UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOLOGIA APLICADA À SAÚDE
LABORATÓRIO DE IMUNOPATOLOGIA KEIZO ASAMI

DENIS ARTHUR PINHEIRO DE MOURA

Análise preditiva *in silico* da susceptibilidade funcional a alterações da proteína PiT2

Recife
2018

DENIS ARTHUR PINHEIRO DE MOURA

Análise preditiva *in silico* da susceptibilidade funcional a alterações da proteína PiT2

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Biologia Aplicada à Saúde, do Laboratório de Imunopatologia Keizo Asami (LIKA) da Universidade Federal de Pernambuco, como parte dos requisitos à obtenção do título de Mestre em Biologia Aplicada à Saúde.

Área de Concentração: Biologia Aplicada à Saúde

Orientador: Dr. João Ricardo Mendes de Oliveira

Departamento de Neuropsiquiatria, CCS/UFPE

Laboratório de Imunopatologia Keizo Asami – LIKA/UFPE

Recife

2018

Catalogação na fonte
Elaine Barroso
CRB 1728

Moura, Denis Arthur Pinheiro de

Análise preditiva *in silico* da susceptibilidade funcional a alterações da proteína PiT2 / Denis Arthur Pinheiro de Moura- 2018.

51 folhas: il., fig., tab.
Orientador: João Ricardo Mendes de Oliveira
Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Biociências. Programa de Pós-Graduação em Biologia Aplicada à Saúde, Recife, 2018.
Inclui referências e anexos

1. Neurogenética 2. Bioinformática 3. Genômica I. Oliveira, João Ricardo Mendes de (orient.) II. Título

576.5 CDD (22.ed.) UFPE/CB-2018-088

DENIS ARTHUR PINHEIRO DE MOURA

Análise preditiva *in silico* da susceptibilidade funcional a alterações da proteína PiT2

Dissertação de Mestrado apresentada ao Programa de Pós-Gradução em Biologia Aplicada à Saúde, do Laboratório de Imunopatologia Keizo Asami (LIKA) da Universidade Federal de Pernambuco, como parte dos requisitos à obtenção do título de Mestre em Biologia Aplicada à Saúde

Aprovado em: 28 de Fevereiro de 2018

Banca Examinadora

_____

PROF. DR. JOÃO RICARDO MENDES DE OLIVEIRA
PRESIDENTE (ORIENTADOR)
DEPATAMENTO DE NEUROPSIQUIATRIA – CCS/UFPE

_____

PROF. DR. MICHEL SATYA NASLAVSKY
1º TITULAR
INSTITUTO DE BIOCIÊNCIAS – UNIVERSIDADE DE SÃO PAULO

_____

PROF. DR. SERGIO CROVELLA
2º TITULAR
DEPARTAMENTO DE GENÉTICA – CCB/UFPE

_____

PROF. DRª CÍNTIA RENATA COSTA ROCHA
1º SUPLENTE
DEPARTAMENTO DE BIOQUÍMICA– CCB/UFPE

_____

PROF. DR. MANUELA BARBOSA RODRIGUES DE SOUZA
2º SUPLENTE
CENTRO DE CIÊNCIAS BIOLÓGICAS E SAÚDE – UNIVERSIDADE CATÓLICA DE PERNAMBUCO

AGRADECIMENTOS

Foram dois anos de investimento total em um projeto de mestrado em Biologia Aplicada à Saúde. Processo longo, mas com bons frutos e que permitiu acentuar o crescimento acadêmico e pessoal que hoje possuo. Tenho a agradecer pessoas que facilitaram todo o processo, seja pelo convívio agradável e despretensioso, pelas companhias em praticamente todos os momentos, ou pelo suporte financeiro e emocional. Pessoas que tornam a vida leve e prazerosa, e fazem momentos como esse acontecerem de forma natural.

Agradeço primeiramente a minha família, sobretudo meus pais, os quais sem eu não teria a menor possibilidade de investir este tempo em estudos acadêmicos. Tanto pelo suporte financeiro, como pela paciência e compreensão de que a carreira científica demanda um longo período de estudos e pouco retorno econômico.

Agradeço aos meus colegas do laboratório de Neurogenética do LIKA, com os quais formei laços de amizade. Darlene, Eraldo, Eriton, Joana, Laura, Lylyan, Rayssa e Roberta. Pessoas que tornam a rotina do laboratório agradável e divertida, facilitando demais o processo da pós-graduação. Pelas horas pontuais do café, às 15h, pelas brincadeiras e pelos momentos de aprendizagem durante os seminários semanais. Reflexo da orientação amigável e eficiente do professor João Ricardo Mendes de Oliveira, a quem devo agradecimentos especiais. Um mestre sem o qual não seria nem possível iniciar, quiçá finalizar com sucesso, o processo que foi o mestrado.

Finalmente, agradeço a minha noiva, Amanda Melo, por todo apoio e incentivo prestado durante esses sete anos que estamos juntos. A pessoa que está por trás de qualquer sucesso que obtive e terei por todos esses anos. Pelos auxílios em *design* gráfico, pelas broncas, conselhos e carinhos. Obrigado, meu amor!

RESUMO

Neste trabalho são apresentados resultados de uma análise *in silico* sobre os efeitos das alterações no gene *SLC20A2* no desenvolvimento da doença neurodegenerativa Calcificação Cerebral Familial Primária (CCFP). O *SLC20A2* codifica o transportador de fosfato sódio-dependente PiT2. PiT2 é composto de 12 domínios transmembrana e 12 domínios topológicos intra- e extracelulares. Recentemente, mutações nesse gene foram ligadas como a maior causa da CCFP, uma neuropatologia caracterizada por calcificações bilaterais no cérebro. Mais de 50 variantes do *SLC20A2* foram associadas a CCFP. Conhecer a estrutura do PiT2 é importante para a compreensão do desenvolvimento da doença. Nesse aspecto, ferramentas de predição de proteínas são utilizadas para informar possíveis efeitos de variações no gene em sua proteína codificada. Nesse trabalho foi utilizado o banco de dados dbNSFP para traçar o perfil de susceptibilidade a variações da PiT2 através da sobreposição de 10 diferentes algoritmos de predição. Um total de 4545 anotações do *SLC20A2*, compreendendo variantes benignas e ligadas a doença, foram prospectadas do banco de dados dbNSFP. Esses dados foram organizados, visualizados e analisados na linguagem R. A combinação de algoritmos de predição revelou os exons 3, 4, 5, 10 e 11 como tendo proporcionalmente o maior número de variantes preditas como danosas, enquanto os exons 7, 6, 2 e 8 apresentando os menores números. Nenhuma correlação foi encontrada entre o comprimento do exon e a proporção de variantes danosas. A proporção média geral de variantes danosas encontrada para o *SLC20A2* foi de 8,63%. Os exons 3, 4 e 5 tiveram a maior proporção de variantes danosas (~15% cada), seguidos pelo exon 10 (13,29%). Isso indica uma maior susceptibilidade a variações nessas posições para a funcionalidade da proteína. Da mesma forma, os exons 7 e 6 (0 e 1,5%) foram considerados mais tolerantes a variações. A literatura recente apoia a alta proporção de variantes danosas para o exon 10 em pacientes com CCFP, enquanto nenhuma variante *missense* foi encontrada para os exons 7 e 6. O PiT2 foi analisado em relação a sua susceptibilidade a variações em cada posição. Exons 3, 4, 5 e 10 foram considerados mais susceptíveis, enquanto exons 7 e 6 foram considerados mais tolerantes. Esse método de análise do perfil de susceptibilidade pode ser utilizado em outras proteínas, possivelmente levando a um melhor entendimento de suas estruturas genéticas.

