



**HAL**  
open science

## **SGCE isoform characterization and expression in human brain: implications for Myoclonus-Dystonia pathogenesis?**

Katja Ritz, Barbera van Schaik, Marja Jakobs, Antoine van Kampen, Eleonora Aronica, Marina Tijssen, Frank Baas

### ► **To cite this version:**

Katja Ritz, Barbera van Schaik, Marja Jakobs, Antoine van Kampen, Eleonora Aronica, et al.. SGCE isoform characterization and expression in human brain: implications for Myoclonus-Dystonia pathogenesis?. *European Journal of Human Genetics*, 2010, <10.1038/ejhg.2010.206>. <hal-00600445>

**HAL Id: hal-00600445**

**<https://hal.science/hal-00600445v1>**

Submitted on 15 Jun 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

**SGCE isoform characterization and expression in human brain: implications  
for Myoclonus-Dystonia pathogenesis?**

Katja Ritz<sup>1,4</sup>, Barbera DC van Schaik<sup>2</sup>, Marja E Jakobs<sup>1</sup>, Antoine H van Kampen<sup>2,5</sup>, Eleonora Aronica<sup>3</sup>, Marina A Tijssen<sup>4</sup>, Frank Baas<sup>1</sup>

<sup>1</sup>Department of Genome Analysis, <sup>2</sup>Bioinformatics Laboratory, <sup>3</sup>Department of (Neuro)Pathology, <sup>4</sup>Department of Neurology, Academic Medical Center, University of Amsterdam, Meibergdreef 9, 1105AZ Amsterdam, The Netherlands, <sup>5</sup>Biosystems Data Analysis, Swammerdam Institute for Life Science, University of Amsterdam, Nieuwe Achtergracht 166, 1018WV Amsterdam, The Netherlands

**Address for correspondence**

Frank Baas, PhD, MD

Department of Genome Analysis

Academic Medical Center

Meibergdreef 9

1105AZ Amsterdam

The Netherlands

Phone +31-205663846

Fax +31-205669312

Email: f.baas@amc.uva.nl

## 1 **Abstract**

2 Myoclonus-Dystonia (M-D) is a neurological movement disorder with involuntary jerky  
3 and dystonic movements as major symptoms. About 50% of M-D patients have a  
4 mutation in epsilon-sarcoglycan (*SGCE*), a maternally imprinted gene that is widely  
5 expressed. Since little is known about *SGCE* function one can only speculate about the  
6 pathomechanisms of the exclusively neurological phenotype in M-D. We characterized  
7 different *SGCE* isoforms in the human brain using ultra deep sequencing. We show that a  
8 major brain-specific isoform is differentially expressed in the human brain with a notably  
9 high expression in the cerebellum, namely in the Purkinje cells and neurons of the dentate  
10 nucleus. Its expression was low in the globus pallidus and moderate to low in caudate  
11 nucleus, putamen and substantia nigra. Our data is compatible with a model in which  
12 dysfunction of the cerebellum is involved in the pathogenesis of M-D.

13

14

15 Keywords: Myoclonus-Dystonia, *SGCE*, deep sequencing, alternative splicing,  
16 imprinting, cerebellum

## 17 Introduction

18 Myoclonus-Dystonia (M-D) is a movement disorder characterized by myoclonic  
19 jerks and dystonic features, usually affecting the upper part of the body. Little is known  
20 about the pathological basis of M-D symptoms. Electrophysiological studies in man  
21 suggest that myoclonic symptoms are of subcortical origin.<sup>1-3</sup> In general, dystonia is  
22 thought to arise from dysfunction of the basal ganglia.<sup>4</sup> About 50% of M-D patients that  
23 were classified as definite M-D carry a mutation in the widely expressed  $\epsilon$ -sarcoglycan  
24 (*SGCE*).<sup>5-7</sup> The genetic cause in the remaining patients is still unclear. A second locus has  
25 been reported in one large M-D family (DYT15, 18p11), but no gene has been identified  
26 yet.<sup>8,9</sup> SGCE is part of the sarcoglycan family that consists of N-glycosylated  
27 transmembrane proteins. Six different sarcoglycans have been identified so far ( $\alpha$ -,  $\beta$ -,  $\gamma$ -,  
28  $\delta$ -,  $\epsilon$ -,  $\zeta$ -) but little is known about the function of particular sarcoglycan members and their  
29 function in different tissues.<sup>10</sup> In muscle, sarcoglycans form a heterotetrameric complex  
30 that is constituent of the dystrophin-associated protein complex. This complex mediates the  
31 structural stability of the plasma membrane and interactions between the extracellular  
32 matrix and the cytoskeleton.<sup>11</sup> Mutations in other sarcoglycans,  $\alpha$ -,  $\beta$ -,  $\gamma$ - and  $\delta$ -  
33 sarcoglycan, lead to different forms of limb-girdle muscular dystrophy, characterized by  
34 progressive muscle weakness.<sup>10</sup> Little is known about composition and function of the  
35 dystrophin-associated protein complex in brain. *SGCE* is highly homologous to  $\alpha$ -  
36 sarcoglycan,<sup>12</sup> but no muscle or myocardial muscle abnormalities have been identified in  
37 M-D patients.<sup>13</sup>

38 Understanding the exclusively neurological phenotype of the disease is still a major  
39 challenge. *SGCE* is an imprinted gene<sup>14,15</sup> meaning that a loss of function mutation in the

40 expressed allele will be dominant. All reported mutations behave as null alleles, since they  
41 result in premature stop codons and are predicted to lead to nonsense-mediated decay.  
42 There is also evidence that *SGCE* missense mutations lead to loss of function.<sup>16</sup> *SGCE* is  
43 widely expressed which suggests either redundancy for this protein in non-brain tissue or a  
44 brain-specific function for *SGCE*.

45 There are four known alternatively spliced exons in the *SGCE* gene (exon 2, 8, 10  
46 and 11b, Fig. 1A) with exon 11b as a brain-specific exon.<sup>17</sup> Recently, a new brain-specific  
47 alternatively spliced exon (exon 11c, an elongated exon 11b) has been identified in mice.<sup>18</sup>  
48 Transcripts containing either exon 11b or 11c encode proteins with a different C-terminal  
49 sequence containing a PDZ-binding motif. This motif is a protein-interaction domain and  
50 thus may contribute to a unique *SGCE* function in brain. The ubiquitous *SGCE* protein and  
51 the brain-specific protein (exon 11b) are localized in different synaptosomal membrane  
52 fractions: post- and pre-synaptic membrane, respectively.<sup>17</sup> It has been speculated that both  
53 isoforms play different roles at neuronal synapses.

54 The aim of our study was to characterize the quantitative and qualitative expression  
55 pattern of *SGCE* isoforms in human brain in order to identify brain regions associated with  
56 M-D.

57

## 58 **Methods**

### 59 **Sample collection and preparation**

60 Human tissue was obtained from six control subjects without neurological  
61 symptoms (Table 1). In addition, two blood samples were obtained from control subjects.  
62 *SGCE* mutations were excluded by direct sequencing of the coding region. Human tissues

63 were obtained from the Department of neuropathology of the Academic Medical Center  
64 (University of Amsterdam, The Netherlands) and informed consent was obtained for  
65 research purposes in all control subjects. Mouse (cerebral cortex), rat (cerebral cortex) and  
66 zebrafish (100 embryos, 24 hours after fertilization) tissue was obtained from wild type  
67 animals (approved by local committee). DNA extractions from frozen brain tissues and  
68 whole-blood samples were performed using standard procedures. Twenty to 40 sections  
69 (depending on tissue size, 20µm) were cut per sample and dissolved in TRIzol Reagent  
70 (Invitrogen, Breda, The Netherlands) for subsequent RNA isolation, or in SE buffer  
71 (750mM NaCl pH 8, 250mM EDTA, 1% SDS) containing Proteinase K for DNA isolation.  
72 Total RNA was isolated with QIAcube instrument (Qiagen, Venlo, The Netherlands) using  
73 RNeasy Mini protocol or using the PAXgene system (Qiagen) for blood samples.

