studies are included in the Leiden Open Source Variant Database (LOVD) of SLC6A8 (www.LOVD.nl/SLC6A8).

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

Cerebral creatine deficiency syndromes (CCDS) are caused by arginine:glycine amidinotransferase (AGAT) deficiency, guanidinoacetate methyltransferase (GAMT) deficiency or creatine transporter (SLC6A8) deficiency. All three deficiencies are characterized by cerebral creatine deficiency, intellectual disability (ID), language delay, epilepsy and autistic-like behavior. GAMT deficiency, which affects the second enzyme in creatine biosynthesis, was the first inborn error in the creatine metabolism to be identified [18]. AGAT deficiency (MIM 612718), the second CCDS reported, was found in two female siblings with intellectual disability [12]. The discovery of creatine transporter deficiency being caused by pathogenic mutations in the SLC6A8 gene appeared to complete the search of the genetic cause of primary inborn errors of creatine metabolism [16].

In all forms of CCDS, the cerebral creatine levels are reduced which can be detected by brain proton magnetic resonance spectroscopy (MRS). Urinary creatine:creatinine ratio is increased in males with SLC6A8 deficiency and creatine uptake in cultured fibroblasts is impaired. In the vast majority of males affected with SLC6A8 deficiency, direct genomic DNA sequencing analysis reveals a pathogenic mutation in SLC6A8. The definitive diagnosis is often based on at least two of the above-mentioned analyses. Since the creatine transporter defect is an X-linked disorder, in males a hemizygous mutation is expected, whereas in females heterozygous mutations are expected. The identification of females is complex since 1) they have been shown to have varying phenotypes, 2) they often have normal urinary creatine:creatinine ratios, 3) as a group they have lower than average cerebral creatine levels, but individual values are within the normal range, 4) functional studies in fibroblasts are hampered by the fact that skewed X-inactivation may result in expression of only the wildtype allele and thus both mRNA studies as well as creatine uptake studies can remain inconclusive, 5) the detection of heterozygous mutations in the SLC6A8 gene is slightly more difficult than with learning disabilities from SLC6A8 deficiency [19,20]. In both males and females, when a novel variant without a clear pathogenicity (i.e. missense, intronic) is detected, the variant needs to be classified. Previously, we reported methods for characterizing novel missense variants by site-directed mutagenesis, followed by transfection in cultured fibroblasts [15] and for the characterization of novel intronic and synonymous variants by in silico analysis and/or overexpression and mRNA analysis of minigenes [3,15]. In the present study, we include the functional characterization of 21 missense variants, of
which 15 novel, as well as the validation of the detection of variants in SLC6A8 by denaturing high performance liquid chromatography (DHPLC), most suitable for detecting heterozygous variants in females. All variants reported in this study are included in the LOVD database (www.lovd.nl/SLC6A8).

2. Materials and methods

2.1. Samples

For the DHPLC analysis of SLC6A8, 33 novel variants and 34 previously published variants, all found and/or confirmed in our diagnostic department, were used. The samples analyzed consisted of 6 intronic and 61 exonic variants. For the functional analysis of the SLC6A8 missense variants, using the overexpression assay, we analyzed a total of 21 variants.

2.2. PCR

The amplification and sequencing of SLC6A8 were performed under standard conditions as described earlier. For DHPLC analysis, all amplicons, consisting of a single exon and the flanking intronic sequences, were amplified using Discoverase DHPLC DNA polymerase (Invitrogen) according to the instructions of the manufacturer. Primers were designed using Navigator 2.0 software (Transgenomic, Omaha, NE) according to the instructions of the manufacturer. Primers were designed using Navigator 2.0 software (Transgenomic, Omaha, NE) according to the instructions of the manufacturer.

2.3. Mutation detection by DHPLC

PCR products of a hemizygous (SLC6A8) mutation were mixed in a 1:1 volume ratio with a wildtype PCR product of the same amplicon, while heterozygous variants were used directly. Samples were heat-denatured for 5 min at 95 °C and gradually cooled to 25 °C in a period of 1 h.

Subsequently, 5 μl of the heteroduplex/homoduplex mixture was loaded on a WAVE 3500HT DNA fragment analysis system (Transgenicomic, Omaha, NE). Elution of the PS-DVB DNAsep column (Transgenomic, Omaha, NE) was performed in high-throughput mode with a linear gradient of increasing acetonitrile (ACN) concentration with a flow rate of 1.5 ml/min. Gradient was realized with the use of buffer A (0.1 M triethylammonium acetate [TEAA] and 0.025% ACN) and buffer B (0.1 M TEAA and 25% ACN). Navigator 2.0 software (Transgenicomic, Omaha, NE) was used to calculate the ratio of buffers A and B during the gradient, as well as the optimal partially denaturing temperature(s) (Tm) for each amplicon, depicted in Table 1.

2.4. Functional analysis of missense variants by overexpression studies in SLC6A8 deficient fibroblasts

To assess the effect of the missense variants found in SLC6A8, they were all introduced in the SLC6A8 open reading frame by site-directed mutagenesis, using Phusion polymerase (NEB). The reaction was performed in a total volume of 50 μl, including 50 ng template DNA, 3% DMSO and 125 ng of forward and reverse primer (For primer sequences see Supplementary Table 1). The thermocycling program consisted of an initial 2′ at 98 °C, followed by 22 cycles of 15′ at 98 °C, 60′ at 66 °C and 90′ at 72 °C, with a final step of 10′ at 72 °C. The product was then subcloned in the pEGFP-N1 vector. Prior to transfection, polyethyleneimine and plasmid were mixed in a 1:3 ratio. Subsequently, SLC6A8-deficient primary fibroblasts were transiently transfected in triplicate with the variant containing vector, as well as a separate transfection of the wildtype and the empty vector. The creatine uptake capacity was then tested through incubation with the physiological creatine concentration of 25 μM as previously described by Ref. [15]. Cells were harvested by trypsinization, divided in aliquots for creatine measurement as well as Western blotting and stored as dry pellets at —80 °C until further use. Intracellular creatine content was measured by stable isotope dilution GC–MS [14]. The presence of the creatine transporter was determined with the use of Western blotting. To prove SLC6A8-EF2 fusion protein expression, cell pellets were lysed in urea lysis buffer (8 M urea/100 mM NaCl/10 mM Tris–HCl, pH 8.0). Protein content was measured using a Bicinchoninic acid protein assay (Sigma-Aldrich, St Louis, MO, USA).

Cell lysates (10 μg protein) were size-separated by 12% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS PAGE) and transferred to a PVDF membrane (Invitrogen, dry blot system). Immunodetection was performed using antibodies directed against the EGFP tag. Immune complexes were detected by enhanced chemiluminescence (LumiLight Plus), conducted according to the manufacturer’s specifications (Roche, Indianapolis, IN, USA).

3. LOVD database

All novel SLC6A8 variants have been checked for proper nomenclature using Mutalyzer and were subsequently included in the LOVD database. The database is localized at a server in Leiden, The Netherlands, which can be reached by the following URL http://www.LOVD.nl/SLC6A8 or through the Variation Databases page of the Human Genome Variation Society (HGVS, www.HGVS).