PALAVRAS-CHAVE: Bioinformática. Neurogenética. Genômica.

ABSTRACT

This dissertation presents results from a *in silico* analysis regarding the effects of alterations in the *SLC20A2* gene in the development of the neudegenerative disease Primary Familial Brain Calcification (PFBC). *SLC20A2* codifies the sodium-dependent phosphate transporter PiT2. PiT2 is composed of 12 transmembrane and 12 intra- and extracellular topological domains. Recently, mutations in this gene were linked as major causative to Primary Familial Brain Calcifications (PFBC), a neuropathology characterized by bilateral calcifications in the brain. Over 50 *SLC20A2* variants were associated to PFBC. Understanding PiT2's structure is important for the comprehension of the disease development. Protein prediction tools are used to inform possible effects of variations in the gene to its encoded protein function. We used the dbNSFP database to profile the PiT2 variation susceptibility by overlaying 10 different prediction algorithms. A total of 4545 *SLC20A2* annotations, comprising benign and disease-related variants, were prospected from dbNSFP database. This data was organized, visualized and analyzed in R language. The combination of the prediction algorithms revealed exon 3, 4, 5, 10 and 11 with the proportionally highest number of predicted damaging variants, while exons 7, 6, 2 and 8 having the lowest numbers. No correlation was found between exon length and proportion of damaging variants. The overall proportion of damaging variants for *SLC20A2* found was 8.63%. Exons 3, 4 and 5 had the highest proportion of damaging variants (~15% each), followed by exon 10 (13.29%). This indicates a higher susceptibility to variation in these positions for protein functionality. Likewise, exons 7 and 6 (0 and 1.5%) were considered more variation-tolerant. Recent literature supports a high proportion of damaging variants for exon 10 in PFBC patients, while no PFBC-related missense variants were found for exons 7 and 6. PiT2 was profiled regarding its variation susceptibility in each position. Exons 3, 4, 5 and 10 were considered highly variation-susceptible, while exons 7 and 6 were considered more variation-tolerant. This method of prediction profiling can be used for other proteins, possibly leading to a better understanding of its structures.

KEYWORDS: Bioinformatics. Neurogenetics. Genomics.

SUMÁRIO

1 INTRODUÇÃO

1.1 PROBLEMATIZAÇÃO

Calcificação Cerebral Familiar Primária (CCFP) é uma doença neurológica rara caracterizada por padrões bilaterais de calcificação do encéfalo. As regiões mais afetadas são as estruturas do gânglio basal, tálamo e cerebelo (núcleo dentado). Os sintomas neurológicos motores e cognitivos associados à CCFP são comparáveis aos de outras doenças neurodegenerativas como parkinsonismo, esquizofrenia, demência, podendo também ser assintomática (OLIVEIRA, 2011; DA SILVA et al., 2013).

Quatro moléculas foram confirmadas como causadoras de CCFP. A primeira foi o transportador de fosfato inorgânico PiT-2, responsável por cerca de 40% dos casos de CCFP (WANG et al., 2012; LEMOS et al., 2013). As outras duas moléculas a serem relacionadas foram a PDGFB (subunidade beta do fator de crescimento derivado de plaquetas) e seu receptor PDGFRB, envolvidos na velocidade de incorporação de fosfato inorgânico pela PiT-1 (NICOLAS et al., 2013b; KELLER et al., 2013; GIANCHELLI et al., 2001; e LEMOS et al., 2013). Em 2015 foi confirmada como causadoras de CCFP mutações o gene *XPR1*, envolvido na exportação de fosfato inorgânico celular (LEGATI et al., 2015).

Algumas outras formas de calcificação cerebral foram ligadas a mutações de herança autossômica dominantes, geralmente com a presença de outros achados clínicos associados, como imunodeficiência primária ou graves más formações cerebrais. O gene mais recente confirmado como relacionado a calcificações cerebrais é o *PCDH12* (NICOLAS et al, 2017), um gene que codifica uma protocaderina envolvida na adesão e integridade da membrana das células, com perfil de expressão cerebral muito semelhante ao gene *SLC20A2* (DA SILVA et al, 2013).

Com quatro, e possivelmente mais, genes ligados à doença e mais de uma variante de cada gene como causa, a CCFP é considerada uma condição poligênica. Além disso, a divergência de sintomas entre cada caso, mesmo com variações no mesmo gene como causa, levanta a questão de como cada um dos genes está envolvido no desenvolvimento da doença. Compreender, portanto, como cada variante afeta cada um dos genes se torna uma peça chave na elucidação dos aspectos de causa e efeito da genética da CCFP. O gene *SLC20A2*, associado ao maior número de casos da condição, torna-se o principal alvo de estudos genéticos que entendamos como seu produto, a

proteína PiT2, é afetado com cada uma das possíveis variações no gene. Nesse aspecto, estudos de bioinformática permitem modelar os possíveis cenários genéticos e emular os efeitos que cada uma das possibilidades vão afetar a funcionalidade da proteína; o que pode ser extrapolado como causa da CCFP.

1.2 JUSTIFICATIVA

Sendo um gene de importância médica, como o principal causador da calcificação cerebral familiar primária, são necessários esforços para conhecer melhor o papel metabólico e o funcionamento do gene *SLC20A2*. Não se sabe ao certo o porquê de cada mutação encontrada nas variantes levarem ao desbalanço iônico que resulta nas calcificações. Dessa forma, uma análise *in silico*, emulando os possíveis cenários de variações genéticas para o gene *SLC20A2*, permitirá o estudo dos possíveis efeitos de cada uma das variações nesse gene através de algoritmos de predição de perda de funcionalidade da proteína. O conhecimento gerado neste projeto certamente será de grande importância para a compreensão da estrutura genética do gene *SLC20A2*, possivelmente auxiliando a compreensão das causas da CCFP e dos diferentes níveis de sintomas da doença.

1.3 OBJETIVOS

1.3.1 Objetivo geral

Gerar e analisar o perfil de susceptibilidade a variações *missense* do gene *SLC20A2* utilizando algoritmos de predição.

1.3.2 Objetivos específicos

- Desenvolver uma metodologia de análise *in silico* dos efeitos de variações genéticas no gene *SLC20A2*, traçando para tal um perfil de susceptibilidade a variações dependente de posição.
- Validar perfil de susceptibilidade traçado *in silico* comparando com dados de pacientes com Calcificação Cerebral Familiar Primária.

2 REFERENCIAL TEÓRICO

2.1 CALCIFICAÇÕES CEREBRAIS

Calcificações ectópicas em tecidos moles são muitas vezes consideradas um processo dependente de idade. Nitschke e Rutsch (2012) sugeriram que uma diversidade de moléculas modula ativamente esse processo. No cérebro, as calcificações patológicas são caracterizadas pelo acúmulo de cálcio e outros minerais nas paredes dos vasos sanguíneos e do parênquima, levando a morte neuronal e gliose (SOBRIDO *et al.*, 2014). Tais calcificações cerebrais são encontradas geralmente no gânglio basal, cerebelo, tálamo e mesencéfalo, e são comumente associadas as várias desordens de limitações cognitivas e motoras, como parkinsonismo, distonia, demência, psicose e mudanças de humor. Além disso, casos de calcificações cerebrais assintomáticas também são comuns (LEMOS *et al*., 2013).