74

#### 75 **Methylation-sensitive high resolution melting assay**

76 To determine the degree of methylation in the promotor region of *SGCE* we applied  
77 a methylation-sensitive high resolution melting (MS-HRM) assay. The region amplified  
78 (nucleotide position -1148 to -773 relative to the start ATG) contains 25 CpG sites and has  
79 been shown to be differentially methylated.<sup>14</sup> 1µg of genomic DNA was chemically  
80 modified with sodium bisulfite using the EZ Methylation kit (Zymo Research, Orange, CA,  
81 USA) and amplified in the presence of ResoLight HRM dye (Roche Diagnostics, Almere,  
82 The Netherlands). Both, methylated and unmethylated, strands were amplified.<sup>14</sup> MS-HRM  
83 analysis was performed with the LightCycler 480 (Roche Diagnostics) and analyzed with  
84 LightCycler 480 gene scanning software module. In each run a set of standards of the  
85 identical *SGCE* promotor region was included (100%, 50%, 25%, 12.5%, 6.25% and 0%

86 methylated). Different brain regions of five control subjects (Subject 1, 2, 4, 5, 6), muscle  
87 tissue of two control subjects (Subject 2, 3) and blood of 12 definite M-D *SGCE* non-  
88 mutation carriers<sup>7</sup> were tested. A subset of PCR amplified samples was cloned (pGEM-T  
89 Easy vector, Promega, Leiden, The Netherlands) and sequenced to verify results of the  
90 methylation assay (ABI big dye v3.1 chemistry, ABI 3730 sequencer, Applied Biosystems,  
91 Foster City, CA, USA).

92

### 93 **Ultra deep amplicon sequencing**

94 cDNA was synthesized from 1µg of total RNA with oligodT<sub>12</sub>-VN primers and 1µl  
95 cDNA was subjected to a 10µl PCR reaction. PCR reactions were performed using fusion  
96 primers consisting of a 19 bp fixed sequence (Roche/454 GS FLX, A or B sequence at the  
97 5'end) and a target-specific sequence (3'end). We added a 5 nucleotide **multiplex**  
98 **identifier-tag** to allow for multiplexing of samples (primer sequences and PCR conditions  
99 are available on request). Amplicons were processed and ultra deep sequenced using the  
100 454 GS FLX system (Roche Diagnostics) according to manufacturer's instructions.

101

### 102 **Data analysis**

103 In order to identify all alternatively spliced events, sequences were first grouped by  
104 a **multiplex identifier-tag**. Secondly, we used the BLAT algorithm (BLAST-like alignment  
105 tool) to compare each sequence with all other sequences to identify similar exon  
106 combinations per sample. Sequences were grouped when they encompassed at least 210 nt,  
107 had a percent identity of at least 98%, a score above 105 and a query coverage above 98%.  
108 The resulting groups were mapped to the respective chromosome (BLAT algorithm, hg18

109 build) and analyzed in the UCSC genome browser and using CodonCode Aligner software  
110 3.0.1 (Dedham, MA, USA).

111

### 112 **Real-time quantitative PCR**

113 1µl cDNA was amplified in triplicates using the LightCycler480 (Roche  
114 Diagnostics) in a final volume of 10µl containing SYBR Green I Master mix (Roche  
115 Diagnostics). Intron spanning primers were designed to target either all *SGCE* isoforms or  
116 the brain-specific isoform containing exon 11b. All quantitative PCR (qPCR) results were  
117 normalized against the housekeeping genes *GAPDH* and *eEF-1 alpha*. Data was analyzed  
118 using LinRegPCR analysis program.<sup>19</sup>

119

### 120 ***In situ* hybridization**

121 *In situ* hybridization (ISH) was performed as previously described.<sup>20</sup> Two 5'-fluorescein-  
122 labeled 19mer antisense oligonucleotides containing locked nucleic acid and 2'-O-methyl  
123 (2OME) RNA moieties (Ribotask ApS, Odense, Denmark) were designed: one targeting  
124 the brain-specific alternatively spliced exon 11b (5'-AacGaaAauCucCugTagT-3', **locked**  
125 **nucleic acid** residues are given in capital letters, **2'-O-methyl**-RNA in lower case) and one  
126 detecting all isoforms targeting the constitutive *SGCE* exon 3 (5'-  
127 TagGacTccAucAcuAuaT-3'). Brain tissue was treated with Proteinase K (200µg/ml,  
128 Roche Diagnostics) at 50°C for 10 minutes. Probe annealing was performed at 53°C for 1.5  
129 hours. Peroxidase activity was visualized using NovaRED (Vectorlabs, Burlingame, CA)  
130 and counterstained with hematoxylin. In an adult control subject we tested different human  
131 brain regions (cerebellum, mesencephalon, caudate nucleus, putamen, globus pallidus,

132 cerebral cortex and hippocampus). Skeletal muscle, heart, kidney and liver sections were  
133 used to confirm specificity of the brain-specific probe. As a control for specific binding,  
134 brain material of a *SGCE* mutation carrier (c.835\_839del, p.Lys280SerfsX16, 50 h post  
135 mortem) was analyzed. This tissue was obtained from the NICHD Brain and Tissue Bank  
136 (Baltimore, MD, USA).

137

## 138 **Results**

### 139 ***SGCE* imprinting and M-D**

140 *SGCE* is widely expressed and there is thus far no explanation for the brain-specific  
141 phenotype of *SGCE* mutation carriers. One explanation could be that brain regions which  
142 are not affected are rescued by expression of the maternally imprinted allele. Therefore, we  
143 analyzed the *SGCE* imprinting throughout the human brain with a MS-HRM assay. We  
144 observed that the imprinting pattern was maintained in all brain regions and tissues tested  
145 (Supplementary Fig. 2). Results of the MS-HRM assay were confirmed by cloning and  
146 direct sequencing of a subset of samples. Only minor variations in the number of  
147 methylated/unmethylated CpGs were detected by cloning but not by the MS-HRM assay.

148

### 149 ***SGCE* mRNA expression by ultra deep amplicon sequencing**

150 Different *SGCE* isoforms have been reported so far, but little is known about their  
151 distribution and expression levels throughout the brain. To qualitatively and quantitatively  
152 characterize alternatively spliced *SGCE* exons, we performed a systematic analysis of  
153 *SGCE* cDNA by ultra deep amplicon sequencing. This technique provides a powerful and  
154 accurate tool to investigate relative mRNA expression levels and detects rare and unknown

155 splice events. To cover all exon-exon boundaries, four overlapping *SGCE* amplicons were  
156 made (exon 1 to 5, exon 4 to 6, exon 5 to 9, exon 7 to 12) and sequenced with a 454 GS  
157 FLX sequencer. Ultra deep sequencing was achieved by aiming for 10 000 sequence reads  
158 per amplicon (average ( $\pm$  SD): 11 107 ( $\pm$ 4 273) reads, Supplementary Table 1).

159

#### 160 *SGCE mRNA isoforms*

161 To get an overview of all alternative splicing events and their tissue specificity, we  
162 analyzed alternative splicing events of the entire *SGCE* gene in one control subject (Subject  
163 1) in the primary somatosensory cortex (SM1), heart and blood. Twenty-three different  
164 splicing events of the *SGCE* gene were detected, but only four of them occurred at  
165 frequencies above 1%: exon 1c and the known alternatively spliced exons 2, 8 and 11b  
166 (Fig. 1B, Table 2). In-frame exon 1c was expressed in brain (5.7%), but also in non-brain  
167 tissue (2% - 2.3%). The in-frame exon 2 showed an overall high inclusion level in all  
168 tissues tested, whereas exon 8 was highly represented in muscle and blood and low in  
169 brain. Exon 11b was mainly expressed in brain (34%); it was not expressed in blood and at  
170 very low levels in muscle (0.05%). These four major variants result in a partially altered  
171 protein coding region and are extended into the last exon. Therefore, we expect these to be  
172 translated and stable. Of the low frequency variants only one, exon 4a, retains a long open  
173 reading frame. All other mRNA isoforms, except all 11b variants, are expected to result in  
174 nonsense mediated decay.<sup>21</sup> Splice site predictions and positions of known and novel exons  
175 are listed in Table 2. Most of the new low frequent variants lead to a frameshift (17/19)  
176 resulting in a premature stop codon and were conserved among primate lineage only (Table  
177 2, Supplementary Table 2). All known constitutive *SGCE* exons were present to 100%

178 except for exon 11: Skipping of this exon was observed in all tissues and species tested,  
179 leading to a premature stop after amino acid 409, albeit at low frequency (brain: 0.8%,  
180 muscle: 1.4%, blood: 0.3%). The recently identified brain-specific exon 11c was expressed  
181 at very low levels in brain but also non-brain tissue ( $\leq 0.7\%$ , Table 2). Also, in mouse and  
182 rat brain it was expressed at very low levels ( $< 2.1\%$ ), it was not expressed in zebrafish.

