

Identification of partial SLC20A2 deletions in primary brain calcification using whole-exome sequencing

S. David, Jorge Ferreira, O. Quenez, A. Rovelet-Lecrux, A.-C. Richard, M. Verin, S. Jurici, I. Le Ber, Anne Boland, Jean-François Deleuze, et al.

► To cite this version:

S. David, Jorge Ferreira, O. Quenez, A. Rovelet-Lecrux, A.-C. Richard, et al.. Identification of partial SLC20A2 deletions in primary brain calcification using whole-exome sequencing. *European Journal of Human Genetics*, Nature Publishing Group, 2016, 24 (11), pp.1630–1634. 10.1038/ejhg.2016.50 . hal-01397791

HAL Id: hal-01397791

<https://hal-univ-rennes1.archives-ouvertes.fr/hal-01397791>

Submitted on 26 Jan 2017

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.

Identification of partial *SLC20A2* deletions in primary brain calcification using whole exome sequencing

Stéphanie David^{1,2,*}, Joana Ferreira^{1,2,3,4*}, Olivier Quenez^{1,2,5}, Anne Rovelet-Lecrux^{1,2}, Anne-Claire Richard^{2,5}, Marc Vérin⁶, Snejana Jurici⁷, Isabelle Le Ber^{8,9}, Anne Boland¹⁰, Jean-François Deleuze¹⁰, Thierry Frebourg^{1,2,11}, João Ricardo Mendes de Oliveira^{3,12}, Didier Hannequin^{1,2,5,11,13}, Dominique Champion^{1,2,5,14}, and Gaël Nicolas^{1,2,5,11,§}

¹Inserm U1079, Rouen University, IRIB, Normandy University, Rouen, France

²Normandy Centre for Genomic Medicine and Personalized Medicine, Rouen, France

³Keizo Asami Laboratory, Federal University of Pernambuco, Recife, Brazil

⁴Biological Sciences Graduate Program, Federal University of Pernambuco, Recife, Brazil

⁵CNR-MAJ, Rouen University Hospital, Rouen, France

⁶Department of Neurology, Pontchaillou Hospital, Rennes University Hospital, Rennes, France

⁷Department of Neurology, Perpignan Hospital, Perpignan, France

⁸Sorbonne Universités, UPMC Univ Paris 06 ; Inserm UMR S 1127 ; CNRS UMR 7225 ; ICM F-75013, Paris, France

⁹AP-HP, Hôpital de la Pitié-Salpêtrière, Centre de Référence des Démences Rares, Paris, France; AP-

HP, Hôpital de la Pitié-Salpêtrière, Département des maladies du système nerveux, Paris, France

¹⁰Centre National de Génotypage, Institut de Génomique, CEA, Evry, France

¹¹Department of Genetics, Rouen University Hospital, Rouen, France

¹²Neuropsychiatry Department, Federal University of Pernambuco, Recife, Brazil

¹³Department of Neurology, Rouen University Hospital, Rouen, France

¹⁴Department of Research, Rouvray Psychiatric Hospital, Sotteville-lès-Rouen, France

*The authors contributed equally to the work

§corresponding author: Gaël Nicolas, Inserm U1079, faculté de médecine, 22, boulevard Gambetta,

76183 Rouen Cedex, Tel.: +33235148280, Fax: +33235148237

Running title: *SLC20A2* deletions cause brain calcification

The authors have no conflicts of interest to declare.

ABSTRACT

Primary Brain Calcification (PBC) is a dominantly-inherited calcifying disorder of the brain. *SLC20A2* loss of function variants account for the majority of families. Only one genomic deletion encompassing *SLC20A2* and 6 other genes has been reported.