Em geral, as calcificações cerebrais encontradas em neurodegenerações, independentemente das diferentes etiologias, têm uma estrutura cristalina composta de um conjunto de hidroxiapatita, feita principalmente por fosfato de cálcio (MAHY *et al*., 1999; RODRÍGUEZ *et al*., 2009). Com a crescente disponibilidade de exames de neuroimagem, está cada vez mais frequente o número de casos esporádicos de calcificações cerebrais, fato discutido por Oliveira e Steinberg (2010) como sugestão de que tal fenótipo é mais comum do que se imaginava. Ainda sobre o diagnóstico das calcificações cerebrais, Oliveira (2011) e Sobrido *et al* (2014) argumentam que uma grande bateria de exames deve ser realizada. Esses exames incluem os níveis de paratormônio, cálcio, fósforo, eletrólitos do soro, hormônios da tireóide e de crescimento, cortisol, taxa de sedimentação, testes de função renal e hepática, níveis de ceruloplasmina e marcadores virais. Para condições mais raras, como encefalomiopatia, acidose lática e outras desordens genéticas, pode ser necessário analisar o cariótipo, biópsia muscular e análises de mutações. A tomografia computadorizada (em inglês, *computerized tomography – CT*) é considerado o método mais apropriado de neuroimagem para detectar calcificações cerebrais (YAMADA *et al*., 2013).

Análises de neuroimagem e comparações dos padrões de calcificação cerebral em diferentes síndromes induziram Livingston *et al* (2013), e outros autores, a tentar classificar as calcificações baseando-se nas estruturas mais afetadas. Apesar das limitações admitidas pelos autores, em relação as variações comuns dos procedimentos e as análises heterógenas dos dados de

neuroimagem, as classificações sugeridas foram: mecanismo provavelmente destrutivo; calcificação associada à polimicrogiria; calcificação familiar com envolvimento do globo pálido, membro posterior da cápsula interna, geno do corpo caloso e matéria branca profunda; e por fim, calcificação familiar dominante (calcificação cerebral familiar primária). Apesar de variações intrafamiliares serem comuns, inclusive em uma mesma geração, já há evidência de padrões de calcificações idênticos em uma mesma família (LEMOS *et al*., 2013).

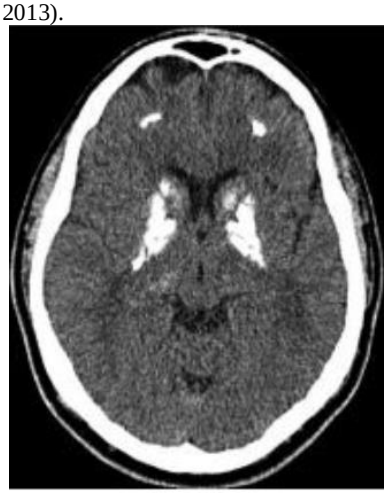
2.2 CALCIFICAÇÃO CEREBRAL FAMILIAR PRIMÁRIA (CCFP)

A CCFP é uma desordem neurológica com sintomas motores e cognitivos, caracterizada pela calcificação bilateral mais comumente encontrada na região do gânglio basal, mas também em outras regiões como o tálamo e o cerebelo (Figura 1). É considerada uma condição rara, com prevalência de menos de um caso por milhão de pessoas. As calcificações são insidiosas e expandem progressivamente pelos vasos e parênquima (OLIVEIRA, 2011). Os sintomas neuropsiquiátricos se apresentam como parkinsonismo, demência, psicose e mudanças de humor, podendo também ser assintomática.

Numerosas desordens neurológicas podem levar a calcificações cerebrais, porém as calcificações caracterizadas como “síndrome de Fahr” têm caráter genético dominante (familiar) e um padrão de calcificação bilateral simétrico na região do estriadopalidodentado. Apesar do caráter genético dominante, casos idiopáticos, associados as triagens bioquímicas e endocrinológicas normais também são comuns. Por essa característica, a Calcificação Cerebral Familiar Primária (CCFP, em inglês: “*Primary Familial Brain Calcification*” - PBFC) já teve como nome oficial Calcificação Idiopática do Gânglio Basal e Calcificação Idiopática Familiar do Gânglio Basal (em inglês, “*Familial Idiopathic Basal Ganglia Calcification* – FIBGC”). A denominação “síndrome de Fahr” foi feita a partir do médico Karl Theodor Fahr. Tal nomenclatura está caindo em desuso pela tendência da nosologia em abolir o uso excessivo de desordens que referenciam o nome de uma pessoa, a fim de evitar epônimos, e também por não ter sido o médico o primeiro a reportar essa doença (OLIVEIRA, 2011). O termo Calcificação Cerebral Familiar Primária foi sugerido por Sobrido *et al*. (2014), com base nas causas genéticas agora conhecidas e o desenvolvimento das calcificações em outras regiões do cérebro.

A região conhecida como gânglio basal, área de maior efeito da CCFP, agrupa um conjunto de corpos neuronais (gânglios) que formam estruturas subcorticais imersas na massa branca cerebral. Essa região é responsável por um grande número de processos motores e cognitivos, incluindo o equilíbrio, coordenação e memória implícita, sendo essa região envolvida também em diversas outras condições neuropsiquiátricas como doença de Parkinson, distonia, coreia e transtorno obsessivo compulsivo (OLIVEIRA, 2011).

Figura 1: Padrão de calcificação cerebral encontrado em pacientes com Calcificação Cerebral Familiar Primária. Figura de paciente com mutações no gene *PDGFB* (Keller *et al.*, 2013).



Fonte: Keller *et al.*, 2013

2.3 MECANISMO GENÉTICO DA CCFP

2.3.1 Gene *SLC20A2*

Um padrão autossômico dominante de segregação é encontrado na maioria das famílias com CCFP, com poucos casos apresentando uma herança recessiva (LEMOS *et al.*, 2013). Recentemente os casos de CCFP foram relacionados a causas genéticas, como o gene *SLC20A2* que codifica a proteína transportadora de fosfato inorgânico (PiT-2), com padrões de expressão e correlação neuroanatômica com as áreas mais afetadas pelas calcificações. A relação entre o gene e a doença também a relaciona ao mecanismo de transporte de fosfato (WANG *et al.*, 2012). Entretanto, só é possível atrelar aproximadamente 40% dos casos estudados até hoje ao gene *SLC20A2*, evidenciando a heterogeneidade genética da síndrome (LEMOS *et al.*, 2013).

A PiT-2 é uma proteína transmembrana identificada previamente como um receptor para o retrovírus anfotrópico de murinos (Ram-1). Essa proteína tem um papel no transporte de sódio-fosfato inorgânico, com função da manutenção da homeostase de fosfato inorgânico. É envolvida na mineralização dos ossos e na calcificação ectópica (KAVANAUGH e KABAT, 1996; COLLINS *et al*., 2004; VIRKKI *et al*., 2007).

Mutações encontradas para o gene *SLC20A2* em pacientes com CCFP vão desde sentido alterado, alterações do quadro de leitura, deleções, mutações sem sentido, mutações sinônimas, e até de sítio de *splicing*. Segundo Oliveira *et al*. (2013), as análises *in silico* de diferentes algoritmos de bioinformática predizem que a maioria dessas mutações é deletéria. Hsu *et al*. (2013) encontrou que nove entre catorze mutações eram preditas como introdutores de um códon de terminação, apontando haploinsuficiência como mecanismo de causa. Zhang *et al*. (2013) demonstrou por análise da expressão relativa do gene *SLC20A2* com uma mutação de deleção produzia uma proteína PiT-2 truncada, com terminação prematura. Essa alteração resultou em uma redução de aproximadamente 30% na expressão desse gene (níveis de RNA mensageiro) em relação a indivíduos do grupo controle. Mutações *de novo* no gene *SLC20A2* também foram encontradas e relacionadas como causas de CCFP (FERREIRA *et al*., 2014).