Table 1

<table>
<thead>
<tr>
<th>Exon</th>
<th>Basepairs</th>
<th>Melting temperature(s)</th>
<th>Primer sequence (5′–3′)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>346</td>
<td>62.2</td>
<td>cctacactgactacccagtc</td>
</tr>
<tr>
<td>3</td>
<td>416</td>
<td>61.8; 63.2; 63.8</td>
<td>gaggttggcgcaggaagtc</td>
</tr>
<tr>
<td>4</td>
<td>328</td>
<td>63.2</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>5</td>
<td>273</td>
<td>63.9</td>
<td>gagagtgcgcagaggtct</td>
</tr>
<tr>
<td>6</td>
<td>288</td>
<td>63.3; 64.2</td>
<td>gtcatacaagacccgcgcg</td>
</tr>
<tr>
<td>7</td>
<td>311</td>
<td>63.9</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>8</td>
<td>295</td>
<td>65.4</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>9</td>
<td>320</td>
<td>61.6</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>10</td>
<td>339</td>
<td>64.6; 65.1</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>11</td>
<td>228</td>
<td>64.6; 65.1</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>12</td>
<td>437</td>
<td>63.1; 63.8; 64.9</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
<tr>
<td>13</td>
<td>344</td>
<td>64.1</td>
<td>cggccgccaagcgcgcgctctc</td>
</tr>
</tbody>
</table>

Exons 2 to 13 of SLC6A8 were amplified in 12 separate amplicons with the corresponding primer pairs. Primers were designed using the Navigator 2.0 software (Transgenicomic, Omaha, NE) and GC clamp (underlined in primer sequence) were added if necessary. Each amplicon was run with the corresponding melting temperature. In cases where the sequence could not be analyzed with one melting temperature, multiple injections were performed with different melting temperatures. When analyzing male patients, PCR reactions were mixed in a 1:1 volumetric ratio with the corresponding wildtype amplicon to ensure detection of hemizygous variants.
org). The database has been developed with the recently described LOVD software, which is in agreement with the HGVS guidelines. Currently, the SLC6A8 database contains 80 pathogenic mutations reported in 138 male patients and 82 female carriers, and 74 variants that seem not to be associated with SLC6A8 deficiency.

4. Results

4.1. Functional and conservative analysis of novel missense variants

The transfection of the 21 missense variants resulted in definite characterization of the variants. In total we found 16 variants to be pathogenic, indicated by the lack of restoration of creatine uptake compared to the wildtype transfected fibroblasts. We found 3 variants, c.1271G>A; p.(Gly424Asp), c.1631C>T; p.(Pro544Leu) and c.1699T>C; p.(Ser567Pro), to have residual activity ranging from 21% to 39% of the measured activity of the wildtype transfected cells. Two variants, c.1442C>T; p.(Thr481Ile) and c.1516G>A; p.(Asp506Asn), were characterized as not associated with disease since the transfection with these vectors did result in a restoration of the creatine uptake. All triplicate transfections for each variant resulted in detectable EGFP-SLC6A8 fusion protein on Western blot. Some variants resulted in different patterns as compared to the wildtype and although the blots were not quantified, the level of expression of variants without restored SLC6A8 activity showed a generally reduced expression (Fig. 1).

All missense variants studied by overexpression in SLC6A8 deficient fibroblasts, consisting of 20 pathogenic, 6 non-disease associated and 3 with residual activity, are depicted in the model of SLC6A8, based on the LeuT model of Ref. [21], together with the computed conservation of the 12 transmembrane domains of SLC6A8 throughout different species as well as the SLC6 family (Fig. 2).

4.2. Detection of variants by DHPLC

For the DHPLC analysis, all but one of the 67 mutations and variants of SLC6A8 analyzed were identified with a clear separation of hetero- and homoduplexes (Supplementary Fig. F1) and all variants resulted in distinctive unique peak patterns. The c.541T>C variant not being separated at a range of melting temperatures was already expected at onset and indeed the analysis at partial denaturing temperatures, ranging from 61.7 °C to 64.1 °C, did not result in clear separation.

4.3. Detection of low level somatic mosaicism by DHPLC

One mutation, c.1667G>A;Trp556X, was detected by DNA sequence analysis in an index SLC6A8 deficient male. Subsequent sequencing of his mother did not give a clear result of the presence of the variant. However, using DHPLC, we showed a low level of the variant present in DNA extracted from EDTA blood, indicating low-level somatic mosaicism (Fig. 3).

5. Discussion

Currently, the methods of definitive diagnosis for CCDS include a wide range of biochemical and molecular analyses. Previously, we set up guidelines for the diagnosis of CCDS in males, starting from patients suffering from intellectual disability through to the definitive molecular analysis [15]. In the present paper we have investigated...
the effect of 21 missense variants on the function of the creatine transporter. Of these, 16 showed reduced creatine uptake comparable with untransfected SLC6A8 deficient fibroblasts. The presence of the EGFP–SLC6A8 fusion proteins was shown by Western blotting, indicating proper transfection. These data classified all 16 variants as pathogenic.

Included in the pathogenic variants is c.1141G>C, p.[Gly381Arg, Val377GlyfsX15]. In a previous study, cDNA analysis, which showed two splice products, together with impaired creatine uptake in the patients fibroblasts proved the pathogenic nature of this variant [3,9]. In this study we detected the c.1141G>A variant, which causes the same amino acid substitution, but due to the lack of additional materials of this patient we could not show the actual alternative splicing effect of the variant. However, in the functional analysis of the c.1141G>C variant, where we only tested the effect of the amino acid substitution, a pathogenic effect was seen. Therefore, regardless of any alternative splicing effect, we can conclude that the c.1141G>A variant is also a pathogenic event.

Of the other 5 missense variants, 2 variants, c.1442C>T; p.(Thr481Ile) and c.1516G>A; p.[Asp506Asn], resulted in creatine uptake in the same range of the wildtype transfected cells, indicating that these variants do not cause SLC6A8 deficiency. The remaining 3 variants, c.1271G>A; p.(Gly424Asp), c.1631C>T; p.(Pro544Leu) and c.1699T>C; p.(Ser567Pro), resulted in a residual uptake of 29%, 39% and 37% of the wildtype transfection level respectively and are depicted in a blue hexagon. Two variants, c.1472G>A and c.1473C>G, substitute the 491st amino acid, Cysteine, to a Tryptophan and Tyrosine respectively and are only indicated by a single red box. The overall amino acid conservation of SLC6A8 throughout different species (Rat, Mouse, Cow, Rabbit, Marbled Electric Ray) and the SLG6 neurotransmitter transporter family (A1 = GABA1, A2 = noradrenaline, A3 = dopamine, A4 = serotonine, A5 = glycine, A6 = taurine, A7 = proline, A9 = glycine, A11 = GABA3, A12 = betaine/GABA, A13 = GABA2 and A14 = ATB0+) is computed for each transmembrane domain (TM). The number of identical, functionally conserved and non-conserved amino acids was divided by the total amount of amino acids of each TM and results are depicted in black (identical), dark grey (functionally conserved) and light grey (non-conserved). Please note that while some intra- and extracellular domains are illustrated in the same manner as the TM domains, they are not included in the amino acid conservation calculations. Also, the size of the illustrated TMs does not correlate with the actual number of amino acids.

Several studies have been performed with creatine and creatine precursors supplementation to patients with SLC6A8 deficiency, but unfortunately no consistent improvement in clinical outcome was reported [1,2,4,6,7,14]. However, patients in whom mutations are found with residual activity may have a better response to creatine and/or creatine precursors supplementation. So far 2 families including 3 patients with the c.1631C>T; p.Pro544Leu mutation have been treated with either both l-arginine and creatine or just with l-arginine. In both families, no significant clinical improvement was observed [8,19,20]. There is no treatment data available for the patients harboring the other 2 variants with residual activity found in this study, c.1271G>A; p.(Gly424Asp) and c.1699T>C; p.(Ser567Pro).