183 In order, to exclude that these low frequency variants were experimental artifacts,  
184 we sequenced a synthetic *SGCE* RNA using the same procedure. Analysis of 12 000 reads  
185 gave only sequence reads identical to the input RNA (data not shown).

186

#### 187 *Brain region-specific SGCE expression*

188 *SGCE* exon 11b was the most abundant and a highly conserved brain-specific splice  
189 variant. Therefore, we analyzed exon 11b expression in more detail and tested nine  
190 different human brain regions of two control subjects and three different organisms (mouse  
191 and rat brain, zebrafish) by ultra deep amplicon sequencing.

192 Exon 11b showed differential expression among the different brain regions in both  
193 control subjects. Levels were highest in SM1 and motor cortex (M1), low in the globus  
194 pallidus, intermediate in caudate nucleus and substantia nigra, and interindividual variation  
195 was found in the cerebellum and putamen (Table 3). In mouse and rat brain, we observed a  
196 higher rate of exon 11b inclusion compared to human brain (SM1: 50% in mouse, 73% in  
197 rat vs. 35% in human). In whole zebrafish RNA, exon 11b was present in 20% of  
198 transcripts.

199

#### 200 ***SGCE* mRNA expression by qPCR**

201 Ultra deep amplicon sequencing revealed that exon 11b is the major brain-specific  
 202 alternatively spliced *SGCE* exon and that it is differentially expressed in the human brain.  
 203 Similar trends in expression levels of exon 11b among the different brain regions were  
 204 identified, but we also observed interindividual variability for some regions. To further  
 205 characterize the expression pattern of *SGCE* transcripts we tested the same brain regions of  
 206 more control subjects by qPCR (Fig. 2). We found similar trends of exon 11b expression in  
 207 five control subjects: Expression in the cerebellum was significantly higher compared to  
 208 globus pallidus (Fig.2,  $p < 0.001$ , Friedman test, Dunn's multiple comparison test) and  
 209 substantia nigra ( $p < 0.05$ ). Also SM1 and pallidum were significantly different ( $p < 0.05$ ).  
 210 The total *SGCE* expression level (all isoforms together) was determined in all five control  
 211 subjects for the various brain regions. Their means did not differ significantly in the  
 212 different brain regions ( $p = 0.5$ ), suggesting that high exon 11b expression is indeed due to  
 213 differential isoform expression and not only due to differences in total *SGCE* expression  
 214 levels. Overall, *SGCE* and *SGCE* exon 11b expression levels varied among individuals.  
 215 Normalization with reference genes *GAPDH* or *eEF-1* showed comparable results.  
 216 Furthermore, the 454 sequencing and qPCR results are in support of each other  
 217 (Supplementary Fig. 1).

218

### 219 **Localization of *SGCE* mRNA isoforms in human brain**

220 In order, to study *SGCE* mRNA localization of the major brain-specific and the  
 221 ubiquitously expressed transcripts we performed isoform-specific ISH in human brain (Fig.  
 222 3). *SGCE* exon 11b and total *SGCE* expression was most prominent in neurons in all  
 223 regions tested: high signal in the cerebral cortex, cerebellum (Purkinje cells, dentate

224 nucleus), hippocampus (pyramidal neurons of all CA regions) and moderate signal in the  
225 mesencephalon (substantia nigra) and basal ganglia. There was only a faint signal in glial  
226 cells and no staining in the white matter. The exon 11b probe did not detect a signal in  
227 skeletal muscle, heart, kidney and liver, in contrast to the *SGCE* probe targeting all  
228 isoforms. Probe specificity was confirmed with brain tissue of a *SGCE* null mutation  
229 carrier (c.835\_839del, p.Lys280SerfsX16) showing no staining for both probes, whereas it  
230 stained for control genes (*CD68*, *TSEN54*).

231 In view of the localized expression of *SGCE* in the cerebellum, we tested whether  
232 the interindividual variation of *SGCE* exon 11b expression in the cerebellum is due to  
233 variation in Purkinje cell number in the tissue examined. qPCR analysis of Purkinje cell  
234 marker Calbindin-D28k showed no correlation with *SGCE* expression (data not shown).

235

### 236 **Mutation analysis of novel *SGCE* exons**

237 We performed mutation analysis in our cohort of definite M-D patients where no  
238 mutation was identified<sup>7</sup> and sequenced novel alternatively spliced *SGCE* exons. This  
239 included all new in-frame exons (1c, 4a) and frameshift exon 11b with all extended  
240 versions (exon 11d, 11c, 11f). Also, frameshift exons 3b and 3d were sequenced since their  
241 combination was observed in a transcript, and maintains the reading frame. No mutations  
242 were identified in the 12 patients analyzed.

243

### 244 **Discussion**

245 Mutations in *SGCE* lead to M-D, but how mutations in this ubiquitously expressed  
246 gene result only in **neurological impairments** is unknown. In this study, we propose a link  
247 between the cerebellum and the M-D phenotype.

248 M-D is an autosomal-dominant disease with a reduced penetrance due to maternal  
249 imprinting. Monoallelic paternal *SGCE* expression has been shown in human blood  
250 leukocytes and in human and mice brain.<sup>15;18</sup> Several genes are known to show tissue-  
251 specific imprinting<sup>22</sup> and brain cell-type or region-specific imprinting.<sup>23</sup> The latter has not  
252 yet been investigated for *SGCE*, but is important to address, since only regions in which the  
253 imprint is maintained can be involved in the M-D pathogenesis. In this study, we show that  
254 the *SGCE* imprinting pattern is retained in different human brain regions and different  
255 human tissues (muscle, blood) similar to what has been shown in mice.<sup>18</sup> Thus, **the brain-**  
256 **specific** M-D phenotype cannot be explained by a brain region-specific or tissue-specific  
257 imprinting. Since not all M-D patients have mutations in *SGCE*, silencing of *SGCE*  
258 expression due to imprinting defects could also be a possible disease mechanism. We tested  
259 our cohort of patients in which no mutation was identified and found a normal imprint in  
260 blood. Imprinting defects in the brain cannot be excluded, however.

261 **More likely explanations for the brain-specific symptoms in M-D are** the presence  
262 of a brain-specific *SGCE* isoform with a brain-specific function or **a brain-specific function**  
263 **for ubiquitous *SGCE***. Our analysis of *SGCE* expression in human brain using ultra deep  
264 amplicon sequencing shows the presence of 23 alternatively spliced exons, of which 19 at  
265 very low frequencies (Table 2). Five of the 23 events retained an **open reading frame** and  
266 are expected to be translated (exon 1c, 2, 4a, 8, and 10). Exon 1c and 4a have not been  
267 identified thus far. The majority of the new splice variants have splice sites implying

268 involvement of the splicing machinery. However, most of the new isoforms lead to a  
269 frameshift and a premature stop codon which makes them likely a target for **nonsense**  
270 **mediated decay**. Despite the overall low frequency of all new exons, we cannot exclude  
271 that they contribute to the physiological function of SGCE.

272         Of the four major alternatively spliced exons (exon 1c, 2, 8, 11b), exon 11b was of  
273 most interest with respect to the disease: It has a brain-specific expression pattern, was  
274 highly conserved and expressed at high levels. **Exon 8 expression was low in brain (7.2%)**  
275 **compared to heart and muscle (96%), but its expression in brain appears to be variable.** We  
276 further analyzed expression of SGCE exon 11b throughout the human brain. We identified  
277 differential expression of exon 11b in human brain areas, whereas total SGCE expression  
278 levels did not differ. Quantitative PCR analysis confirmed this observation and revealed  
279 trends: A consistently high expression in the cerebellum, moderate to high expression in  
280 SM1, moderate expression in substantia nigra and putamen and consistently low levels in  
281 globus pallidus. We observed a high interindividual variability of exon 11b expression in  
282 the cerebellum. Interindividual differences in expression of isoforms could account for this  
283 observation<sup>24;25</sup>.