2.3.2 Outros genes causadores da CCFP

Pouco após a descoberta da associação entre mutações no *SLC20A2* e a Calcificação Cerebral Familiar Primária, Nicolas *et al*. (2013) encontrou duas famílias apresentando CCFP com mutações no gene *PDGFRB*, que codifica um receptor de tirosina-quinase classe III, e Keller *et al*. (2013) encontrou outras seis famílias com CCFP e mutações no gene *PDGFB*. As mutações no gene *PDGFRB* afetavam resíduos conservados no domínio quinase, enquanto as mutações do *PDGFB* não têm um mecanismo de atuação compreendido. O PDGFB é envolvido em muitos mecanismos incluindo angiogênese, sobrevivência de pericitos e a manutenção barreira hematoencefálica (BHE). Um dos mecanismos propostos pelos quais os genes *PDGFB* e *PDGFRB* são envolvidos nos padrões de calcificação da CCFP seria pela modulação da velocidade de captação do fosfato inorgânico pelo transportador de fosfato PiT-1 (GIANCHELLI *et al*., 2001; LEMOS *et al*., 2013).

Camundongos com alterações no gene *PDGFB* são utilizados para o estudo da funcionalidade da BHE e como modelos para CCFP. Esses animais desenvolvem naturalmente calcificações

cerebrais similares a dos pacientes com a doença, sobretudo o camundongo *Pdgfb*$^{ret/ret}$. Essa linhagem tem uma mutação em homozigose no gene *PDGFB*, com um produto biologicamente ativo, compatível com a vida adulta, apesar de ser hipomórfico pela ausência da região C-terminal (KELLER *et al*, 2013; VANLANDEWIJCK *et al*, 2015; VIILLASEÑOR *et al*, 2017). Moura, Lemos e Oliveira (2017) reuniram informações de diferentes autores sobre os cérebros dos camundongos *Pdgfb*$^{ret/ret}$ e construíram um modelo em três dimensões com dados de extravasamento de BHE e calcificações cerebrais (Figura 2). Foi observado que regiões mais passíveis de calcificar (mesencéfalo, diencéfalo, prosencéfalo basal e ponte) apresentam uma menor taxa de extravasamento de BHE em relação a outras regiões neuroanatômicas. Tal resultado indica que a formação das calcificações pode não estar relacionada com falhas na BHE. A relação entre BHE, pericitos e calcificações cerebrais permanece incompletamente entendida, sendo necessários mais estudos para entender a patofisiologia da CCFP causada por alterações no *PDGFB* e *PDGFRB*.

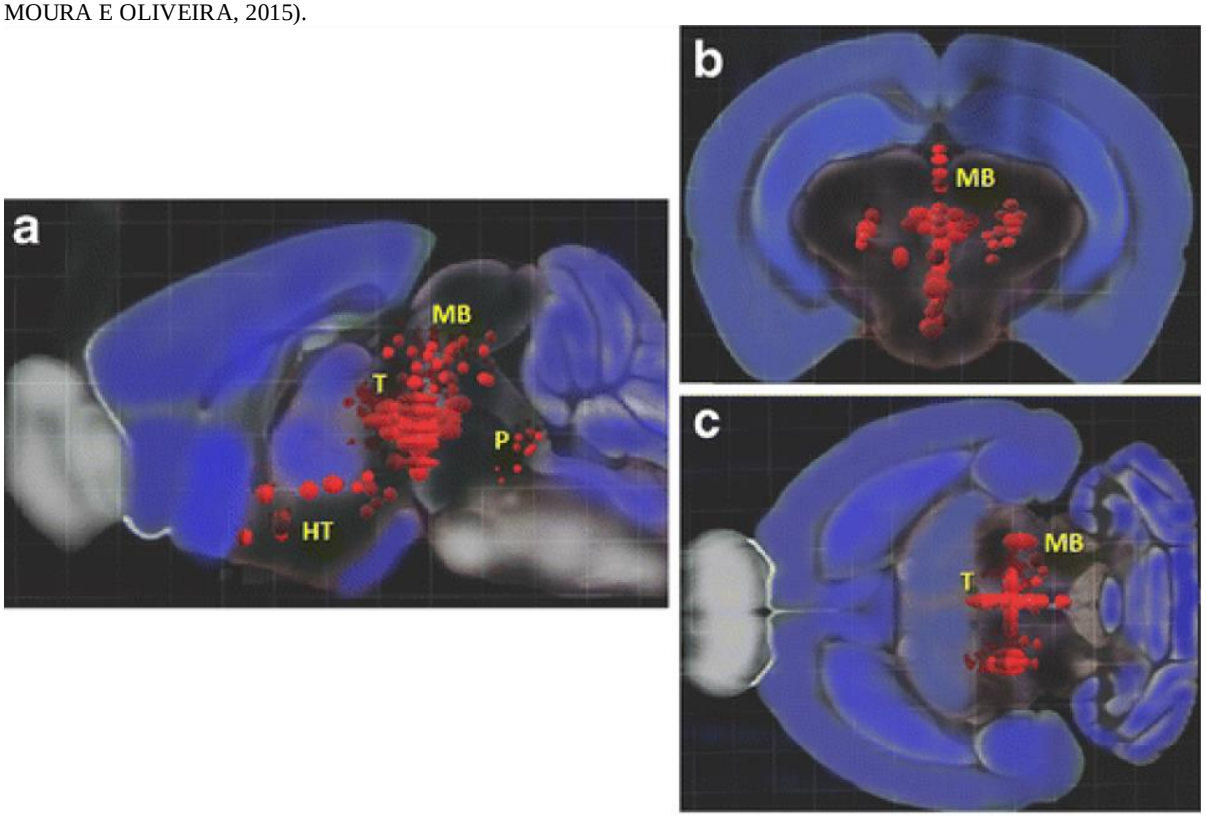
Mutações no gene *XPR1* também foram associadas à CCFP (LEGATI *et al*, 2015). Múltiplas famílias não relacionadas geneticamente apresentaram mutações no *XPR1* no domínio SPX, relacionado a exportação de fosfato celular, ou em suas proximidades. Diferente do mecanismo de calcificação pelo *SLC20A2* (que codifica a proteína transportadora de fosfato inorgânico PiT-2), onde a inibição da introdução de fosfato leva a deposição de fosfato de cálcio na matriz extracelular dos vasos, as mutações no XPR1 impedem a exportação do fosfato celular, causando acúmulo desse elemento dentro da célula. Esse gene também foi associado a outras doenças neurodegenerativas como esquizofrenia, esclerose lateral amiotrófica, degeneração do lobo frontal (ALS/FTD) e Doença de Alzheimer (XU *et al*, 2012; FUJIOKA *et al*, 2013; MANAVALAN *et al*, 2013)

2.4 GENÔMICA E BIOINFORMÁTICA

A utilização de ferramentas de bioinformática tem possibilitado a integração de dados genômicos, proteômicos, metabolômicos, entre outros ômicos, na formação de redes genéticas e metabólicas para melhor elucidar condições fisiológicas e patológicas. Dentro desse cenário, doenças complexas, que são condições causadas por mais de um fator genético, ambiental ou

combinação desses, são ainda mais beneficiadas com a utilização de ferramentas de genômicas e de bioinformática (MIDDLETON *et al*., 2007; LOWE E REDDY, 2015).