All of the currently known variants, with the exception of large deletions and deep intronic variants, were used for the validation of DHPLC. This resulted in a highly sensitive 99% (66 out of 67) detection rate. The c.541T>C variant was not detected. Prior to measurement
this was expected since no partial denaturing temperature could be predicted, as a result of the high GC-content of this region. This problem could not be circumvented with the use of lengthy GC-clamps and this was also the case for exon 1 of \textit{SLC6A8}. The conclusion is that these regions should be analyzed by direct sequence analysis. DHPLC is most suitable for the detection of heterozygous mutations and in the case of low-level mosaicism DHPLC can easily detect a 1\% occurrence of the mutant allele, while sequencing may not be able to reach this. Henceforth we propose using DHPLC for the screening of females with (mild) intellectual disability. Furthermore, due to the small amount of handling time required, it is a very convenient and suitable technique for mutation analysis of large cohorts. However, the detection of homozygous or hemizygous mutations does require an additional step of adding an equal amount of WT PCR product, increasing the handling time for each sample. In general, DNA analysis for heterozygous variants has its caveats; large deletions can be missed, allele dropout can cause a wildtype signal and in the specific case of \textit{SLC6A8}, where only the open reading frame is sequenced, variants outside this region can be overseen.

All our findings are reported in the LOVD database, with which we hope to aid in and simplify the process of diagnosing \textit{SLC6A8} deficiency. Moreover, the database is a great tool for clinicians and researchers to get an up-to-date overview of the type of mutations, the frequency of specific variants and more details. An added value is the option to hide information on mutations that have not been made public yet, but still allows contact between clinicians. Since all data is carefully screened, researchers and clinicians are invited to submit their data to the database building up a vast collection of accessible information.

We report a total of 67 variants, of which 33 novel, and their effect on the activity of the creatine transporter, shown by transient transfection in \textit{broblasts}. In addition, we developed a fast and efficient method for the DNA analysis of \textit{SLC6A8}, suitable for screening of large cohorts of males with intellectual disability of unknown etiology. This method is very useful for the detection of (somatic) mosaicism in females and in the screening for mutations of potential female carriers since it does not require a mixing step with a wildtype DNA.

Supplementary materials related to this article can be found online at doi:10.1016/j.ymgme.2011.12.022.

\textbf{Conflict of interest}

All authors have nothing to disclose.

\textbf{Acknowledgments}

The authors would like to thank their fellow investigators of the Creatine Transporter Research Group: \textit{Belgium:} H. van Esch, M.C. Nassogne, Y. Sznajer, Université Catholique de Louvain, Leuven; \textit{Canada:} S. Mercimek-Mahmutoglu, University of British Colombia, Vancouver; \textit{Estonia:} K. Ounap, Tartu University Hospital, Tartu; \textit{France:} D. Cheillan, L. Lion-Francois, G. Pitelet, Hôpital Debrousse, Lyon; A. Munnich, Institut National de la Santé et de la Recherche Médicale, Montpellier; S. Quijano-Roy, Hôpital Raymond-Poincaré, Paris; V. Valayannopoulos, Necker-Enfants Malades Hospital, Paris; \textit{Germany:} S. Burkhard, Kinderkrankenhaus der Stadt Köln, Cologne; B.R. Gebhardt, Johann Wolfgang Goethe University of Frankfurt, Frankfurt; U. Moog, Universitätsklinikum Heidelberg,
Heidelberg; J. Muecke, Kliniken St. Ingbert des Saarpfalz-Kreises, Saarbrücken; P. Wolf, DRK-Kinderklinik Siegen, Siegen; Italy: U. Caruso, C. Schiaffino, Instituto Giannina Gaslini, Genova; Latvia: Z. Krumina, Children’s Hospital Gaigezers, Riga; Portugal: M. Rodrigues do Rosario, C. Valongo, L. Vilarinho, Centro de Genética Médica Jacinto de Magalhães, Lisbon; Spain: J. Campistol, M.A.Vilaseca, Hospital Sant Joan de Deu Barcelona, Barcelona; A. Ribes, University of Barcelona, Barcelona; The Netherlands: A. de Brouwer, R. Wevers, H. Yntema, Radboud University Nijmegen Medical Centre, Nijmegen; S.G.M. Frints, Maastricht University Medical Centre, Maastricht; P.M. van H Jesselt, P.F. Ippel, University Medical Center Utrecht, Utrecht; S.G. Kant, Leids University Medical Center, Leiden; S.A. de Man, Amphia Hospital, Breda; G. Mancini, J.A. Maat-Kievit, Erasmus Medical Center, Rotterdam; K.E. Niezen-Koning, University Medical Center Groningen, Groningen; J. van de Kamp, Free University Medical Center, Amsterdam; Tunisia: H. Azzouz, La Rabta Hospital, Tunis; UK: J.A. Cook, D. Johnson, Sheffield Children’s Hospital, Sheffield; USA: T. Degrauw, Cincinnati Children’s Hospital Medical Center, Cincinnati; OH; N. Longo, University of Utah Health Care, Salt Lake City, UT; C. Schwartz, Greenwood Genetic Center, Greenwood, SC; A. Vanderver, Cincinnati Children’s Hospital Medical Center, Cincinnati; OH; N. Longo, University of Utah Health Care, Salt Lake City, UT; C. Schwartz, Greenwood Genetic Center, Greenwood, SC; A. Vanderver, Children’s National Medical Center, Washington D.C.

The work of Ofer T Betsalel and Gaja S Salomons is supported by the Dutch Society for Scientific Research (ZonMW/NWO), VIDI grant number 917.56.349.