We performed whole exome sequencing (WES) in 24 unrelated French patients with PBC, negatively screened for sequence variant in the known genes *SLC20A2*, *PDGFB*, *PDGFRB* and *XPR1*. We used the CANOES tool to detect copy number variations (CNVs). We detected 2 deletions of exon 2 of *SLC20A2* in two unrelated patients, which segregated with PBC in one family. We then reanalyzed the same sample using a QMPSF assay including one amplicon in each exon of *SLC20A2* and detected two supplemental partial deletions in two patients: one deletion of exon 4 and one deletion of exons 4 and 5. These deletions were missed by the first screening step of CANOES but could finally be detected after readjustment of bioinformatic parameters and use of a genotyping step of CANOES.

This study reports the first partial deletions of *SLC20A2* and strengthens its position as the major PBC-causative gene. It is possible to detect short CNVs from WES data, although the sensitivity of such tools should be evaluated in comparison with other methods.

KEYWORDS

Primary Familial Brain Calcification, *SLC20A2*, deletion, CNV, exome, CANOES, Fahr disease

INTRODUCTION

Primary Brain Calcification (PBC), also known as Primary Familial Brain Calcification (PFBC, OMIM 213600) or formerly Fahr's disease, is defined by the presence of microvascular calcification affecting at least the basal ganglia with no cause following an extensive clinical, biological, and imaging etiological assessment. Calcification carriers may exhibit diverse neurologic or psychiatric symptoms. Among them, the most frequent three categories are: psychiatric disturbances (mainly mood, psychotic or personality disorders), cognitive impairment (involving mainly memory and executive functions) and movement disorders (mainly extrapyramidal signs)¹. Patients may present other symptoms, such as gait disorder, cerebellar syndrome, dysarthria, and rarely seizures. Cephalalgia and especially migraine are frequent circumstances allowing the identification of brain calcification.¹

PBC is inherited as an autosomal dominant trait. In this context, causative variants have been identified in 4 genes: *SLC20A2* (OMIM 158378)², *PDGFRB* (OMIM 173410)³, *PDGFB* (OMIM 190040)⁴ and *XPR1* (OMIM 605237)⁵. Two Copy Number Variants (CNVs) were previously detected as PBC-causing: a partial deletion of *PDGFB* in one patient⁶ and a deletion encompassing *SLC20A2* entirely and 6 other genes in one family⁷. *SLC20A2* and *XPR1* encode inorganic phosphate transporters, the importer SLC20A2 (also known as PiT2) and the exporter XPR1, respectively. Loss of function of either gene leads to microvascular mural and perivascular calcification involving vascular smooth muscle cells and pericytes. Loss of function variants of *PDGFB* or *PDGFRB* are responsible for an alteration of the blood brain barrier (BBB). Whether the BBB alteration is the direct cause of microvascular calcification or there is other links with inorganic phosphate metabolism remains to be determined⁸. In our French series of PBC patients, more than 50% of probands or sporadic cases do not exhibit any causative variant in one of these genes after Sanger sequencing(unpublished data).

We performed whole exome sequencing (WES) in 24 patients with PBC negatively screened for the four genes. Although WES was developed to detect single nucleotide variants and short indels, we can now take advantage from WES data to detect CNVs. We report here the results of CNV analysis of these four genes in these 24 WES.

MATERIALS AND METHODS

Patients

Patients were included following the previously described criteria⁹. Briefly, probands exhibited (1) at least bilateral lenticular calcification, (2) the total calcification score (TCS) using our own visual rating scale was above the age-specific threshold, and (3) an extensive etiological assessment was normal.

Genetic analyses were performed after written, informed consent. This study was approved by our ethics committee. The entire coding sequence of *SLC20A2*, *PDGFB*, *PDGFRB*, and *XPR1* was assessed by Sanger sequencing.

We included for WES 24 patients with PBC (14 probands of unrelated families, 10 apparently sporadic cases) negatively screened for all four genes.

Whole exome sequencing

Exomes were captured using the Agilent Sureselect All Exons Human V5 Kit (Agilent technologies, Santa Clara, CA, USA). Final libraries were sequenced on a HiSeq2000 with paired ends, 100-bp reads. Reads were mapped to the 1000 Genomes GRCh37 build using BWA 0.7.5a.¹⁰ Picard Tools 1.101 was used to flag duplicate reads. We applied GATK for indel realignment, base quality score recalibration and SNPs and indels discovery using the Haplotype Caller across all samples simultaneously according to GATK 3.3 Best Practices recommendations.¹¹ The joint variant calling file (VCF) was annotated using Annovar (date of version used: 2015/03/22 (<http://www.openbioinformatics.org/annovar/>)).