Figura 2: Modelo representativo dos efeitos da perda de função da proteína *PDGFB* no cérebro de camundongos (*Pdgfb*$^{ret/ret}$). São representadas áreas com maior (azul) e menor (preto) extravasamento de barreira hematoencefálica, áreas de calcificação (vermelho) e áreas sem informações disponíveis (cinza). A) Vista sagital. B) Vista coronal. C) Vista horizontal. 013; e MOURA E OLIVEIRA, 2015).



Fonte: Figura retirada de Moura, Lemos e Oliveira (2017).

Dentro do conjunto de ferramentas de bioinformática, os algoritmos de predição de função são particularmente importantes para analisar resultados de estudos genômicos e elucidar os efeitos de variações genéticas no desenvolvimento de processos patológicos (LIU, JIAN e BOERWINKLE, 2011; LIU *et al*., 2016).

3 MÉTODOS

3.1 BUSCANDO O GENE *SLC20A2* NO BANCO DE DADOS dbNSFP

O dbNSFP (versão 3.29, LIU et al, 2016) foi utilizando para baixar todas as variantes preditas para o gene *SLC20A2* contidas no banco de dados em outubro de 2016. Para isso foi necessário instalar o ambiente Java e baixar o banco de dados inteiro, um grande arquivo de 89 Gigabytes e a criação de um arquivo de entrada contendo o nome do gene 'slc20a2'. O *prompt* de comando do Windows foi utilizado para acessar o banco de dados com a linha 'java search_dbNSFP32a -i slc20a2.in -o slc20a2.out'. Esse comando gerou uma lista de 4544 diferentes anotações de variantes para o gene *SLC20A2*, cada uma testada em 30 diferentes algoritmos de predição em relação ao seu efeito na funcionalidade da proteína.

3.2 ANALISANDO AS VARIANTES PREDITAS DO GENE *SLC20A2*

A linguagem R (versão 3.3.3) foi utilizada junto ao *software* RStudio (versão 1.0.143) para analisar os dados gerados pelo dbNSFP em relação as variantes do gene *SLC20A2*. Primeiro, os dados foram organizados por posição cromossômica, separando por exon.

Para gerar uma análise superficial do perfil do gene, o algoritmo de predição SIFT foi utilizado para traçar os resultados para todo o gene. Dos 30 diferentes algoritmos de predição, apenas aqueles que caracterizam a variação como 'danosa' ou 'tolerada' foram utilizados nessa análise. Dois subconjuntos de dados foram criados com os dados considerados 'danosos' (D) pelos algoritmos SIFT, Polyphen2 HDIV, Polyphen2 HVAR, LRT, MutationTASTER, FATHHM, PROVEAN, MetaSVM, MetaLR e FATHHM-MKL e as variantes remanescentes foram consideradas toleradas. O número de variantes toleradas e danosas foi contabilizado para cada exon e, finalmente, os exons foram comparados entre si em relação ao número de variantes preditas como danosas.

Para estatisticamente os dados foi realizado um teste T de Student, comparando o número de variantes preditas como danosas para cada exon. Uma análise de regressão linear foi realizada no R com o comando "lm" para checar se o tamanho do exon tem relação com o número de variantes preditas como danosas. Por fim, foi realizada uma análise de correlação de Pearson utilizando a fórmula "cor" do R. O *script* completo do R é apresentado em 3.1.

3.3 REVISÃO DE LITERATURA SOBRE VARIANTES *MISSENSE* DO GENE *SLC20A2* LIGADAS A CCFP

Uma pesquisa aprofundada foi realizada para reunir possivelmente todos os dados disponíveis de pacientes com CCFP associados a variações *missense* no gene *SLC20A2*. As palavras-chave "SLC20A2", "primary familial brain calcification", "idiopathic basal ganglia calcification" e "PiT2" foram buscadas nos índices de publicações acadêmicas PubMed e Google Scholar, entre o período de janeiro de 2012 à dezembro de 2017. Variantes *missense* do gene *SLC20A2* associadas a pacientes com CCFP foram anotadas e reunidas para este trabalho.

4 RESULTADOS

4.1 ARTIGO SUBMETIDO AO PERIÓDICO PROTEINS: STRUCTURE, FUNCTIONS AND BIOINFORMATICS - ‘Profiling PiT2 protein variants effects on functionality by prediction algorithms’

TITLE: Profiling PiT2 protein variants effects on functionality by prediction algorithms

SHORT TITLE: PiT2 variation susceptibility profile

KEYWORDS: Bioinformatics; Neurogenetics; Genomics; Proteomics; Brain Calcification

AUTHORS: Denis AP Moura[1], Roberta R Lemos[1] João RM Oliveira[1,2]

[1]Laboratório de Imunopatologia Keizo Asami – Universidade Federal de Pernambuco – Brazil

[2]Departamento de Neuropsiquiatria – Universidade Federal de Pernambuco - Brazil

All work was performed at the Universidade Federal de Pernambuco – Brazil.

CORRESPONDING AUTHOR:

João RM Oliveira

Av. Prof. Moraes Rego, 1235 - Cidade Universitária, Recife - PE - CEP: 50670-901. Laboratório de Neurogenética, LIKA/UFPE.

joao.ricardo@ufpe.br

Abstract

SLC20A2 codifies the sodium-dependent phosphate transporter PiT2. Mutations in this gene are a major cause to Primary Familial Brain Calcifications (PFBC), neuropathology characterized by bilateral brain calcifications. Over 50 *SLC20A2* variants are currently associated to PFBC. Understanding PiT2's structure is important to comprehend the disease development. Protein prediction tools are useful to inform possible effects of variations in the gene to its encoded protein function. dbNSFP database was used to profile PiT2 variation susceptibility by overlaying 10 different prediction algorithms. A total of 4545 *SLC20A2* annotations, comprising benign and disease-related variants were prospected from dbNSFP database. This data was organized, visualized and analyzed in R language. No correlation was found between exon length and proportion of damaging variants. The overall proportion of damaging variants found was 8.63%. Exons 3, 4 and 5 had the highest proportion of damaging variants (~15% each), followed by exon 10 (13.29%). This indicates a higher susceptibility to variation in these positions for protein functionality. Likewise, exons 7 and 6 (0 and 1.5%) were considered more variation-tolerant. Recent literature supports a high proportion of damaging variants for exon 10 in PFBC patients, while no PFBC-related missense variants were found for exons 7 and 6. In conclusion, PiT2 was profiled regarding its variation susceptibility in each position. Exons 3, 4, 5 and 10 were considered highly variation-susceptible, while exons 7 and 6 were considered more variation-tolerant. This method of prediction profiling can be used for other proteins, possibly leading to a better understanding of its structures.

Introduction

Gene *SLC20A2* codifies a transmembrane protein PiT-2. This protein, along with its homologous PiT-1 (*SLC20A1*), forms up the NaPiT3 protein family of type-3 phosphate transporters. These proteins`s function is in inorganic phosphate homeostasis and are they involved directly in bone mineralization and ectopic calcifications[1,2]. PiT2 was previously identified as a receptor for Ram-1, an anfotropic murine retrovirus[3]. Phosphate transporters share a PH04 superfamily domain and PiT2 has two phosphate transporter domains close to the N and C terminal regions of the protein. These domains are located at positions ~5-225 and ~490-625. Within these domains there is a highly conserved sequence named GANDVANA. Within this sequence, the glycine and aspartic acids at positions two and five are the most conserved residues in most lifeforms of earth. And mutations in these residues severely impair the phosphate transport function of PiT2 while maintaining the viral receptor function of this protein[4]. In PiT2, this sequence is located at both PH04 domains, in positions 25-32 and 503-510. These sequences are comprised in exons 2 and 9, respectively. Part of the coding sequence within these exons are directly related to the phosphate transport function of the protein.