References

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>c.291_292insAGGG; p.(Ala98ArgfsX92)</td>
<td>c.306A&gt;G; p.(=)</td>
<td>c.321_323delCTT; p.(Phe107del)</td>
</tr>
<tr>
<td></td>
<td>c.370T&gt;C; p.(Trp124Arg)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 3</td>
<td>Wild type 63.6</td>
<td>c.395G&gt;T; p.(Gly132Val)</td>
<td>c.428_430delACT; p.(Tyr143del)</td>
</tr>
<tr>
<td></td>
<td>c.467_469del; p.(Phe156del)</td>
<td>c.495C&gt;T; p.(Thr165Thr)</td>
<td>c.541T&gt;C; p.(Cys181Arg)</td>
</tr>
<tr>
<td></td>
<td>c.570_571delTG; p.(Ala191GlnfsX10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 4</td>
<td>Wild type 63.2</td>
<td>c.676G&gt;T; p.(Glu226X)</td>
<td>c.729G&gt;T; p.(Trp243Cys)</td>
</tr>
<tr>
<td>Exon 5</td>
<td>Wild type 63.9</td>
<td>c.778-2A&gt;G; p.(spl)</td>
<td>c.786 C&gt;G; p.(Tyr262X)</td>
</tr>
<tr>
<td>Exon 6</td>
<td>Exon 7</td>
<td>Exon 7</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>----------------</td>
<td>----------------</td>
<td></td>
</tr>
<tr>
<td>c.813C&gt;T; p.Val217Val</td>
<td>c.878_879delTC; p.(Lys293fsX3)</td>
<td>c.884_885delCT; p.(Pro295ArgfsX169)</td>
<td></td>
</tr>
<tr>
<td>c.912C&gt;A; p.Ile260_Gln304del</td>
<td>Wild type 63.3</td>
<td>c.917G&gt;A; p.(Trp306X)</td>
<td></td>
</tr>
<tr>
<td>c.942_944delCTT; p.(Phe315del)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.950dupA; p.(Tyr317X)</td>
<td>c.974_975delCA; p.(Thr325SerfsX139)</td>
<td>c.986G&gt;T; p.(Ser329Ile)</td>
<td></td>
</tr>
<tr>
<td>c.1006_1008delAAC; p.(Asn336del)</td>
<td>c.1011C&gt;G; p.(Cys337Trp)</td>
<td>c.1059_1061delCTT; p.(Phe354del)</td>
<td></td>
</tr>
<tr>
<td>c.1067G&gt;T; p.(Gly356Val)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1090G&gt;T; p.(Gly364Cys)</td>
<td>c.1105G&gt;A; p.(Glu369Lys)</td>
<td>c.1141G&gt;C; p.(Gly381Arg, Val377GlyfsX15)</td>
<td></td>
</tr>
<tr>
<td>c.1141G&gt;A; p.(Gly381Arg)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 8</td>
<td>Wild type 62.4</td>
<td>c.1145C&gt;T; p.(Pro382Leu)</td>
<td>c.1162G&gt;A; p.(Ala388Thr)</td>
</tr>
<tr>
<td>----------------</td>
<td>---------------</td>
<td>--------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td></td>
<td>c.1166A&gt;G; p.(Tyr389Cys)</td>
<td>c.1169C&gt;T; p.(Pro390Leu)</td>
<td>c.1206G&gt;A; p.(Trp402X)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.1222_1224delTTC; p.(Phe408del)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 9</td>
<td>Wild type 63.1</td>
<td>c.1261G&gt;C; p.(Gly421Arg)</td>
<td>c.1271G&gt;A; p.(Gly424Asp)</td>
</tr>
<tr>
<td></td>
<td>c.1299_1309delCCCCGGCTCCT; p.(Pro434LeufsX27)</td>
<td>c.1318dupC; p.(Arg440ProfsX25)</td>
<td>c.1381A&gt;G; p.(Met461Val)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>c.1372_1375delGATC; p.(Asp458SerfsX4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 10</td>
<td>Wild type 64.6</td>
<td>c.1417_1420delTTTG; p.(Phe473ThrfsX37)</td>
<td>c.1428C&gt;G; p.Tyr476X</td>
</tr>
<tr>
<td></td>
<td>c.1432dupG; p.(Ala478GlyfsX24)</td>
<td>c.1456C&gt;T; p.Gln468X</td>
<td>c.1472G&gt;A; p.(Cys491Tyr)</td>
</tr>
<tr>
<td>c.1473C&gt;G; p.(Cys491Trp)</td>
<td>c.1495+5G&gt;C; p.Gly465GlufsX12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type 64.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1496-7G&gt;A; p.(=)</td>
<td>c.1496-8C&gt;T; p.(=)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1516G&gt;A; p.(Asp506Asn)</td>
<td>c.1519_1543delATTGCCTGTATGATCGGGTACCGAC; p.(Ile507LeufsX5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1540C&gt;T; p.(Arg514X)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1554G&gt;A; p.(Trp518X)</td>
<td>c.1596+1G&gt;A; p.(spl)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 12</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type 64.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1607_1612delTCTTCAI; p.(Ile536_Phe537del)</td>
<td>c.1631C&gt;T; p.(Pro544Leu)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1661C&gt;T; p.Pro554Leu</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1667G&gt;A; p.(Trp556X)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1678A&gt;G; p.(Met560Val)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type 64.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1685G&gt;A; p.(Trp562X)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>c.1885G&gt;A; p.(Val629Ile)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exon 11</td>
<td>Exon 12</td>
<td>Exon 13</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>c.1473C&gt;G; p.(Cys491Trp)</td>
<td>c.1495+5G&gt;C; p.Gly465GlufsX12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type 64.2</td>
<td>c.1496-7G&gt;A; p.(=)</td>
<td>c.1496-8C&gt;T; p.(=)</td>
<td></td>
</tr>
<tr>
<td>c.1516G&gt;A; p.(Asp506Asn)</td>
<td>c.1519_1543delATTGCCTGTAT GATCGGGTACCGAC; p.(Ile507LeufsX5)</td>
<td>c.1540C&gt;T; p.(Arg514X)</td>
<td></td>
</tr>
<tr>
<td>c.1554G&gt;A; p.(Trp518X)</td>
<td>c.1596+1G&gt;A; p.(spI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type 64.9</td>
<td>c.1607_1612delTCTTCAI; p.(Ile536_Phe537del)</td>
<td>c.1631C&gt;T; p.(Pro544Leu)</td>
<td></td>
</tr>
<tr>
<td>c.1661C&gt;T; p.Pro554Leu</td>
<td>c.1667G&gt;A; p.(Trp556X)</td>
<td>c.1678A&gt;G; p.(Met560Val)</td>
<td></td>
</tr>
<tr>
<td>Wild type 64.1</td>
<td>c.1685G&gt;A; p.(Trp562X)</td>
<td>c.1885G&gt;A; p.(Val629Ile)</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2.3

Characterization of novel SLC6A8 variants with the use of splice-site analysis tools and implementation of a newly developed LOVD database.

*European Journal of Human Genetics, 2011*
Characterization of novel SLC6A8 variants with the use of splice-site analysis tools and implementation of a newly developed LOVD database

Offr T Betsalel, Efraim H Rosenberg, Ligia S Almeida, Tjitske Kleefstra, Charles E Schwartz, Vassili Valayannopoulos, Omar Abdul-Rahman, Nicola Poplawski, Laura Vilarinho, Philipp Wolf, Johan T den Dunnen, Cornelis Jakobs and Gajja S Salomons

The X-linked creatine transporter defect is caused by mutations in the SLC6A8 gene. Until now, 66 synonymous and intronic variants in SLC6A8 were detected in our laboratory. To gain more insight in the effect of the detected variants, we applied five free web-based splice-site analysis tools to 25 published variants that were stratified as (non-)disease causing. All were correctly predicted to have no effect (n=18) or to cause erroneous splicing (n=7), with the exception of a pathogenic de novo 24-bp intronic deletion. Second, 41 unclassified variants, including 28 novel, were subjected to analysis by these tools. At least four splice-site analysis tools predicted that three of the variants would affect splicing as the mutations disrupted the canonical splice site. Urinary creatine/creatinine ratio and impaired creatine transporter deficiency in five patients (four families), including one female. Another variant was predicted to moderately affect splicing by all five tools. However, transient transfection of a minigene containing the variant in a partial SLC6A8 segment showed no splicing errors, and thus was finally classified as non-disease causing. This study shows that splice tools are useful for the characterization of the majority of variants, but also illustrates that the actual effect can be misclassified in rare occasions. Therefore, further laboratory studies should be considered before final conclusions on the disease-causing nature are drawn. To provide an accessible database, the 109 currently known SLC6A8 variants, including 35 novel ones, are included in a newly developed LOVD DNA variation database.

European Journal of Human Genetics (2011) 19, 56–63; doi:10.1038/ejhg.2010.134; published online 18 August 2010

Keywords: SLC6A8; XLMR; splicing; LOVD

INTRODUCTION

In 2001, the X-linked mental retardation (XLMR) syndrome, SLC6A8 deficiency, was identified because of a creatine deficiency in the brain of an infant with severe developmental delay and autistic behavior. Laboratory hallmarks of SLC6A8 deficiency include a reduction of the creatine signal in the proton magnetic resonance spectroscopy (H-MRS) of the brain, an increased urinary creatine/creatinine ratio and impaired creatine uptake in cultured fibroblasts. In female carriers, learning disabilities of varying degrees, including one female. Another variant was predicted to moderately affect splicing by all five tools. However, transient transfection of a minigene containing the variant in a partial SLC6A8 segment showed no splicing errors, and thus was finally classified as non-disease causing. This study shows that splice tools are useful for the characterization of the majority of variants, but also illustrates that the actual effect can be misclassified in rare occasions. Therefore, further laboratory studies should be considered before final conclusions on the disease-causing nature are drawn. To provide an accessible database, the 109 currently known SLC6A8 variants, including 35 novel ones, are included in a newly developed LOVD DNA variation database.