CNV calling

In order to detect CNVs from WES data, we used the CANOES suite¹². Its algorithms are dedicated to the detection of CNVs using read depth comparisons between different samples. In a first step, we calculated the read depth for each target using bedtools. We merged close targets (less than 30 bp) to reduce the rate of false positives and removed all events localized in known regions of segmental duplication. As a second step, we used the GenotypeCNVs function of CANOES, which allows

determining the probability of a specific event. The objective was to analyze each target intersecting any exon of *SLC20A2*, *PDGFB*, *PDGFRB*, and *XPR1* by considering each target as a putative independent event.

For exon numbering, we considered as exon 1 the first one from 5' to 3' in the c.DNA following the given transcript. The three *SLC20A2* partial deletions were submitted to the LOVD database (<http://databases.lovd.nl/shared/genes/SLC20A2/>) with following variant IDs: 0000089490, 0000089491, and 0000089492.

***SLC20A2* QMPSF assay**

We designed a Quantitative Multiplex PCR of Short Fluorescent Fragments (QMPSF) assay to confirm the presence of the deletion encompassing exon 2 of *SLC20A2* in two unrelated individuals and to screen all other patients. This assay included one control amplicon in *HMBS* and one amplicon in each exon of *SLC20A2* (NM_006749.4) plus one amplicon in alternative (non-coding) exon 1 (NM_001257180.1), which maps between non-coding exon 1 and (coding) exon 2 of transcript NM_006749.4 in the genome. Primers are available in Supplementary Table 1.

RESULTS

Among the 24 WES of patients with PBC, mean depth of coverage was 135x (a mean of 98.97% of targeted bases were covered more than 15x). We extracted all variants mapping to *SLC20A2*, *PDGFB*, *PDGFB*, and *XPR1* genes and confirmed that no single nucleotide variant or short insertion or deletion probably affecting function had been missed by Sanger sequencing.

After exome-wide CNV calling using the CANOES tool, we detected a total of 273 events without any filtration on genes (11.4 per individual on average [range: 6-26], with a mean estimated size of 50.5kb [range 0.099kb - 796.27kb]; 96.7% overlapped known CNVs from the database of genomic variants (<http://dgv.tcag.ca/dgv/app/home>). In total we identified 120 duplications and 153 deletions. Of these, 110 duplications and 134 deletions were identified only once in a single patient i.e. unique CNVs. We then focused on PBC genes and identified a heterozygous deletion of exon 2 of *SLC20A2* in two unrelated probands (ROU 1129 and EXT 383, Fig. 1A and 1B, respectively). We confirmed the presence of both deletions by QMPSF and showed that this deletion only encompassed exon 2 and not exon 1 (NM_006749.4), alternative exon 1 (NM_001257180.1), or exon 3. In family 1129, the presence of the deletion was confirmed in an affected brother and his affected daughter. DNA was neither available for other affected relatives from this family nor for relatives from family EXT 383. ROU 1129 and EXT 383 families both originate from the same region of France. However, pedigree information does not suggest that they are related and we confirmed the absence of close relatedness by estimating identity by descent through Plink software (see Supplementary Note 1). Clinical and imaging findings are summarized in Table 1. A more detailed description of the phenotype is available in Supplementary Note 2.

The sensitivity of CNV detection has been estimated to be 74% in the first report of CANOES using WES data obtained with another capture kit¹². With the aim to assess the presence or absence of CNV in each *SLC20A2* exon, we developed a QMPSF assay including one amplicon in each exon of *SLC20A2*. After assessment of the 22 remaining patients, we detected two new heterozygous deletions: one encompassing exon 4 only (EXT 434, Fig. 1C) and the other one encompassing exons 4 and 5 (EXT 451, Fig. 1D). The deletion of exon 4 was also found in the the proband's affected sister. DNA

from relatives of the patient carrying the exon 4-5 deletion was not available. RNA from these patients was not available. None of the three deletions is reported in the database of genomic variants.

