**Supplementary Material**

***AP4S1* splice-site mutation in a case of spastic paraplegia type 52 with polymicrogyria**

Susana Carmona, PhD, Clara Marecos, MD, Marta Amorim, MD, Ana Cristina Ferreira, MD, Carla Conceição, MD, José Brás, PhD, Sofia Temudo Duarte, MD, PhD, Rita Guerreiro, PhD

**Material and Methods**

**Exome sequencing**

Genomic DNA samples from the proband and both parents were subjected to whole-exome sequencing. Capture was performed using Agilent SureSelect v4 and sequencing was done on Illumina’s HiSeq 4000 according to manufacturer’s instructions. After sequencing, reads were aligned to the hg19/GRCh37 reference genome using BWA v0.7.12[1](https://paperpile.com/c/HAAJVU/sK9W) and variants were called using GATK best practices v3.3-0[2](https://paperpile.com/c/HAAJVU/kpdj).

***In silico* analyses**

SNPs and Indels were filtered to select rare exonic and splice-site variants predicted as deleterious. This process was performed using Exomiser v7.2.1[3](https://paperpile.com/c/HAAJVU/P2vw) with the following parameters: autosomal recessive inheritance pattern, minor allele frequency (MAF) inferior to 1%, and HP:0001263 (global developmental delay), HP:0009062 (infantile axial hypotonia), HP:0001250 (seizures), HP:0012650 (perisylvian polymicrogyria), HP:0002079 (hypoplasia of the corpus callosum) and HP:0002119 (ventriculomegaly) human phenotype ontology (HPO) IDs. The annotated variants were also compared to the genome genotyping results and variants located in large regions of loss of heterozygosity (LOH) were prioritized.

The functional predicted impact of missense and nonsense variants was evaluated using SIFT[4](https://paperpile.com/c/HAAJVU/SXS4), PolyPhen-2[5](https://paperpile.com/c/HAAJVU/S6fM), MutationTaster2[6](https://paperpile.com/c/HAAJVU/Pbjm) and CADD v1.3[7](https://paperpile.com/c/HAAJVU/Nv2o) software. *In silico* analysis of the splice-site variants was performed using Human Splicing Finder (HSF) v3.1[8](https://paperpile.com/c/HAAJVU/oQpM), MaxEntScan[9](https://paperpile.com/c/HAAJVU/Pg0I), Splice Site Prediction by Neural Network (NNsplice) v0.9[10](https://paperpile.com/c/HAAJVU/IV4m) and SpliceView[11](https://paperpile.com/c/HAAJVU/yb9T) software. Scores obtained with the wild-type allele were compared with the scores generated in the presence of the variant. ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) was also used in the variants’ classification process.

**Genome-wide genotyping**

The proband’s genome was genotyped using Illumina Infinium technology to identify the presence of any large structural variants (>50 Kb) and large regions of homozygosity (>1 Mb). The sample was genotyped using the HumanOmniExpress BeadChip according manufacturer’s instructions and data were visualized using the GenomeStudio Data Analysis Software (Illumina Inc.).

**Transcript size and sequence analysis**

Total RNA from the proband, both parents and a wild-type individual was extracted from peripheral blood. After RNA integrity evaluation and quantification, 1 µg of total RNA was converted into cDNA using High Capacity RNA-to-cDNA Kit (Thermo Fisher Scientific) according manufacturer’s instructions. A fragment including the end of exon 2, exons 3, 4 and 5 and the beginning of exon 6 was amplified by polymerase chain reaction (PCR) using 1x FastStart PCR Master (Roche Diagnostics Corp), 0.4 µM of forward primer 5’-CTCGATCCAATGAACAATGCTCT-3’, 0.4 µM of reverse primer 5’-TCAACAGAATTTAAAGTGCGACAG-3’ and 2.2 ng/µl of cDNA in the following conditions: an initial denaturation step at 94°C for 5 min, followed by 30 cycles including initial denaturation at 94°C for 15 s, followed by annealing at 58°C for 25 s and extension at 72°C for 30 s, and final extension step at 72°C for 5 min. PCR products of all samples were run in a 1.5% agarose gel. Transcript sizes were compared between the proband, both parents and the wild-type individual.