European Journal of Human Genetics (2011) 19, 56–63; doi:10.1038/ejhg.2010.134; published online 18 August 2010

Keywords: SLC6A8; XLMR; splicing; LOVD

INTRODUCTION

In 2001, the X-linked mental retardation (XLMR) syndrome, SLC6A8 deficiency, was identified because of a creatine deficiency in the brain of an infant with severe developmental delay and autistic behavior. Laboratory hallmarks of SLC6A8 deficiency include a reduction of the creatine signal in the proton magnetic resonance spectroscopy (H-MRS) of the brain, an increased urinary creatine/creatinine ratio and impaired creatine uptake in cultured fibroblasts. In female carriers, learning disabilities of varying degrees have been noted. Two studies of males with XLMR estimated the prevalence of SLC6A8 deficiency to be 2.1% (CI: 0.44–3.76) and 1.5% (CI: 0–4.46). A third study of males with MR revealed a prevalence of 0.8% (CI: 0.02–1.7%). Within the last decade many novel variants (including 28 novel). These variants are difficult to classify as pathogenic or non-disease causing because the variants are either located in the coding region, but are synonymous, or are caused by mutations in the creatine transporter (SLC6A8) gene (MIM 300036). SLC6A8 has been mapped to Xq28 and is a member of the solute-carrier family 6 (neurotransmitter transporters). The clinical presentation of males affected with SLC6A8 deficiency is mental retardation (MR), expressive speech and language delay, epilepsy, laboratories interested in the SLC6A8 gene, we developed a novel development and autistic behavior. Laboratory hallmarks LOVD database (http://www.LOVD.nl/SLC6A8), which includes a reduction of the creatine signal in the proton magnetic clinically and genetically relevant data.

MATERIALS AND METHODS

Subjects

In our diagnostic unit, a total of 1900 patients with a differential diagnosis of SLC6A8 deficiency were analyzed by DNA sequence analysis. This resulted in the detection of 66 individuals with intronic or synonymous variants (including 28 novel). These variants are addressed in this study. Also, in the SLC6A8 gene have been detected. These variants are difficult to five novel patients affected with SLC6A8 deficiency are reported.

1Metabolic Unit, Department of Clinical Chemistry, VU University Medical Center, Amsterdam, The Netherlands; 2Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; 3Medical Genetics Center, National Institute of Health, INSA, Porto, Portugal; 4Department of Human Genetics, Radboud University Nijmegen Medical Center, Nijmegen, The Netherlands; 5Greenwood Genetics Center, Greenwood, SC, USA; 6Department of Metabolic Disorders, Necker-Enfants Malades Hospital, Paris, France; 7Department of Pediatrics, University of Mississippi Medical Center, Jackson, MS, USA; 8South Australian Clinical Genetics Service, Women’s and Children’s Hospital, North Adelaide, South Australia, Australia; 9Department of Neur opediatrics, DRK Children Clinic Siegen, Siegen, Germany; 10Department of Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands

*Correspondence: Dr GS Salomons, Department of Clinical Chemistry (Metabolic Unit PK 1 X 009), VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands. Tel: +31 20 4442738; Fax: +31 20 4440305; E-mail: g.salomons@vumc.nl

Received 10 March 2010; revised 23 June 2010; accepted 24 June 2010; published online 18 August 2010

Keywords: SLC6A8; XLMR; splicing; LOVD

INTRODUCTION

In 2001, the X-linked mental retardation (XLMR) syndrome, SLC6A8 deficiency, was identified because of a creatine deficiency in the brain of an infant with severe developmental delay and autistic behavior. Laboratory hallmarks of SLC6A8 deficiency include a reduction of the creatine signal in the proton magnetic resonance spectroscopy (H-MRS) of the brain, an increased urinary creatine/creatinine ratio and impaired creatine uptake in cultured fibroblasts. In female carriers, learning disabilities of varying degrees have been noted. Two studies of males with XLMR estimated the prevalence of SLC6A8 deficiency to be 2.1% (CI: 0.44–3.78) and 1.5% (CI: 0–4.46). A third study of males with MR revealed a prevalence of 0.8% (CI: 0.02–1.7%). Within the last decade many novel variants (including 28 novel). These variants are addressed in this study. Also, in the SLC6A8 gene have been detected. These variants are difficult to classify as pathogenic or non-disease causing because the variants are either located in the coding region, but are synonymous, or are caused by mutations in the creatine transporter (SLC6A8) gene (MIM 300036). SLC6A8 has been mapped to Xq28 and is a member of the solute-carrier family 6 (neurotransmitter transporters). The clinical presentation of males affected with SLC6A8 deficiency is mental retardation (MR), expressive speech and language delay, epilepsy, laboratories interested in the SLC6A8 gene, we developed a novel development and autistic behavior. Laboratory hallmarks LOVD database (http://www.LOVD.nl/SLC6A8), which includes a reduction of the creatine signal in the proton magnetic clinically and genetically relevant data.

MATERIALS AND METHODS

Subjects

In our diagnostic unit, a total of 1900 patients with a differential diagnosis of SLC6A8 deficiency were analyzed by DNA sequence analysis. This resulted in the detection of 66 individuals with intronic or synonymous variants (including 28 novel). These variants are addressed in this study. Also, in the SLC6A8 gene have been detected. These variants are difficult to five novel patients affected with SLC6A8 deficiency are reported.
Methods

PCR of exon 1–13 of SLC6A8 (NM_005629.1). Exon 2 to 13 and flanking intronic sequences of SLC6A8 were amplified using HotStarTaq Polymerase (Qiagen, Valencia, CA, USA). Amplification consisted of an initial denaturation step at 95 °C for 15 min, followed by 38 cycles of 94 °C for 45 s, 66 °C for 45 s and 72 °C for 80 s. For the amplification of exon 1 and its flanking sequences, Takara LA-Taq (Takara Bio Inc., Otsu, Shiga, Japan) Polymerase was used. After an initial denaturing step for 1 min at 94 °C, amplification was allowed in 35 cycles composed of 95 °C for 30 s, 66 °C for 30 s and 72 °C for 80 s.

RT-PCR analysis of SLC6A8 cDNAs and minigenes. RNA was isolated from lymphoblasts or fibroblasts using the SV RNA kit (Promega, Madison, WI, USA) or from Pax blood tubes (Qiagen). cDNA was synthesized from isolated RNA using Omniscript Reverse Transcriptase (Qiagen) according to the manufacturer’s instructions. RT-PCR was performed using cDNA primers designed for the exons of interest of the SLC6A8 gene. To rule out possible amplification of genomic DNA, RT-PCR for each mRNA was also performed without reverse transcriptase.

Construction and transfection of the minigenes. Owing to the large size of the SLC6A8 gene and its highly GC-rich 3’ region, we only cloned the region of interest of the SLC6A8 gene. As templates for the PCR, the patients’ and a wild-type genomic DNAs were used. The fragment, covering exons 3–7 including 53 and 55 nucleotides of the flanking 5’ and 3’ intronic regions respectively, was amplified with forward primer 5’-CCGGAATTCCATACGGCAG-3’ and reverse primer 5’-GCTGGAGATCCTTGGTCCG-3’ containing an EcoRI site and reverse primer 5’-GGGACCGTTCGCGCTTG-3’ and reverse primer 5’-GGGACCGTTCGCGCTTG-3’ containing an XhoI site and an XbaI site for cloning into the pBABE-puro plasmid. The fragment was amplified using Takara LA-Taq (Takara Bio Inc.), cloned into the TOPO-TA (Invitrogen, Breda, The Netherlands) vector and either wild-type or mutant sequence was confirmed by sequencing. For the construction of the positive control (ie, c.777+2T>A), site directed mutagenesis was performed in the TOPO vector containing the wild-type fragment with forward primer 5’-AAATCCACGGGAAAGGAACCACTAGAGGCATGC-3’ and reverse primer 5’-GTCGACGACGGGAAGCTTTGATTCGATCCATC-3’. The obtained sequences were analyzed using the Mutation Surveyor software package (Softgenetics, State College, PA, USA).