284         We show that exon 11b is the major brain-specific alternatively spliced SGCE exon  
285 and that it is differentially expressed in five control subjects. Its expression pattern is  
286 compatible with a role for the brain-specific SGCE isoform in the cerebellum and a link to  
287 M-D. This hypothesis is supported by recent literature: (1) there is a shift from dysfunction  
288 of the basal ganglia towards involvement of the cerebellum in dystonia pathogenesis;<sup>26</sup> (2)  
289 several animal models linked dystonia to abnormal cerebellar signaling and cerebellar  
290 defects, and cerebellar lesions leading to dystonia have been observed in patients;<sup>26;27</sup> (3)

291 fMRI analysis of genetically confirmed M-D patients suggests involvement of thalamus  
292 and cerebellum (dentate nucleus)<sup>28</sup> and involvement of different cortical areas and  
293 cerebellum.<sup>29</sup>

294 A role of the cerebellum in M-D is in contrast with striatal changes on dopamine  
295 levels in SGCE deficient mice and a reduced striatal D2 receptor binding in M-D  
296 patients.<sup>30;31</sup> Moderate expression levels of the major brain-specific *SGCE* isoform in the  
297 striatum and lowest expression levels in the globus pallidus were found in all control  
298 subjects. We propose that observed striatal changes may be secondary due to abnormal  
299 cerebellar signaling. This is supported by an anatomical link between the cerebellum and  
300 the basal ganglia, namely projections from the dentate nucleus to the striatum,<sup>32</sup> and by  
301 cerebellar lesions or stimulations that have been shown to alter striatal dopamine  
302 signaling.<sup>33-35</sup> Our isoform-specific ISH showed that the major brain-specific *SGCE*  
303 transcript is highly expressed in Purkinje cells and neurons of the dentate nucleus in the  
304 cerebellum. No signals were detected in the cerebellar granule cell layer in the cerebellum  
305 as it has been shown in mouse brain.<sup>17</sup> Interestingly, alcohol consumption alleviates the  
306 symptoms in M-D and the cerebellum reacts very sensitively to alcohol.<sup>36</sup> Also, symptoms  
307 of patients with essential tremor, a common neurological movement disorder with  
308 suggested involvement of the cerebellum,<sup>37-39</sup> are known to respond to alcohol as seen in  
309 M-D.

310 We propose that the loss of function of the brain-specific *SGCE* isoform underlies  
311 the exclusively neurological M-D phenotype. We hypothesize that the general function of  
312 the abundant *SGCE* isoform is redundant or not essential since no non-neurological

313 symptoms have been reported. The brain-specific protein may have a unique function and  
 314 cannot be replaced.

### 315 **Fundings**

316 This study was supported by Nederlandse Organisatie voor Wetenschappelijk  
 317 Onderzoek (NWO) VIDI (project 0160.056.333) and by the Prinses Beatrix Fund.

318

### 319 **Acknowledgements**

320 We thank Edward J Bradley for assistance with 454 sequencing and Dr Kees Fluiters  
 321 for help with ISHs. We thank Dr. Anneloor ten Asbroek for helpful discussion and critical  
 322 reading of the manuscript. Tissue from the M-D mutation carrier was obtained from the  
 323 Brain and Tissue Bank for Developmental Disorders at the University of Maryland under  
 324 contracts N01-HD-4-3368 and N01-HD-4-3383.

325

326 **Conflict of interest.** The authors declare no conflict of interest.

327 **Supplementary information is available at EJHG's website.**

328

### 329 **Reference List**

330

331 1 Klein C: Myoclonus and myoclonus-dystonias. *In Pulst SM, ed Genetics of Movement*  
 332 *Disorders* 2003; San Diego, USA: Academic Press/Elsevier Science. 449-469.

333 2 Nardocci N, Zorzi G, Barzaghi C *et al*: Myoclonus-dystonia syndrome: clinical  
 334 presentation, disease course, and genetic features in 11 families. *Mov Disord* 2008;  
 335 **23**: 28-34.

336 3 Marelli C, Canafoglia L, Zibordi F *et al*: A Neurophysiological Study of Myoclonus  
 337 in Patients with DYT11 Myoclonus-Dystonia Syndrome. *Mov Disord* 2008; **23**:  
 338 2041-2048.

339 4 Breakefield XO, Blood AJ, Li Y, Hallett M, Hanson PI, Standaert DG: The  
 340 pathophysiological basis of dystonias. *Nat Rev Neurosci* 2008; **9**: 222-234.

- 341 5 Zimprich A, Grabowski M, Asmus F *et al*: Mutations in the gene encoding epsilon-  
342 sarcoglycan cause myoclonus-dystonia syndrome. *Nat Genet* 2001; **29**: 66-69.
- 343 6 Grunewald A, Djarmati A, Lohmann-Hedrich K *et al*: Myoclonus-dystonia:  
344 significance of large SGCE deletions. *Hum Mutat* 2008; **29**: 331-332.
- 345 7 Ritz K, Gerrits MCF, Foncke EMJ *et al*: Myoclonus-dystonia: clinical and genetic  
346 evaluation of a large cohort. *J Neurol Neurosurg Psychiatr* 2009; **80**: 653-658.
- 347 8 Grimes DA, Han F, Lang AE, George-Hyssop P, Racacho L, Bulman DE: A novel  
348 locus for inherited myoclonus-dystonia on 18p11. *Neurology* 2002; **59**: 1183-1186.
- 349 9 Han F, Racacho L, Lang AE, Bulman DE, Grimes DA: Refinement of the DYT15  
350 locus in myoclonus dystonia. *Mov Disord* 2007; **22**: 888-892.
- 351 10 Ozawa E, Mizuno Y, Hagiwara Y, Sasaoka T, Yoshida M: Molecular and cell  
352 biology of the sarcoglycan complex. *Muscle & Nerve* 2005; **32**: 563-576.
- 353 11 Waite A, Tinsley CL, Locke M, Blake DJ: The neurobiology of the dystrophin-  
354 associated glycoprotein complex. *Ann Med* 2009; **41**: 344-359.
- 355 12 McNally EM, Ly CT, Kunkel LM: Human epsilon-sarcoglycan is highly related to  
356 alpha-sarcoglycan (adhalin), the limb girdle muscular dystrophy 2D gene. *FEBS Lett*  
357 1998; **422**: 27-32.
- 358 13 Hjermind LE, Vissing J, Asmus F *et al*: No muscle involvement in myoclonus-  
359 dystonia caused by epsilon-sarcoglycan gene mutations. *Eur J Neurol* 2008; **15**: 525-  
360 529.
- 361 14 Muller B, Hedrich K, Kock N *et al*: Evidence that paternal expression of the epsilon-  
362 Sarcoglycan gene accounts for reduced penetrance in myoclonus-dystonia. *Am J Hum*  
363 *Genet* 2002; **71**: 1303-1311.
- 364 15 Grabowski M, Zimprich A, Lorenz-Depiereux B *et al*: The epsilon-sarcoglycan gene  
365 (SGCE), mutated in myoclonus-dystonia syndrome, is maternally imprinted. *Eur J*  
366 *Hum Genet* 2003; **11**: 138-144.
- 367 16 Esapa CT, Waite A, Locke M *et al*: SGCE missense mutations that cause myoclonus-  
368 dystonia syndrome impair epsilon-sarcoglycan trafficking to the plasma membrane:  
369 modulation by ubiquitination and torsinA. *Hum Mol Genet* 2007; **16**: 327-342.
- 370 17 Nishiyama A, Endo T, Takeda S, Imamura M: Identification and characterization of  
371 epsilon-sarcoglycans in the central nervous system. *Mol Brain Res* 2004; **125**: 1-12.
- 372 18 Yokoi F, Dang MT, Mitsui S, Li Y: Exclusive paternal expression and novel  
373 alternatively spliced variants of epsilon-sarcoglycan mRNA in mouse brain. *FEBS*  
374 *Lett* 2005; **579**: 4822-4828.