Several mutations in *SLC20A2* were considered causative for Primary Familial Brain Calcification (PFBC), a rare neuropsychiatric condition involving symmetrical bilateral calcifications in the brain, affecting mostly the midbrain, thalamus, basal forebrain, pons and cerebellum[5]. PFBC-linked mutations of *SLC20A2* vary from missense to nonsense mutations, as

well as splicing site mutations[6]. Hsu et al[7] found that nine in fourteen of these mutations were predicted to introduce a premature stop codon, pointing protein haploinsuficiency as a causal mechanism for the disease. In cases where both nonsense and missense mutations are cause for a particular condition, because nonsense mutations effects are too extreme, understanding the effects of missense mutations might lead to a better understanding of the protein involvement in the pathomolecular mechanism of the condition.

Protein prediction tools are useful to inform the possible effects of a newly found variant in the gene to its encoded protein. Several different prediction algorithms are currently available for this end. The dbNSFP database groups 14 of the available prediction algorithms into a single database, making it incredibly easy to assess information about prediction of protein functionality by missense variations[8]. We have used this database to profile the *SLC20A2* gene *per* exon by overlaying different prediction algorithms.

Materials and Methods

Querying SLC20A2 gene on dbNSPF database

The dbNSFP (version 3.2[9]) was used to download all predicted variants for *SLC20A2* gene contained in the database in October 2016. This required the installation of the Java environment and the download of the whole database, a large file of 89 Gb and the creation of the input file containing the gene name 'slc20a2'. Windows' command prompt was used to access the database with the line "java search_ dbNSFP32a -i slc20a2.in -o slc20a2.out". This command generated a list of 4544 different variant annotations for *SLC20A2* gene, each one tested in 30 different prediction algorithms regarding its effects on protein functionality.

Analyzing SLC20A2 predicted variants

R language (version 3.3.3) was used along with RStudio software (version 1.0.143) to analyze the dbNSFP generated data regarding *SLC20A2* gene variants. First the data was selected by chromosome position, separating it by exon (Table S1).

To have a superficial analysis of the gene's profile, SIFT prediction algorithm was used by plotting the results for the whole gene (Figure 1). Of the 30 different prediction algorithms, only those that characterize the variant effect as damaging or tolerable were used in this analysis. Two subsets were created with the data considered damaging ('D') by algorithms SIFT, Polyphen2 HDIV, Polyphen2 HVAR, LRT, MutationTASTER, FATHHM, PROVEAN, MetaSVM, MetaLR and FATHHM-MKL and the variants remaining variants as tolerable. Then the damaging or tolerable variants were accounted for each exon, and finally the exons were compared regarding the number of predicted damaging variants. Student T test was used to compare each exon. A linear regression analysis was conducted in R with the command 'lm' to check whether the exon size is related to the number of predicted damaging variants, followed by a Pearson correlation analysis using R's formula 'cor'. The complete R script can be found in the Supplementary material.

Literature review on PFBC-linked SLC20A2 missense variants

A thorough research was conducted in order to summarize *SLC20A2* missense variants associated to PFBC in patients' data. The keywords "SLC20A2", "primary familial brain calcification", "idiopathic basal ganglia calcification" and "PiT2" where searched upon academic publication indexes PubMed and Google Scholar, from 2012 to December 2017. Missense variants in the gene of PFBC-positive patients were annotated and are presented in Table 3.

Results

SLC20A2's variant damage predicted by SIFT, one of the 30 prediction algorithms available in the dbNSFP database, was analyzed first (Figure 1). Each variant has a damage score attached varying from 0 to 1, and variants with score closer to 1 are considered more damaging. Exons 2, 3, 9, 10 and 11 have apparently the highest number of predicted damaging variants based on SIFT algorithm. Exon 7, in comparison, has none predicted as damaging variants.

In this analysis a variant was considered damaging when all the 10 prediction algorithms agreed, while the remaining variants were considered tolerable, or less damaging (Table 1). Exon 8 had the highest number of considered damaging variants, followed by exons 3, 11 and 9 (50, 49, 42 and 39, respectively). Exon 7 had no considered damaging variants shared by all 10 prediction algorithms, and exons 6 and 2 had the lowest number of damaging variants (4 and 10, respectively). A linear regression model was tested with the data to assess if there is a correlation between exon size and number of damaging variants (Figure S1). An intercept of 21.96 with inclination of 0.01383 and Pearson correlation coefficient (r^2) of 0.2688 were found for this analysis, revealing a weak linear relationship between the data. The relationship between the number of damaging variants and the size of the exon is thus weakly supported.

The proportion of considered damaging over total number of variants reveals exons 3, 4, 5, 10 and 11 as most susceptible to variations (15.5%, 15.07%, 15.02%, 13.29% and 11.41%, respectively), while the average proportion of damaging variants for *SLC20A2* gene was found around 5.3. Exons 7, 6 and 2 are more tolerable regarding missense variations in *SLC20A2* (0%, 1.5% and 1.51%, respectively). Figure S2 represents these proportions for the entire *SLC20A2* gene.

Lastly, we compared the proportions of predicted as damaging variants for each exon by conducting a Student's T test. This test revealed that the proportions are statistically different between each exon ($p = 0.0019$).

Comparing the predicted data with patient data

Currently, there are 28 missense *SLC20A2* variants associated to PFBC (Table 2). All *SLC20A2* data gathered in this work was summarized in a scheme, consisting in the gene structure by

exons, with the 12 transmembrane domains, every reported PFBC-linked *SLC20A2* gene missense variants and the 'GANDVANA' sequences positions (Figure 2). The scheme also contains the number of predicted as damaging over total number of variants annotated in dbSNFP for each exon, the positions of the PHO4 phosphate transporter superfamily homologies in the sequence and a pie graph presenting each exon contribution for the total number of predicted damaging variants for *SLC20A2* gene.

Discussion

Variations in coding sequences can lead to a large spectrum of effects on protein functionality. Since the effects of number variations are extreme, we focused our analysis on missense variants.

Lemos et al[6] analyzed the number of *SLC20A2* gene variants related to Primary Familial Brain Calcification and found that these variants are well distributed along the protein. Exon 8 had the highest number of variants (~27%), followed by exon 2 (~14.5%) and exons 9, 10 and 11 (~10.5 each). Exon 8 and exon 2 are the largest exons of the gene, a possible reason for them to have proportionally higher counts of disease-related variants[6]. The variant profile revealed by our analysis agrees with this statement by Lemos and colleagues, showing both exons with low proportion of damaging over total number of annotated variants (~3.6% for exon 8 and ~1.5% for exon 2).

Our results show no correlation between exon and the proportion of damaging variants (r^2 of 0.1617). Concomitant to this finding, exon 10 while being the smallest exon of *SLC20A2* gene had a proportion of ~13.3% of damaging variants over total. This is comparable to what Lemos and colleagues[6] found when normalizing the annotated variants for the gene by exon size. They found exon 10 as having the highest proportion of variants linked to PFBC (~6%) and exon 7 with the lowest proportion (~1%).