DNA sequence analysis. Sequence analysis was performed using BigDye v3.1 terminator and an ABI 3130xl (Applied Biosystems, Nieuwerkerk aan de IJssel, The Netherlands). The obtained sequences were analyzed using the Mutation Surveyor software package (Softgenetics, State College, PA, USA).

Analysis of amplicons with DHPLC. For DHPLC analysis, optimal hetero-meric sequences were used for the amplicon. The Tm of this amplicon was determined as 64 °C using GenscanW9 (http://genes.mit.edu/GENSCAN.html) and FSplice (http://linux1.softberry.com/berry.phtml?topic=fs splice&group=programs&subgroup=genfnd). The parameters used for analysis were the default settings of all tools.

LOVD database. The target of the LOVD database is to include all published and unpublished mutations and variants. It is localized at a server in Leiden, The Netherlands, which can be reached by the following URL http://www.LOVD.nl/SLC6A8 or through the Variation Databases page of the Human Genome Variation Society (HGVS, www.HGVS.org/). The SLC6A8-specific database has been developed with the recently described LOVD software,10 which is in agreement with the HGVS guidelines.11 All variants are checked and approved by Mutalyzer.12 So far 38 pathogenic mutations have been reported in 44 male patients affected with SLC6A8 deficiency.

RESULTS

Validation of splice-site analysis using previously reported variants

We analyzed seven pathogenic mutations (Table 1) detected in patients affected with SLC6A8 deficiency by applying five splice-site analysis tools (see methods). In six out of seven pathogenic mutations at least four of the five tools predicted a reduced recognition of the splice sites, varying with a reduction of the probability score between 9 and 100%. Three mutations had significant reduction scores of >99% in four of the five tools, whereas three other variants had reduction of their scores in a much lower range. One mutation, c.1392+24_1393-30del, reducing the intron from 76 to 52 bp, was not recognized by all the tools. Although analysis of this mutation with Splice Predictor resulted in the loss of both rho and gamma values, this was not ascribed to the mutation itself but to Splice Predictor’s precondition of an intron size of at least 60 bp.

In addition, 18 previously reported variants that are not disease causing were correctly predicted not to cause erroneous splicing by all five splice-site analysis tools, with the exception of one synonymous variant that showed in four tools minor reduction scores of equal or <7%. These variants are all confirmed to be non-disease causing, because we now either established at the cDNA level that no erroneous splicing occurred or that the variants were detected in healthy control males or were published in dbSNP (see Table 1; see www.LOVD.nl/SLC6A8).

Detection and classification of novel variants

Within the last decade, we have analyzed about 1900 individuals (referred to our department because of, eg, MR, increased urinary creatine/creatinine) by DNA sequence analysis of the SLC6A8 gene to exclude/confirm creatine transporter defect. In this endeavor, we identified 41 variants including 28 novel synonymous and intronic variants that were not included in the dbSNP database and of which their (non-) disease-causing nature needed to be established. All were subjected to the five splice-site analysis tools. In total, three novel pathogenic mutations (c.263-1G>C, c.778-2A>G and c.1596+1G>A) were detected that were predicted to cause aberrant splicing with reduction scores of 100% by at least three splice-site analysis tools (Table 2). Twenty-seven variants showed no effect by all five tools. By only one or two tools eight variants were predicted to decrease the possibility of correct splicing by 10% or less. One synonymous unclassified variant (c.780C>T) showed in three splice-site analysis tools reduction scores that appear significant in terms of percentage, but if the normal range is taken into account these differences are considered minimal (eg, the score by the FruitFly tool shows a decrease in 32% while the actual score decreases from 0.22 to 0.15 with the range of low probability to high probability being 0–1.00). All these variants were considered non-disease causing. Only one intronic unclassified variant (c.777+4C>T) was predicted by all five splice-site

European Journal of Human Genetics
### Table 1: Proven pathogenic mutations and non-disease associated variants in SLC6A8

<table>
<thead>
<tr>
<th>Location</th>
<th>cDNA</th>
<th>RNA</th>
<th>Protein</th>
<th>Controls</th>
<th>Netgene2</th>
<th>Fructify</th>
<th>Splice predictor</th>
<th>Genomen W</th>
<th>FSplice</th>
<th>Path.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>c.262+14del8</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>1.00 (0-100)</td>
<td>0.99-0.01 (99%)</td>
<td>0.99-0.01 (100)</td>
<td>170.39-13 (31)</td>
<td>12.4-0.1 (100)</td>
<td>+/-</td>
<td>Winkel et al.</td>
</tr>
<tr>
<td>02</td>
<td>c.253-3G</td>
<td>r.263.32del81</td>
<td>p.Glu88[Glu108del61]</td>
<td>0.280</td>
<td>0.95-0.01 (100)</td>
<td>0.61-0.01 (100)</td>
<td>0.97-0.01 (100)</td>
<td>21.9-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Schwarzenbacher et al.</td>
</tr>
<tr>
<td>03</td>
<td>c.1141+40del4</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Loni-Franco et al.</td>
</tr>
<tr>
<td>04</td>
<td>c.912-1G</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Loni-Franco et al.</td>
</tr>
<tr>
<td>05</td>
<td>c.913+10G</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Loni-Franco et al.</td>
</tr>
<tr>
<td>06</td>
<td>c.1016+17C</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Loni-Franco et al.</td>
</tr>
<tr>
<td>07</td>
<td>c.1141+18delA</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
<td>+/-</td>
<td>Loni-Franco et al.</td>
</tr>
</tbody>
</table>

### Table 2: Twenty-eight novel intronic variants and 13 previously reported unclassified variants in SLC6A8

<table>
<thead>
<tr>
<th>Location</th>
<th>cDNA</th>
<th>Deduced effect</th>
<th>Netgene2</th>
<th>Fructify</th>
<th>Splice predictor</th>
<th>Genomen W</th>
<th>FSplice</th>
<th>Path.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>c.262+65delA</td>
<td>p.(?)</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>+/-</td>
<td>Wilcken et al.</td>
</tr>
<tr>
<td>02</td>
<td>c.253-1G</td>
<td>r.263.32del81</td>
<td>p.Glu88[Glu108del61]</td>
<td>0.280</td>
<td>0.95-0.01 (100)</td>
<td>0.61-0.01 (100)</td>
<td>0.97-0.01 (100)</td>
<td>21.9-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
</tr>
<tr>
<td>03</td>
<td>c.1141+40del4</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
</tr>
<tr>
<td>04</td>
<td>c.912-1G</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
</tr>
<tr>
<td>05</td>
<td>c.913+10G</td>
<td>r.S(?)</td>
<td>p.(?)</td>
<td>0.280</td>
<td>0.99-0.01 (100)</td>
<td>0.99-0.01 (100)</td>
<td>0.97-0.11 (11.3)</td>
<td>24.3-5.9 (73%)</td>
<td>12.5-0.1 (100)</td>
</tr>
</tbody>
</table>

### Notes
1. This mutation arose de novo.
2. Location refers to the intron (i) or exon (two digits) number in which the variant was found. In the upper panel the proven pathogenic variants are denoted. The lower panel includes both intronic and synonymous variants. NC: no change (ie, the variant did not influence the predicted score compared with the canonical site). Path.: reported and concluded pathogenicity. +/-: no known pathogenicity; +/-: probably no pathogenicity; +: pathogenic; +/+: pathogenic.
3. cDNA denotation is based on the reference sequence NM_005629.3, where +1 corresponds to the first nucleotide of the initiation codon. All variants are annotated according to the guidelines of Den Dunnen and Antonarakis (http://www.genomic.unimelb.edu.au/mid/mutnomen/).