Amplicons of all samples were also sequenced with Applied Biosystems BigDye terminator version 3.1 sequencing chemistry in an ABI3730XL genetic analyser as per the manufacturer’s instructions (Applied Biosystems). The sequences were analysed using Sequencher software version 4.2 (Gene Codes) using the NM\_001128126.2 transcript as a reference. The nucleotide transcript sequence of the proband and the wild-type individual were translated into amino acids using the Translate tool (<https://web.expasy.org/translate/>) and aligned using Clustal Omega v1.2.4[12](https://paperpile.com/c/HAAJVU/0PMw).

**Results**

**Genome genotyping data**

No CNVs predicted as deleterious were found in the patient. Fourteen loss of heterozygosity regions larger than 1 Mb were identified and these included 156 genes. We filtered out all exome sequencing variants not located in these LOH regions. Only the *AP4S1* splice-site variant c.294+1G>T was left after filtering for rare and high impact variants. This variant was also the top hit from the Exomiser analysis. It is located in a 2.4 Mb LOH region in chromosome 14 (LOH between positions 29977571 and 32364165) (Figure e-1).

**Discussion**

When comparing the brain imaging data from our patient with images from Roubertie A et al., there is a difference between patients regarding basal ganglia signal in T2 weighted images: in the images from Roubertie *et al*., globus pallidus is more hypo-intense than the putamen and in our patient globus pallidus is slightly hyperintense in comparison to the putamen[13](https://paperpile.com/c/HAAJVU/90dp). However, in our patient there is a thin hypo-intense strip in the junction of the posterior part of these two portions of the lentiform nucleus, very subtle, and in the current MRI, of uncertain significance. We could speculate that this thin strip represents the beginning of pathological brain iron accumulation, but in Roubertie *et al.* the hyposignal is more pronounced in the medial part of the globus pallidus and not in the lateral part, what could suggest that the medial part is where the iron deposition begins. Thus, we think that a follow up MRI examination would be required to confirm or not these deposits, knowing that the progression of brain iron accumulation is generally slow in most processes. However, a new MRI would require anaesthetic procedures and, at the moment, we consider that there is no significant clinical indication to perform such test.

**References**

1. [Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–1760.](http://paperpile.com/b/HAAJVU/sK9W)

2. [Van der Auwera GA, Carneiro MO, Hartl C, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. Curr Protoc Bioinformatics. 2013;43:11.10.1–33.](http://paperpile.com/b/HAAJVU/kpdj)

3. [Smedley D, Jacobsen JOB, Jäger M, et al. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat Protoc. 2015;10:2004–2015.](http://paperpile.com/b/HAAJVU/P2vw)

4. [Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4:1073–1081.](http://paperpile.com/b/HAAJVU/SXS4)

5. [Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–249.](http://paperpile.com/b/HAAJVU/S6fM)

6. [Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11:361–362.](http://paperpile.com/b/HAAJVU/Pbjm)

7. [Kircher M, Witten DM, Jain P, O’Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46:310–315.](http://paperpile.com/b/HAAJVU/Nv2o)

8. [Desmet F-O, Hamroun D, Lalande M, Collod-Béroud G, Claustres M, Béroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37:e67.](http://paperpile.com/b/HAAJVU/oQpM)

9. [Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol. 2004;11:377–394.](http://paperpile.com/b/HAAJVU/Pg0I)

10. [Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. J Comput Biol. 1997;4:311–323.](http://paperpile.com/b/HAAJVU/IV4m)

11. [Rogozin IB, Milanesi L. Analysis of donor splice sites in different eukaryotic organisms. J Mol Evol. 1997;45:50–59.](http://paperpile.com/b/HAAJVU/yb9T)

12. [Sievers F, Wilm A, Dineen D, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.](http://paperpile.com/b/HAAJVU/0PMw)

13. [Roubertie A, Hieu N, Roux C-J, et al. AP4 deficiency: A novel form of neurodegeneration with brain iron accumulation? Neurol Genet. 2018;4:e217.](http://paperpile.com/b/HAAJVU/90dp)

**Figures**

**Figure e-1 – Whole genome genotyping results.**



Plot of the B allele frequencies and log R ratios values showing the region of loss of heterozygosity in the proband located in chromosome 14:29977571-32364165 bp (green area). This region contains 19 genes, including *AP4S1*.