Our results present exon 9 as having one of the highest number of variants predicted as damaging by all 10 prediction algorithms (31, 18.23% of the total number of damaging variants for the whole gene), only behind exon 3 (34, 20%). Curiously, we found exon 2 as having one of the lowest number of damaging variants in the gene (2, 1.7%), a much lower percentage of what was expected given the importance of the PH04 domain and the GANDVANA sequence. This could suggest that the C- terminal PH04 domain is more important for the protein function. And having two phosphate transport domains could serve as a protector mechanism against possible mutations.

The large intracellular central region of PiT2[10] still has an unknown function. This region comprises exons 7 and 8. Exon 7 has no variant considered damaging by all the 10 prediction algorithms combined. Exon 8 is the largest exon of the *SLC20A2* gene, coding for 196 amino acids. Although exon 8 has 28 (16.47%) predicted as damaging variants, it's proportion of damaging over total number of variants is very low (2%). This means that only a small part of this large intracellular portion is functionally important.

The number of PFBC-linked variants for *SLC20A2* is constantly being updated. Up until now there are over 50 variants, with varying effects on the protein and on the patient's phenotype. Since frameshift and truncating mutations are too drastic, missense variants are more interesting to assess the importance of specific residues and protein regions for functionality. Of the current PFBC-linked variants, there are 25 missense variants. These missense variants are limited to exons 2, 3, 5, 8, 9, 10 and 11. Exons 2 and 10 have the same number of variants (6, 22.22%), suggesting that these exons code functionally important residues for PiT2. In our analysis, exon 2, with low proportion of damaging variants, was considered as highly tolerant for missense variations, interpreted as of with lesser importance for protein function.

Exons 4, 6 and 7 still yet have no missense variant associated to PFBC. While this is expected for exons 6 and 7 for they have the lowest number of predicted as damaging variants in our analysis, this is surprising for exon 4. Exon 4 is predicted as having a high importance for protein function, with over 15% proportion of predicted as damaging over total number of annotations. These discrepancies between our prediction model and the actual patient data can be due to a yet small number of PFBC patients with missense mutations. If other types of variants were included, such as frameshift and stop codon mutations, which are much more damaging to the protein structure and function, there still would be no variant in exon 6, only 2 variants for exon 7, one frameshift and one stop codon, and 3 variants for exon 4, two stop codons and one frameshift[7,11,12].

Conclusion

A variant susceptibility profile was drawn for *SLC20A2* gene by using data from 10 prediction algorithms combined. Exons 3 and 10 are the most variant susceptible exons, each exon encoding transmembrane regions at opposite ends of the protein, but without direct function associated to these regions. Exon 10 has, along with exon 2, the highest number of missense variants associated to PFBC. Exon 2 was found to have a low number of variants predicted as damaging (3, 1.76%) while the patient's data show 6 (19.23%) of PFBC-linked variants located in this exon. Exon 7 was found to have no predicted as damaging variants in our analysis, and no missense variant in this exon was linked to PFBC until now. Although some differences were found between the real data and the prediction algorithms, we believe this sort of gene profiling to be useful for increasing our comprehension of this and other genes. It's important to notice that only missense variants were accounted in this analysis, in an attempt to investigate the importance of each position in the gene.

Conflict of interest

The authors of this paper report no conflict of interest.

References

[1] Collins, J. F., Bai, L. & Ghishan, F. K. The SLC20 family of proteins: dual functions as sodium-phosphate cotransporters and viral receptors. Pflugers Arch. 2004;447: 647–652.

[2] Virkki, L. V, Biber, J., Murer, H. & Forster, I. C. Phosphate transporters: a tale of two solute carrier families. Am. J. Physiol. Renal Physiol. 2007;293: F643-54.

[3] Kavanaugh, M. P. & Kabat, D. Identification and characterization of a widely expressed phosphate transporter/retrovirus receptor family. Kidney Int. 1996;49: 959–963.

[4] Bottger, P. & Pedersen, L. Evolutionary and experimental analyses of inorganic phosphate transporter PiT family reveals two related signature sequences harboring highly conserved aspartic acids critical for sodium-dependent phosphate transport function of human PiT2. FEBS J. 2005;272: 3060–3074.

[5] Wang, C. et al. Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis. Nat. Genet. 2012;44: 254–256.

[6] Lemos, R. R. et al. Update and Mutational Analysis of SLC20A2: A Major Cause of Primary Familial Brain Calcification. Hum. Mutat. 2015;36: 489–495.

[7] Hsu, S. C. et al. Mutations in SLC20A2 are a major cause of familial idiopathic basal ganglia calcification. Neurogenetics 2013;14: 11–22.

[8] Liu, X. et al. Novel mutation of SLC20A2 in a Chinese family with primary familial brain calcification. Journal of the neurological sciences 2016;360: 1–3.

[9] Liu, X., Wu, C., Li, C. & Boerwinkle, E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Hum. Mutat. 2016;37: 235–241.

[10] Chien, M. L., Foster, J. L., Douglas, J. L. & Garcia, J. V. The amphotropic murine leukemia virus receptor gene encodes a 71-kilodalton protein that is induced by phosphate depletion. J. Virol. 1997;71: 4564–4570.

[11] Zhang, Y., Guo, X. & Wu, A. Association between a novel mutation in SLC20A2 and familial idiopathic basal ganglia calcification. PLoS One 2013;8: e57060.

[12] Nicolas, G. et al. Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. Neurology 2013;80: 181–187 (a).

[13] Chen, W.-J. et al. Novel SLC20A2 mutations identified in southern Chinese patients with idiopathic basal ganglia calcification. Gene 2013;529: 159–162.

[14] Nicolas, G. et al. Phenotypic spectrum of probable and genetically-confirmed idiopathic basal ganglia calcification. Brain 2013;136: 3395–3407 (b).

[15] Yamada, M. et al. Evaluation of SLC20A2 mutations that cause idiopathic basal ganglia calcification in Japan. Neurology 2014;82: 705–712.

[16] Rubino, E. et al. Three novel missense mutations in SLC20A2 associated with idiopathic basal ganglia calcification. Journal of the neurological sciences 2017;377: 62–64.

[17] Koyama, S. et al. Clinical and radiological diversity in genetically confirmed primary familial brain calcification. Sci. Rep. 2017;7: 12046.

[18] Kasuga, K. et al. A Japanese family with idiopathic basal ganglia calcification with novel SLC20A2 mutation presenting with late-onset hallucination and delusion. Journal of neurology 2014;261: 242–244.

[19] Taglia, I. et al. Primary familial brain calcification: Genetic analysis and clinical spectrum. Mov. Disord. 2014;29: 1691–1695.

[20] Lemos, R. R., Oliveira, M. F. & Oliveira, J. R. M. Reporting a new mutation at the SLC20A2 gene in familial idiopathic basal ganglia calcification. European journal of neurology 2013;20: e43-4.

[21] Kumar, N. & Jog, M. Fahr's Disease Presenting as Late-Onset Levodopa-Responsive Parkinsonism. Can. J. Neurol. Sci. 2017;44: 322–323.

[22] Brighina L, Saracchi E, Ferri F, Gagliardi M, Tarantino P, Morzenti S, Musarra M, Patassini M, Annesi G, Ferrarese C. Fahr's disease linked to a novel SLC20A2 gene mutation manifesting with dynamic aphasia. Neurodegener Dis 2014;14: 133–118.