### Table 2: Twenty-eight novel intronic variants and 13 previously reported unclassified variants in SLC6A8

<table>
<thead>
<tr>
<th>Location</th>
<th>cDNA</th>
<th>Deduced effect</th>
<th>Netgene2</th>
<th>Fructify</th>
<th>Splice predictor</th>
<th>Genomen W</th>
<th>FSplice</th>
<th>Path.</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>c.262+65delA</td>
<td>p.(?)</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>NC</td>
<td>+/-</td>
<td>Wilcken et al.</td>
</tr>
</tbody>
</table>

1. This variant was cloned into a minigene. Overexpression studies indicated that this variant does not affect splicing.
2. None of these variants, including 3 pathogenic mutations and 13 previously reported variants were detected in 280 control alleles. Location refers to the intron (i) or exon (two digits) number in which the variant was found. NC: no change, (ie, the variant did not influence the predicted score compared with the canonical site).
analysis tools to reduce the probability of the canonical donor site of exon 4 (Table 2) by 10, 2, 23, 10 and 15% compared with the canonical donor site score. cDNA was not available from this patient and thus this variant was studied by overexpression of a minigene (ie, a genomic SLC6A8 segment containing the c.777+4C>T variant). Both overexpression of the wild-type minigene as well as the mutant minigene in SLC6A8-deficient primary fibroblasts showed normal splicing as detected by RT-PCR and sequence analysis. This is in contrast to the positive control minigene that contained the c.777+2T>A mutation and resulted in skipping of exon 4 (Figure 1).

Five novel SLC6A8-deficient patients
The first case (patient I) was a young female (DOB January 1993) born at term following a normal pregnancy and delivery with normal weight, length and head circumference. From the age of 2, she developed a behavior disturbance and was assessed by a pedopsychiatric institution that also diagnosed mild MR. At age 14, she was affected with auditory hallucinations and the diagnosis of chronic hallucinatory psychosis was made. Her cognitive impairment was stable with conserved language capacities and social contact, no motor impairment and no specific neurological signs. Her total IQ was tested at 45 (verbal: 61, reason: 45, memory: 53; and working velocity: 66). A brain H-MRS was performed which showed a reduced cerebral creatine level with a normal appearance to choline and N-acetylaspartic acid (Figure 2). Urinary creatine levels were within the upper normal range. Plasma creatine was determined twice resulting in values of 39.9 μmol/l and 75.4 μmol/l (normal 30–124 μmol/l). Genomic DNA sequencing revealed a c.263-1G>C mutation. This mutation arose de novo. Somatic mosaicism was not detected in her parents’ DNA by both DNA analysis and the use of DHPLC (Figure 3). mRNA isolated from PAX tubes, followed by cDNA synthesis showed approximately

![Figure 1 RT-PCR results of spliced products after transfection of the minigenes in SLC6A8-deficient fibroblasts. SLC6A8-deficient fibroblast were transfected with a minigene, containing an exon 3 to 7 fragment of SLC6A8 with either a wild-type sequence, the c.777+2T>A transversion or the c.777+4C>T variant. Cells were harvested and RNA was isolated using the Promega RNA Isolation Kit. Subsequent RT-PCR of the wild-type and the c.777+4C>T minigene transfected fibroblasts resulted in an amplicon of 384 bp, indicating the variant not to cause aberrant splicing. The positive control (c.777+2T>A transfecants) resulted in a 251 bp amplicon. Sequencing analysis showed that exon 4 was skipped in the 251 bp amplicon.](image1)

![Figure 2 Proton MRS of the brain of an index girl with SLC6A8 deficiency (patient I). A proton MRS was performed on a 14-year-old girl with mild mental retardation. Creatine (Cr) was found to be reduced with a normal appearance to choline (Cho) and N-acetylaspartic acid (NAA). Subsequent genomic DNA sequencing revealed a pathogenic c.263-1G>C mutation, which confirmed the diagnosis of SLC6A8 deficiency.](image2)

![Figure 3 Patient I with a c.263-1G>C mutation in the SLC6A8 gene. Somatic mosaicism for the c.263-1G>C mutation was not detected in both parents of patient I. The DHPLC elution profile of the exon 2 amplicon of EDTA blood DNA of patient I with a c.263-1G>C variant in SLC6A8 including her parents and a male control. At the retention time of the heteroduplex peak clearly visible with the patient, no peak is visible with the parents, indicating both are not somatic mosaic for the mutation.](image3)
90% r.263_325del (p.Gly88_Leu108del) and 10% wild-type product. The X-inactivation was studied in blood with the analysis of HhaI digested and undigested DNA followed by PCR of the highly polymorphic CAG repeat of the androgen receptor gene and showed a 90:10 pattern. Also, fibroblasts incubated with a physiological concentration of creatine (25 μM) showed no uptake. All these findings combined confirmed creatine transporter deficiency.

The second case (patient II) was a 7-year-old male with epilepsy, expressive language difficulties and a movement disorder. He was tested for mutations in SLC6A8 after H-MRS of the brain revealed a diminished creatine peak with a normal appearance to NAA and choline and urinary analysis showed an increased Cr/Cr ratio. This resulted in the detection of a hemizygous c.778-2A>T variant. His brother (patient III), currently 10 years old, showed a similar clinical phenotype with an increased urinary Cr/Cr ratio and absence of cerebral creatine. He was also found to be hemizygous for this variant while his mother was proven to be heterozygous.

The fourth patient (patient IV) was a young male with a reduced cerebral creatine level measured by H-MRS. Further investigation resulted in the detection of a c.1596+1G>T variant, which was also found to be heterozygous in DNA from his mother.

The fifth patient (patient V) was unmistakably diagnosed with SLC6A8 deficiency at the age of 6 years. The patient suffered from moderate MR and severe language delay. Quantitative localized single voxel magnetic resonance spectroscopy was performed over the basal ganglia and revealed a markedly diminished creatine level with a normal appearance to choline. Urinary analysis revealed a creatine/creatinine value of 1.7 (normal 0.017–0.72). The creatine uptake ability of the patient’s fibroblasts was significantly decreased in comparison with control cells when incubated at a physiological concentration of creatine (25 μM).

Moreover, the mutation could not be detected in the DNA of the mother and is therefore considered de novo. At the cDNA level, two erroneous transcripts were revealed (r.(c.1392_1393ins1393+1_1393-1;1392+24_1393-30del,1393_1495del)) (Figure 4). This conclusively classified the c.1392+24_1393-30del mutation as the pathogenic event. Also, the same de novo mutation was detected in an unrelated patient.

**LOVD database**

In total, the LOVD database of SLC6A8 lists 44 SLC6A8-deficient families, with a total of 43 mutations. Of these, 38 are proven pathogenic (+/+), whereas 5 others are presumed to be pathogenic (+/+?). The latter are all missense variants, which have not been investigated by overexpression studies yet.