- 375 19 Ruijter JM, Ramakers C, Hoogaars WM *et al*: Amplification efficiency: linking  
376 baseline and bias in the analysis of quantitative PCR data. *Nucleic Acids Res* 2009;  
377 **37**: e45.
- 378 20 Budde BS, Namavar Y, Barth PG *et al*: tRNA splicing endonuclease mutations cause  
379 pontocerebellar hypoplasia. *Nat Genet* 2008; **40**: 1113-1118.
- 380 21 Maquat LE: Nonsense-mediated mRNA decay: splicing, translation and mRNP  
381 dynamics. *Nat Rev Mol Cell Biol* 2004; **5**: 89-99.
- 382 22 Albrecht U, Sutcliffe JS, Cattanach BM *et al*: Imprinted expression of the murine  
383 Angelman syndrome gene, Ube3a, in hippocampal and Purkinje neurons. *Nat Genet*  
384 1997; **17**: 75-78.
- 385 23 Yamasaki K, Joh K, Ohta T *et al*: Neurons but not glial cells show reciprocal  
386 imprinting of sense and antisense transcripts of Ube3a. *Hum Mol Genet* 2003; **12**:  
387 837-847.
- 388 24 Wang ET, Sandberg R, Luo SJ *et al*: Alternative isoform regulation in human tissue  
389 transcriptomes. *Nature* 2008; **456**: 470-476.
- 390 25 Kwan T, Benovoy D, Dias C *et al*: Genome-wide analysis of transcript isoform  
391 variation in humans. *Nat Genet* 2008; **40**: 225-231.
- 392 26 Jinnah HA, Hess EJ: A new twist on the anatomy of dystonia: the basal ganglia and  
393 the cerebellum? *Neurology* 2006; **67**: 1740-1741.
- 394 27 Pizoli CE, Jinnah HA, Billingsley ML, Hess EJ: Abnormal cerebellar signaling  
395 induces dystonia in mice. *J Neurosci* 2002; **22**: 7825-7833.
- 396 28 Nitschke MF, Erdmann C, Trillenber P *et al*: Functional MRI reveals activation of a  
397 subcortical network in a 5-year-old girl with genetically confirmed myoclonus-  
398 dystonia. *Neuropediatrics* 2006; **37**: 79-82.
- 399 29 Beukers RJ, Foncke EM, van der Meer JN *et al*: Disorganized sensorimotor  
400 integration in mutation-positive myoclonus-dystonia: a functional magnetic resonance  
401 imaging study. *Arch Neurol* 2010; **67**: 469-474.
- 402 30 Yokoi F, Dang MT, Li J, Li Y: Myoclonus, motor deficits, alterations in emotional  
403 responses and monoamine metabolism in epsilon-sarcoglycan deficient mice. *J*  
404 *Biochem* 2006; **140**: 141-146.
- 405 31 Beukers RJ, Booij J, Weisscher N, Zijlstra F, van Amelsvoort TA, Tijssen MA:  
406 Reduced striatal D2 receptor binding in myoclonus-dystonia. *Eur J Nucl Med Mol*  
407 *Imaging* 2009; **36**: 269-274.
- 408 32 Hoshi E, Tremblay L, Feger J, Carras PL, Strick PL: The cerebellum communicates  
409 with the basal ganglia. *Nat Neurosci* 2005; **8**: 1491-1493.

- 410 33 Nieoullon A, Cheramy A, Glowinski J: Release of dopamine in both caudate nuclei  
411 and both substantia nigrae in response to unilateral stimulation of cerebellar nuclei in  
412 the cat. *Brain Res* 1978; **148**: 143-152.
- 413 34 Tranchant C, Maquet J, Eber AM, Dietemann JL, Franck P, Warter JM: Cerebellar  
414 cavernous angioma, cervical dystonia and crossed cortical diaschisis. *Rev Neurol*  
415 1991; **147**: 599-602.
- 416 35 Neychev VK, Fan X, Mitev VI, Hess EJ, Jinnah HA: The basal ganglia and  
417 cerebellum interact in the expression of dystonic movement. *Brain* 2008; **131**: 2499-  
418 2509.
- 419 36 Volkow ND, Mullani N, Gould L *et al*: Effects of acute alcohol intoxication on  
420 cerebral blood flow measured with PET. *Psychiatry Res* 1988; **24**: 201-209.
- 421 37 Dupuis MJ, Delwaide PJ, Boucquey D, Gonsette RE: Homolateral disappearance of  
422 essential tremor after cerebellar stroke. *Mov Disord* 1989; **4**: 183-187.
- 423 38 Louis ED, Vonsattel JP, Honig LS *et al*: Essential tremor associated with pathologic  
424 changes in the cerebellum. *Arch Neurol* 2006; **63**: 1189-1193.
- 425 39 Louis ED: Essential tremor: evolving clinicopathological concepts in an era of  
426 intensive post-mortem enquiry. *Lancet Neurol* 2010; **9**: 613-622.  
427

428 **Titles and legends to figures**

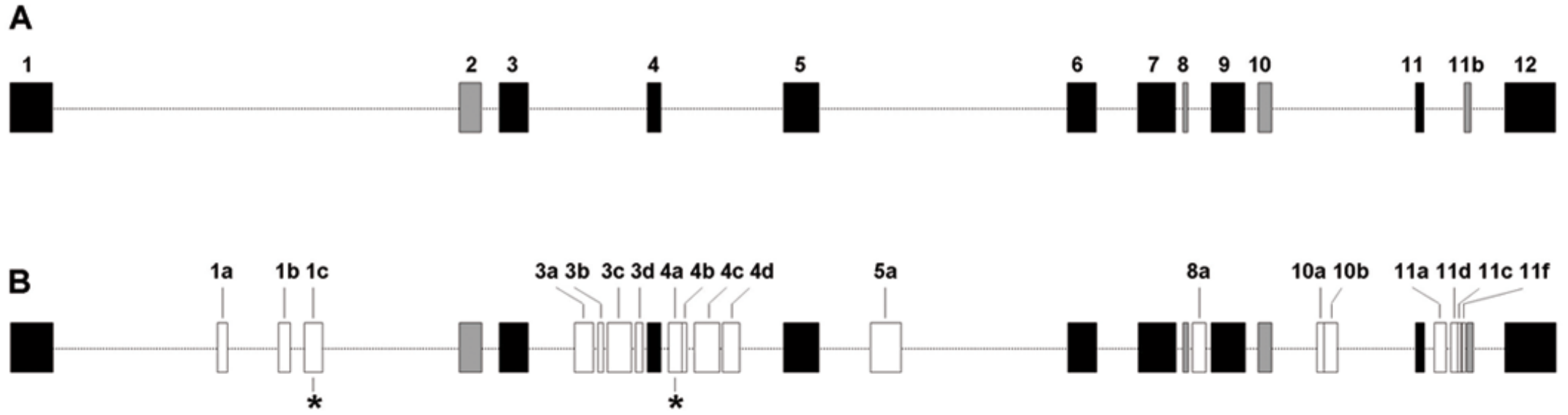
429 **Figure 1: Overview of novel alternatively spliced exons identified by 454 ultra deep**  
430 **sequencing.** A) *SGCE* gene including all known exons. Constitutive exons are depicted in  
431 black, alternatively spliced exons in grey. B) *SGCE* gene including novel alternatively  
432 spliced exons (in white). Asterisks represent novel in-frame exons. In-frame exons were  
433 observed with a higher frequency (range: 0.3-5.7%) than frameshift exons (range: 0.01-  
434 3%). Relative position of the exons and sizes of exons are correct, but exon sizes are not  
435 represented in the same scale as the introns.