As 2 of the 4 deletions were missed by the first CNV screening of CANOES, we went back to WES data. We used another tool from the CANOES suite, called GenotypeCNVs, which is designed to perform targeted detection of CNVs, to test whether it allowed identifying all four deletions. The GenotypeCNVs tool allowed detecting all four deletions only after decreasing quality thresholds. However, this led to the detection of a false positive duplication of exon 4 of *SLC20A2*.

DISCUSSION

SLC20A2 is the major gene causing PBC¹³. It is clearly demonstrated that heterozygous loss of *SLC20A2* function is sufficient to cause PBC². However, only one genomic deletion has been published to date⁷. Among the 7 genes deleted in this family, the *THAP1* gene also retained the attention as loss of function variants of this gene cause dystonia, which was a prominent feature associated with PBC in this family. We report here the first *SLC20A2* partial deletions causing PBC. Among 24 patients with genetically unexplained PBC, we found 4 (16.7%) *SLC20A2* partial deletions, suggesting that such events should be primarily assessed, together with *SLC20A2* coding sequence variants, during the genetic screening of PBC patients. In our series, *SLC20A2* causative variants, including sequence variants and CNVs, are found in 28.5% of unrelated probands (40% when focusing on familial cases).

Interestingly, almost all patients with a partial *SLC20A2* deletion exhibited the same radiological phenotype as patients carrying other loss of function variants, i.e. calcification of lenticular, caudate nuclei, thalami, cerebellar hemispheres and in particular the vermis and cortical sulci (Table 1, Supplementary Figure 1). All were symptomatic, which seems different from the 58% of symptomatic mutation carriers in our previous report¹⁴. Of note, in family ROU-1129, the grandmother developed cognitive decline only after the age of 86 (of unknown cause), and 2 affected relatives were asymptomatic by family interview. This suggests that the apparent high proportion of symptomatic patients in the present report could be an inclusion bias.

Exon 2 of *SLC20A2* is the first coding exon and contains the translation initiation ATG codon as well as very conserved amino acid residues. This deletion probably results in a loss of function. The deletion of exon 4 causes a frameshift and therefore a probable loss of function. A splice variant at the acceptor site of intron 4 has already been described in a French family³ and could result in the same consequences as the genomic deletion of exon 4, i.e. a total skipping of exon 4 in mRNA. However, RNA was not studied in the family with the splice variant and we cannot exclude that a cryptic acceptor site could be used. The deletion of exon 4-5 is in frame and results in the loss of two transmembrane domains (Fig. 1E). Exon 4 is highly conserved during evolution and it has been shown that an artificial deletion encompassing exon 5 and other exons resulted in decreased transport activity¹⁵. Taken together, this suggests that this deletion might also result in a loss of *SLC20A2* function.

Although WES was not developed with the primary aim to detect CNVs, several tools take advantage from multiple comparisons of depth of coverage allowing the detection of deletions as short as 300 bp (e.g., the deletions involving exon 2 of *SLC20A2*; size of the target in the capture kit: 316 bp). Exon 4 of *SLC20A2* is only 85 bp long (size of the target in the capture kit: 175 bp) and was detected by CANOES from WES data only after targeted genotyping. This suggests that the first step of CANOES, performed exome-wide, might be associated with good specificity but false negatives could be misleading. In presence of a reduced number of candidate genes, the genotyping step used with lower quality thresholds should allow reducing false negatives with the disadvantage of detecting false positives. Sensitivity and specificity of bioinformatics tools aiming at calling CNVs from WES data remain to be determined in comparison with other techniques allowing such a high resolution.