The database (Figure 5) provides detailed information on the nature of the variants, but also exon–intron location of the DNA change (both according to current nomenclature and as published), RNA change and protein change. To all variants a unique database number is assigned. The column ‘variant remarks’ provides a detailed description on the nature of the variant. The final conclusion is presented in the first column ‘path.’ in which the first sign provides the conclusion of the report and the second one that of the curator (eg. +/+ meaning confirmed pathogenic). Other columns include detailed information on the variant origin, reference, template, technique, frequency, disease status, patient remarks, times reported, gender, geographic origin and ethnic origin. Moreover, clinical symptoms are continuously archived (non-public). The database is an excellent tool to get an up-to-date overview of the type of mutations, the frequency of specific variants and more details. This data can be found in the different tabs by any user, but an advanced search function is also available. Researchers or clinicians who are interested in the nature of a specific variant can use the database to search if the variant has been detected in other patients. This information may have been published in reports or on the website only, but also may include unpublished variants, which have not yet been made public. The latter allows contact between clinicians even if the data are not publicly released. All scientists/clinicians are invited to submit their data to the database. The identity of the submitter is included in the database.

---

**Figure 4** Schematic representation of the c.1392+24_1393-30del variant in SLC6A8. The c.1392+24_1393-30del variant was found in two independent patients and occurred de novo in both cases. In the patients’ DNA a hemizygous variant, c.1392+24_1393-30del, was found (upper pane). mRNA analysis showed that this variant affects splicing and results in different splice products, r.(c.1392_1393ins1392+1_1393-1;1392+24_1393-30del,1393_1495del) (lower pane), resulting in SLC6A8 deficiency. For the second patient see Hathaway et al. This variant is annotated according to the guidelines of Den Dunnen and Antonarakis (http://www.genomic.unimelb.edu.au/mdi/mutnomen/).
In this study, the nature of 67 intronic and synonymous variants has been studied by free web-based splice-site analysis tools. These have been characterized and are all being included in a newly developed LOVD database. In addition, all previously published mutations are included in this database as well. This is of high relevance for researchers and clinicians who detect difficult to interpret variants in the SLC6A8 gene. In an ideal diagnostic setup, an unclassified variant is further investigated at the mRNA, protein and/or functional level to confirm/exclude its pathogenic nature. However, in diagnostic laboratories time is usually limited and proper materials are often not available. Therefore, alternative methods are warranted for variant classification such as frequently used splice-site analysis tools, which predict the possible effect on RNA splicing of a variant based on known and computed conserved splicing sequences. Here we demonstrate that the use of five splice-site analysis tools, indeed is very helpful for proper classification.

In total, 24 out of 25 (96%) variants were properly classified by the combined use of the five tools. The importance of using more than one splice-site analysis tool is illustrated by the fact that four out of seven proven pathogenic mutations were not predicted to have a strong effect (ie, 90% or higher reduction) by one or more of the individual analysis tools, whereas the combined data of all the tools clearly predicted erroneous splicing for six of these variants (Table 1). In one case, only one tool predicted a reduction in the probability of correct splicing. The fact that this ‘missed’ mutation (c.1392+24_1393-30del) comprises an intronic deletion of 24 bp out of a 76 bp intron, easily explains this omission as none of the five splice-site analysis tools have intron size as a parameter. This caveat is an important finding for diagnostic laboratories that encounter this specific type of mutations (eg, deletions, insertions in small introns). We demonstrated that this deletion indeed results in erroneous spliced transcripts (r.[1392_1393ins1392+1_1393-1;1392+24_1393-30del,1393_1495del]), (Hathaway et al,14 this study).

DISCUSSION

In this study, the nature of 67 intronic and synonymous variants has been studied by free web-based splice-site analysis tools. These have been characterized and are all being included in a newly developed LOVD database. In addition, all previously published mutations are included in this database as well. This is of high relevance for researchers and clinicians who detect difficult to interpret variants in the SLC6A8 gene. In an ideal diagnostic setup, an unclassified variant is further investigated at the mRNA, protein and/or functional level to confirm/exclude its pathogenic nature. However, in diagnostic laboratories time is usually limited and proper materials are often not available. Therefore, alternative methods are warranted for variant classification such as frequently used splice-site analysis tools, which predict the possible effect on RNA splicing of a variant based on known and computed conserved splicing sequences. Here we demonstrate that the use of five splice-site analysis tools, indeed is very helpful for proper classification.
In 17 out of 18 proven variants, no change in prediction scores was detected by any of the five tools. Only one variant that was proven not to interfere with proper splicing, had a reduced score of 7% or less compared with the canonical site by four of the five tools. Owing to this variant, we arbitrarily decided to only consider variants for further molecular workup if at least three tools predicted a mild effect (10<80%), or if one or two tools predicted a significant effect (>80%). In addition, we took into account the prediction score of the wild-type splice site. Thus, a 50% reduction of a canonical site with a 0.99 score is more likely to have an effect on splicing than a 50% reduction in a splice site with a much weaker score of for instance 0.22.

Utilizing these criteria, we analyzed the potential effect of 41 unclassified variants on the SLC6A8 mRNA with the use of these five splice-site analysis tools. This resulted in the identification of three pathogenic mutations (c.263-1G>C, c.777-2A>G and c.1596+1G>A), and 38 variants that have no apparent relation with creatine transporter deficiency. The effect on splicing for two of the three novel pathogenic mutations could not be confirmed at the cDNA level. However, given the fact that all three mutations affect the donor or acceptor splice-site, all were classified as pathogenic. Pathogenicity was in excellent agreement with the fact that the mutations were not detected in 280 control chromosomes, the biochemical findings (increased urinary Cr/Crnm) and/or reduced cerebral creatine detection by H-MRS.

Interestingly, one novel mutation (c.263-1G>C) was found in a girl, who according to our knowledge is the first reported index girl affected with SLC6A8 deficiency. The girl was not correctly diagnosed until the age of 14 when a brain H-MRS revealed a reduced cerebral creatine level. This led to DNA analysis that resulted in the discovery of the pathogenic c.263-1G>C mutation. Additional investigations revealed a urinary creatine to creatinine ratio within the normal range as well as normal values for plasma creatine. The discrepancy between the urinary, plasma and cerebral creatine levels is possibly caused by variation of the X-inactivation pattern in different tissues, which was shown to be skewed in blood of this patient.

In a recently described cohort of female relatives who were heterozygous for a pathogenic mutation in SLC6A8, it was shown that symptoms of SLC6A8 deficiency (eg, MR, learning difficulties and constipation) can occur in female heterozygotes. However, it should be noted that the level of biochemical markers (urinary creatine to creatinine ratio and/or cerebral creatine) usually overlap with that of normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls. This is in concordance with the clinical and biochemical findings of the above mentioned female index patient and normal controls.

In conclusion, splice-site analysis tools are very important in the process of classifying novel variants. For definite classification of novel variants outside the canonical donor and acceptor sites, in vitro experimentation, either with mRNA analysis or with the use of a minigene (if no additional material is available), is an essential procedure.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
We thank Donald E Brass for his excellent reviewing. The work of Otir T Betsalel, Efrain H Rosenberg, Ligia S Almeida and Gaia S Salomons is supported by the Dutch society for Scientific Research (ZonMW/NWO), VIID grant number 917.56.349 and the Horst Bickel Award 2002, that of Tijtse Kleefstra by the European Union grant QLGI3-CT-2002-01819 EURO-MRX, that of Johan T den Dunnen by the European Community’s Seventh Framework Programme (FP/2007-2013) under grant agreement no. 200754 – the GEN2PHEN project and that of Charles Schwartz by grants of NICHD (HD26202) and the South Carolina Department of Disabilities and Special Needs.