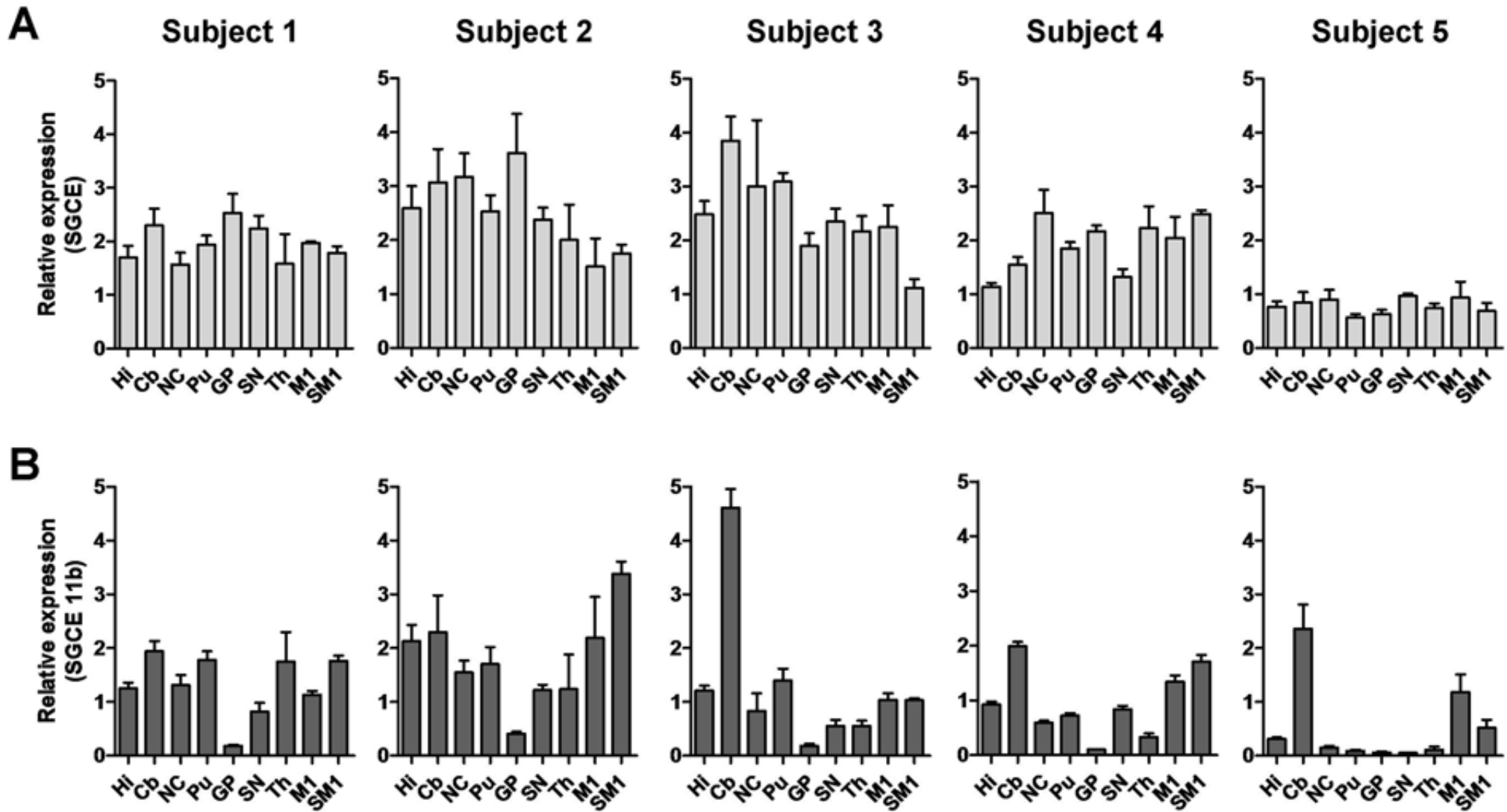
436

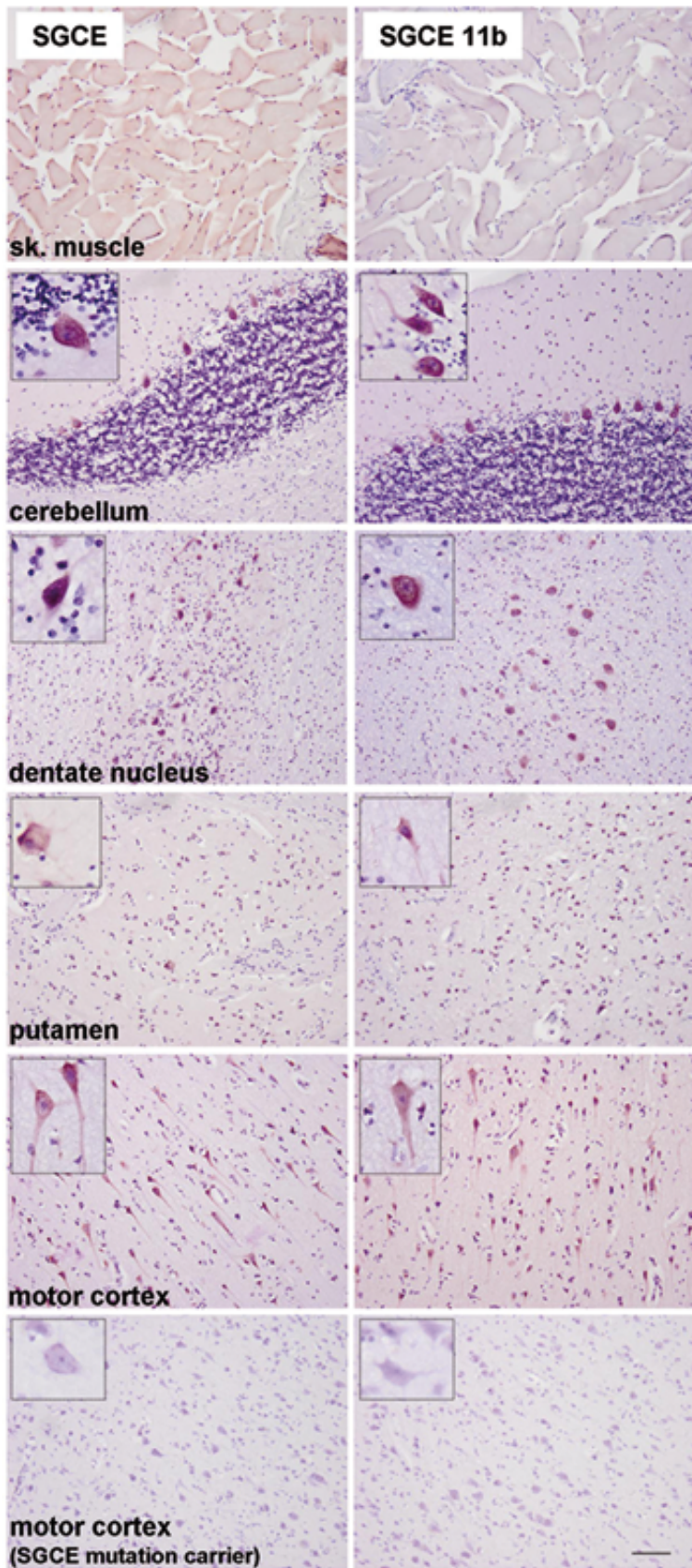
437 **Figure 2: *SGCE* expression levels (qPCR).** Displayed are expression levels for total  
438 *SGCE* (panel A) and *SGCE* exon 11b (panel B) transcripts of the five tested control  
439 subjects normalized to GAPDH. Y-axis shows *SGCE* or *SGCE11b/GAPDH* \*100.  
440 Abbreviations: Hi: Hippocampus, Cb: Cerebellum, NC: Caudate Nucleus, Pu: Putamen,  
441 GP: Globus Pallidus, SN: Substantia Nigra, Th: Thalamus, M1: Primary Motor Cortex,  
442 SM1: Primary Somatosensory Cortex.

443

444 **Figure 3: Localization of total *SGCE* transcripts and those containing exon 11b.**  
445 Depicted are ISH results for total *SGCE* and brain-specific *SGCE* in different human brain  
446 regions and human skeletal muscle. *SGCE* showed neuronal staining in all brain regions  
447 tested as well as staining in muscle. Brain-specific *SGCE* transcripts containing exon 11b  
448 showed the same expression pattern in brain but were not expressed in skeletal muscle. No  
449 *SGCE* mRNA was detected in the motor cortex of an M-D patient carrying a *SGCE*  
450 nonsense mutation. Inserts in each panel show high magnification of neurons. Scale bar  
451 represents 100 $\mu$ m.