ACKNOWLEDGEMENTS

We are grateful to Camille Charbonnier for her help in relatedness estimation and to patients and physicians involved in medical care and follow up of patients: Philippe Baize, Nolwen Delarue, Alexandra Durr, Patrick Louf, Yann Nadjar, Gilles Olivier, Stéphane Schaeffer and Frédéric Sedel.

This study was funded by Conseil Régional de Haute Normandie - APERC 2014 n°2014-19 in the context of Appel d'Offres Jeunes Chercheurs (CHU de Rouen). SD thanks Fondation pour la Recherche Médicale (DEA20140630628) and JF thanks FACEPE (IBPG-0627-2.02/11; AMD-0047-2.00/15) for financial support.

REFERENCES

1. Nicolas G, Charbonnier C, de Lemos RR, et al. Brain calcification process and phenotypes according to age and sex: Lessons from SLC20A2, PDGFB, and PDGFRB mutation carriers. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* 2015.
2. Wang C, Li Y, Shi L, et al. Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis. *Nature genetics* 2012.
3. Nicolas G, Pottier C, Maltete D, et al. Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. *Neurology* 2013;80:181-187.
4. Keller A, Westenberger A, Sobrido MJ, et al. PDGFB mutations cause brain calcifications in humans and mice. *Nature genetics* 2013;doi:10.1038/ng.2723.
5. Legati A, Giovannini D, Nicolas G, et al. Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. *Nature genetics* 2015.
6. Nicolas G, Rovelet-Lecrux A, Pottier C, et al. PDGFB partial deletion: a new, rare mechanism causing brain calcification with leukoencephalopathy. *Journal of molecular neuroscience : MN* 2014;53:171-175.
7. Baker M, Strongosky AJ, Sanchez-Contreras MY, et al. SLC20A2 and THAP1 deletion in familial basal ganglia calcification with dystonia. *Neurogenetics* 2013.
8. Vanlandewijck M, Lebouvier T, Andaloussi Mae M, et al. Functional Characterization of Germline Mutations in PDGFB and PDGFRB in Primary Familial Brain Calcification. *PloS one* 2015;10:e0143407.
9. Nicolas G, Pottier C, Charbonnier C, et al. Phenotypic spectrum of probable and genetically-confirmed idiopathic basal ganglia calcification. *Brain : a journal of neurology* 2013;136:3395-3407.
10. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589-595.
11. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* 2010;20:1297-1303.
12. Backenroth D, Homsy J, Murillo LR, et al. CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic acids research* 2014;42:e97.
13. Hsu SC, Sears RL, Lemos RR, et al. Mutations in SLC20A2 are a major cause of familial idiopathic basal ganglia calcification. *Neurogenetics* 2013.
14. Nicolas G, Charbonnier C, de Lemos RR, et al. Brain calcification process and phenotypes according to age and sex: Lessons from SLC20A2, PDGFB, and PDGFRB mutation carriers. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* 2015;168:586-594.
15. Bottger P, Pedersen L. Mapping of the minimal inorganic phosphate transporting unit of human PiT2 suggests a structure universal to PiT-related proteins from all kingdoms of life. *BMC biochemistry* 2011;12:21.