## Tables 1-3

**Table 1: Sample material**

Subject	Gender	Age (years)	Post-mortem (h)	Brain tissue									Other		Assay	
1	male	77	7	Cb	GP	Hi	M1	NC	Pu	SM1	SN	Th	He	MP	454	qPCR
2	male	76	8	Cb	GP	Hi	M1	NC	Pu	SM1	SN	Th	na		454	qPCR
3	female	76	5	na									He	MP	454	
4	male	74	5	Cb	GP	Hi	M1	NC	Pu	SM1	SN	Th	na			qPCR
5	female	88	12	Cb	GP	Hi	M1	NC	Pu	SM1	SN	Th	na			qPCR
6	female	45	24	Cb	GP	Hi	M1	NC	Pu	SM1	SN	Th	na			qPCR

Abbreviations: Cb: Cerebellum, GP: Globus Pallidus, Hi: Hippocampus, M1: Primary Motor Cortex, NC: Caudate Nucleus, Pu: Putamen, SM1: Primary Somatosensory Cortex, SN: Substantia Nigra, Th: Thalamus, He: Heart, MP: M.Psoas, na: not available

**Table 2: Known and new alternatively spliced *SGCE* exons revealed by screening the entire *SGCE* gene in one control subject with ultra deep sequencing.** Additional information of alternatively spliced *SGCE* exons include: occurrence of respective exons in brain (SM1), heart and blood of control subject 1 in percent, splice sites according to [www.fruitfly.org](http://www.fruitfly.org), NNSPLICE 0.9 version (minimum score for 3'/5' splice site 0.5): splice donor (D)/acceptor (A), frameshift (FS) or in-frame (IF) exons, isoform subject to nonsense mediated decay (NMD), exon size in bp, nucleotide position according to NG\_008893.1. Constitutive exons (100% expression) are not shown. \*: Exons that were sequenced in definite M-D patients where no mutation could be identified.

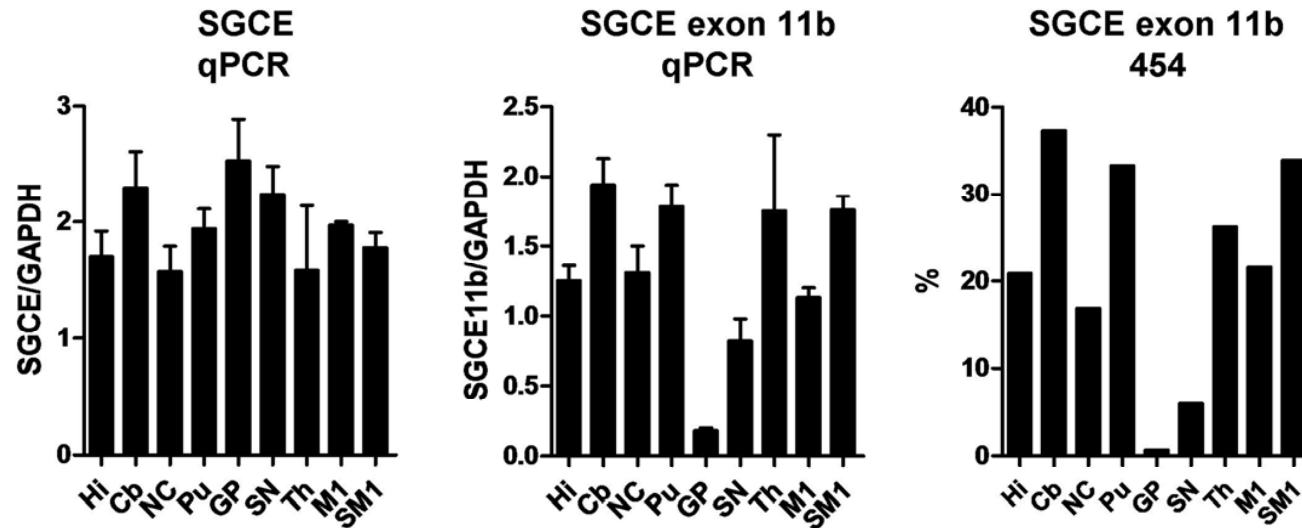
Alternatively spliced exons	Exon	Brain (%)	Heart (%)	Blood (%)	Splice site		FS/IF	NMD	Exon size (Bp)	NG_008893.1	
					D	A				Start	End
<b>known</b>	<b>2*</b>	98.20	94.80	78.50	+	+	IF	-	123	31369	31491
	<b>8*</b>	7.20	96.00	96.00	+	+	IF	-	27	61487	61513
	<b>10*</b>	0.03	0.03	1.25	-	+	IF	-	75	63206	63280
	<b>11b*</b>	33.92	0.05	-	-	-	FS	-	35	73398	73432
<b>new</b>	<b>1a</b>	0.66	0.31	1.08	+	+	FS	+	62	16543	16604
	<b>1b</b>	0.03	0.11	-	+	+	FS	+	68	20246	20313
	<b>1c*</b>	5.70	1.96	2.25	+	-	IF	-	108	21762	21869
	<b>3a</b>	-	0.03	-	+	-	FS	+	102	36163	36264
	<b>3b*</b>	0.43	0.02	-	+	+	FS	+	31	36961	36991
	<b>3c</b>	0.05	-	-	+	+	FS	+	130	37052	37181
	<b>3d*</b>	-	0.17	-	+	-	FS	+	35	37655	37689
	<b>4a*</b>	0.80	0.94	1.66	+	+	IF	-	81	38584	38664
	<b>4b</b>	0.14	0.37	-	-	+	FS	+	110	38584	38693
	<b>4c</b>	0.07	0.12	-	+	+	FS	+	147	39215	39361
	<b>4d</b>	-	0.01	-	+	+	FS	+	94	39504	39597
	<b>5a</b>	-	-	0.06	-	+	FS	+	176	45589	45764
	<b>8a</b>	0.08	0.11	0.51	+	-	FS	+	74	61810	61883
	<b>10a</b>	0.01	-	-	-	+	FS	+	76	66703	66778
	<b>10b</b>	0.02	-	-	+	+	FS	+	115	66664	66778
	<b>11a</b>	0.04	-	-	+	-	FS	-	68	73177	73244
	<b>11d*</b>	0.03	-	0.12	-	-	FS	-	107	73326	73432
<b>11c*</b>	0.70	0.28	0.41	-	-	FS	-	72	73361	73432	
<b>11f*</b>	0.33	-	-	-	-	FS	-	51	73382	73432	

**Table 3: Expression levels of alternatively spliced *SGCE* exon 11b in human brain regions of two control subjects (454 sequencing).**

Expression levels of exon 11b are given as percentage of all analyzed sequence reads obtained by 454 sequencing. Abbreviations: HS: Homo sapiens, RN: Rattus norvegicus, MM: Mus musculus, Hi: Hippocampus, Cb: Cerebellum, NC: Caudate Nucleus, Pu: Putamen, GP: Globus Pallidus, SN: Substantia Nigra, Th: Thalamus, M1: Primary Motor Cortex, SM1: Primary Somatosensory Cortex.