TITLE AND LEGEND TO FIGURES

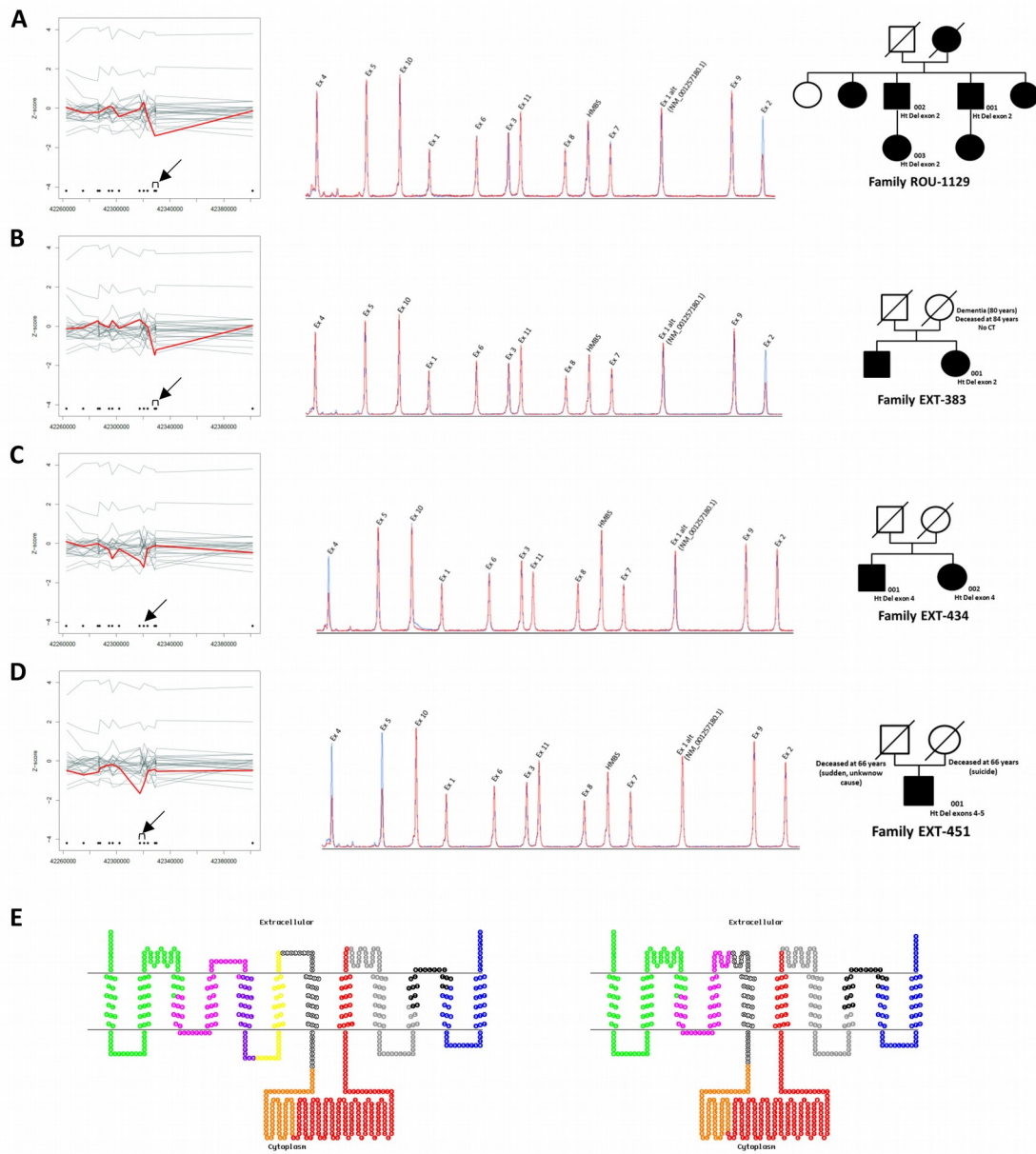


Figure 1. Partial deletions of *SLC20A2*

A-D. On the left part of the panels, Z-score of read-depth comparison among the 24 exomes are represented along genomic positions on chromosome 8. Red lines represent each patient carrying a partial deletion. Black spots represent the targets of the capture kit. Arrows point

the targets of exon 2 (A, B), exon 4 (C), and exons 4 and 5 (D). QMPSF electropherograms of the patients show partial deletions of *SLC20A2*; the patient electropherogram being in red, aligned to one control, in blue (middle part). The corresponding pedigrees are represented on the right part.

E. Schematic representation of the SLC20A2 protein as performed by the TOPO2 online software (<http://www.sacs.ucsf.edu/TOPO2/>) and according to ref.¹⁵. Each color represents one different exon, the first coding exon being exon 2 (green). Monoexonic deletions of exon 2 or 4 are not represented as they are predicted to result in no protein or a frameshift; deletion of exons 4 and 5 is represented in the right part of the panel.