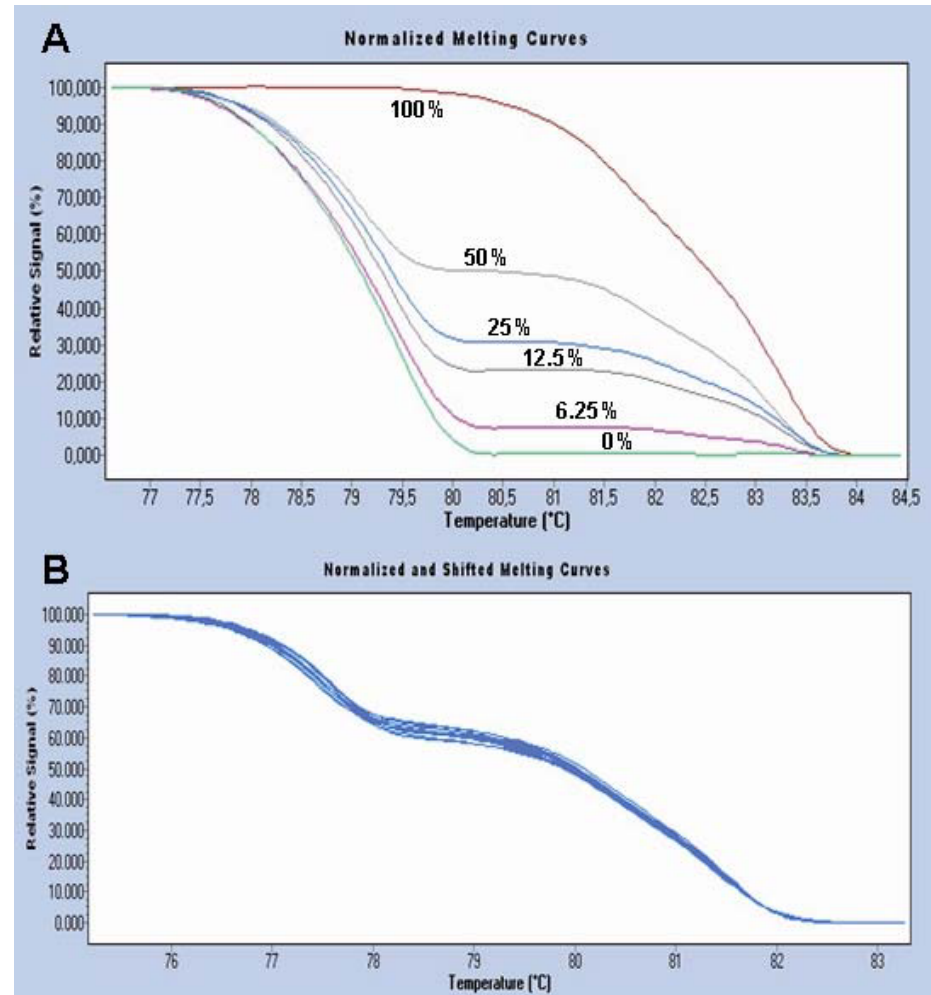
		<b>Exon 11b</b>	
	<b>Region</b>	<b>Subject 1 (%)</b>	<b>Subject 2 (%)</b>
<b>HS</b>	<b>Hi</b>	20.84	16.59
	<b>Cb</b>	37.31	20.99
	<b>NC</b>	16.89	12.48
	<b>Pu</b>	33.29	11.63
	<b>GP</b>	0.68	4.09
	<b>SN</b>	5.98	12.60
	<b>Th</b>	26.22	16.71
	<b>M1</b>	21.58	25.51
	<b>SM1</b>	33.92	36.79
<b>RN</b>	<b>Brain</b>	73.25	
<b>MM</b>	<b>Brain</b>	49.80	
<b>ZF</b>	<b>Embryo</b>	20.38	

## Supplementary Fig. 1



**Comparison of qPCR and 454 sequencing expression data.** Displayed are expression levels of overall *SGCE* and *SGCE* exon 11b transcripts obtained by qPCR and 454 sequencing of a control subject. qPCR results are normalized to *GAPDH* (mean and standard deviation of triplicates). 454 sequencing results show expression levels of transcripts containing exon 11b given as percentage of all sequence reads analyzed. (Hi Hippocampus, Cb Cerebellum, NC Caudate Nucleus, Pu Putamen, GP Globus Pallidus, SN Substantia Nigra, Th Thalamus, M1 Primary Motor Cortex, SM1 Primary Somatosensory Cortex)

## Supplementary Fig. 2



### ***SGCE* imprinting pattern by MS-HRM assay.**

A) Samples with different degrees of methylation of the *SGCE* promoter region were included as standards in each experiment : 100% methylated and unmethylated controls and 50%, 25%, 12.5% and 6.25% methylated samples.

B) Example of melting curves of different brain regions and blood samples all showing a normal imprinting pattern with 50% methylation.

## Supplementary Tables 1-2

**Supplementary Table 1: Number of total sequence reads analyzed with 454 sequencing.** The numbers of sequence reads analyzed are given as average (range) of all sequenced samples (brain regions/tissues) or as absolute numbers in case there were one or two regions sequenced only.

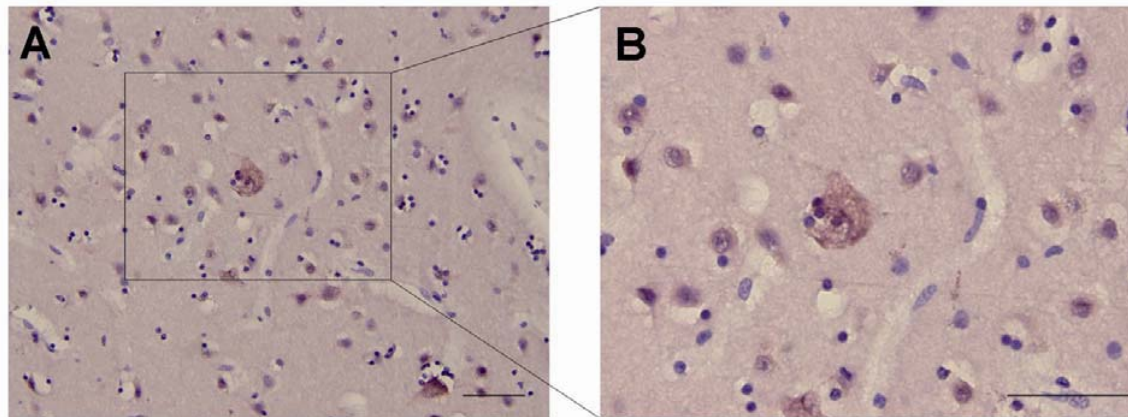
	# Samples	Number of reads analyzed
<b>H.sapiens Subject 1 (Brain, Muscle)</b>	10	15 106 (12 643-17 916)
<b>H.sapiens Subject 2 (Brain)</b>	9	7 084 (4 535-11 677)
<b>H.sapiens Subject 3 (Muscle)</b>	1	6 409
<b>H.sapiens Blood</b>	2	12 423, 13 786
<b>R.norvegicus Brain</b>	1	10 435
<b>M.musculus Brain</b>	1	8 213
<b>D.rerio Embryos</b>	1	14 382

**Supplementary Table 2: Conservation of alternatively spliced *SGCE* exons.**

Conservation was investigated using the “blast assembled genomes” feature and the BLASTn algorithm (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). Numbers indicate the percentage of identical nucleotides relative to the size of the exon in humans.

Exon	1a	1b	1c	2	3a	3b	3c	3d	4a	4b	4c	4d	5a	8	8a	10	10a	10b	11	11a	11d	11c	11f	11b
<b>P.troglodytes</b>	100	100	100	100	100	100	99	100	100	100	99	98	99	100	100	94	97	98	100	98	98	100	100	100
<b>P.pygmaeus</b>	94	97	100	100	98	100	99	100	97	97	97	95	99	100	98	-	97	98	100	97	97	98	100	100
<b>M.mulatta</b>	97	93	100	99	97	100	96	94	95	95	95	96	97	100	100	-	96	96	100	98	98	100	100	100
<b>B.taurus</b>	80	71	96	93	72	-	86	-	89	89	49	-	-	100	93	-	70	54	100	93	92	97	98	100
<b>R.norvegicus</b>	-	-	-	86	-	-	-	-	-	-	-	-	-	96	77	-	42	-	95	89	93	94	98	100
<b>M.musculus</b>	-	-	91	88	75	-	-	-	-	-	-	-	16	92	68	-	47	-	90	91	93	92	96	100
<b>G.gallus</b>	-	-	82	85	-	-	-	-	-	-	25	-	-	100	60	-	-	-	88	-	67	81	88	91
<b>D.rerio</b>	-	-	-	63	-	-	-	-	-	-	-	-	-	96	-	-	-	-	68	-	-	-	-	63

### Optional figure 1:



#### **Localization of brain-specific SGCE in small and large striatal neurons.**

Higher magnifications of ISHs for detection of SGCE exon 11b transcripts as an example to illustrate neuronal staining in the putamen. Scale bar represents 50 $\mu$ m.