[23] Fjaer, R. et al. Generalized epilepsy in a family with basal ganglia calcifications and mutations in SLC20A2 and CHRNB2. Eur. J. Med. Genet. 2015;58: 624–628.

[24] Schottlaender LV, Mencacci N, Koepp M, Hanna M, Hardy J, Lees AJ, Houlden H. Interesting clinical features associated with mutations in the SLC20A2 gene. Eur. J. Neurol. 2012;19: 40.

[25] Kimura, T. *et al.* Familial idiopathic basal ganglia calcification: Histopathologic features of an autopsied patient with an SLC20A2 mutation. *Neuropathology : official journal of the Japanese Society of Neuropathology* 2016;36**:** 365–371.

Figure legends

Figure 1. *SLC20A2* variant damage predicted by SIFT algorithm by exon. The damage score for each variation is depicted, and a trend line has been drawn based on the distribution of these data, allowing a more visual analysis of the score distributions. Exons 2, 3, 9, 10 and 11 have the highest number of predicted damaging variants, while exons 6 and 7 have the smallest number of these variants, predicted by SIFT algorithm.

Figure 2. Scheme representing PiT2's genetic structure in the cell's transmembrane space, separated by exons (A). 12 transmembrane domains are represented, and each exon is painted as a different color. A linear representation of each exon position of PiT2 protein was also included, with the number of predicted as damaging variants over the total number of variants annotated in dbSNFP. Arrows represent PFBC-linked *SLC20A2* missense variants, shown at its approximate position in the protein (red – transmembrane, white – cytoplasmic, blue – extracellular). GANDVANA sequences are represented as a yellow triangle at C and N terminal regions. PHO4 superfamily

homologous sequences are also represented with the orange rectangle. B) A pie graph representing each exon contribution to the total number of predicted as damaging missense variants for *SLC20A2*.

Figure S1. Linear regression model comparing number of variants predicted as damaging over the total number of variants per exon of *SLC20A2* gene. The data fits very weakly into a linear model, with slope of 0.01383. This translates into how the number of damaging variants increases with each increase in the total number of variants. Pearson correlation was also drawn for the same data, resulting in a coefficient (r^2) of 0.2688, a weak positive correlation.

Figure S2. Proportion of variants predicted as damaging by all the 10 prediction algorithms. A) Damaging over total number of variants for *SLC20A2* gene, comparing its exons. The median of proportions of damaging over total number of variants for this gene is ~10%. B) Proportion of damaging by total number of variants for *SLC20A2* gene by exon. Exons 3, 4 and 5 are revealed as having the highest proportion of variants predicted as damaging by all the 10 used algorithms, followed by exon 10. Exons 2, 6 and 7 have a very low proportion of damaging variants in relation to the total.

Tables

Table 1. *SLC20A2* damage prediction profile by exon.

Exon	*Total*	*Tolerable*	*Damaging*	*D/T*	*% for SLC20A2*
2	661	651	10	0.015128593	0.035587189
3	316	267	49	0.155063291	0.174377224
4	199	169	30	0.150753769	0.106761566
5	213	181	32	0.150234742	0.113879004
6	265	261	4	0.01509434	0.014234875
7	457	457	0	0	0
8	1358	1308	50	0.036818851	0.177935943
9	417	378	39	0.09352518	0.138790036
10	188	163	25	0.132978723	0.088967972
11	368	326	42	0.114130435	0.149466192

Table 2. *SLC20A2* missense variants linked to Primary Familial Brain Calcification.

N	*Amino acid position*	*Exon*	*Change*	*Reference*
1	11	2	**p.Ile11Leu**	Chen *et al*[13]
2	28	2	**p.Asp28Asn**	Nicolas *et al*[14] Chen *et al*[13]
3	51	2	**p.Ala51Val**	Yamada *et al*[15]
4	62	2	**p.Leu62Pro**	Chen *et al*[13]
5	63	2	**p.Gly63Asp** **p.Gly63Ser**	Rubino *et al*[16]
6	71	2	**p.Arg71His**	Yamada *et al*[15]
7	90	2	**p.Gly90Val**	Koyama *et al*[17]
8	108	3	**p.Leu108Pro**	Kasuga *et al*[18]
9	115	3	**p.Thr115Met**	Yamada *et al*[15]
10	184	5	**p.Pro184Leu**	Nicolas et al[12]

11	187	5	**p.Tyr187Cys**	Lemos et al[6]
12	194	5	**p.Asn194Ser**	Nicolas *et al*[14]
13	382	8	**p.Arg382Gln**	Hsu *et al*[7]
14	399	8	**p.His399Pro**	Rubino *et al*[16]
15	434	8	**p.Ser434Trp**	Taglia *et al*[19]
16	495	8	**p.Ala495Thr**	Lemos *et al*[20]
17	498	8	**p.Gly498Arg**	Wang *et al*[5]
18	502	8	**p.His502Gln**	Hsu *et al*[7]
19	503	8	**p.Gly503Ser**	Kumar and Jog[21]
20	540	9	**p.Gly540Arg**	Brighina *et al*[22]
21	568	9	**p.Pro568Leu**	Hsu *et al*[7] Fjaer *et al*[23]
22	571	9	**p.Gly571Ser**	Nicolas *et al*[14]
23	575	10	**p.Glu575Lys**	Wang *et al*[5] Schottlaender *et al*[24]
24	585	10	**p.Ala585Thr**	Lemos *et al*[6]
25	595	10	**p.Thr595Met**	Wang *et al*[5] Taglia *et al*[19]
26	597	10	**p.Cys597Tyr**	Liu *et al*[8]
27	601	11	**p.Ser601Leu** **p.Ser601Trp** **p.Ser601Leu**	Wang *et al*[5] Hsu *et al*[7]
28	637	11	**p.Ser637Arg**	Yamada *et al*[15] Kimura *et al*[25]

Table S1. *SLC20A2* chromosome positions

Gene	*Chromosome range*	*Exon*	*E. Chromosome range*
SLC20A2	chr8: 42416462-42501460	1	42501031-42501460
		2	42472102-42472654
		3	42465777-42465918
		4	42463005-42463095
		5	42459896-42459991
		6	42444646-42444762
		7	42439450-42439653
		8	42436989-42437577
		9	42430064-42430249
		10	42428758-42428842
		11	42416475-42417967

Figures

Figure 1

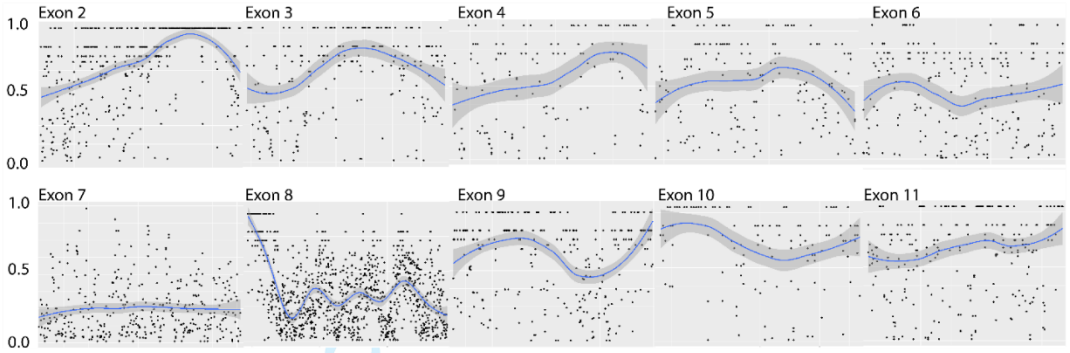


Figure 2

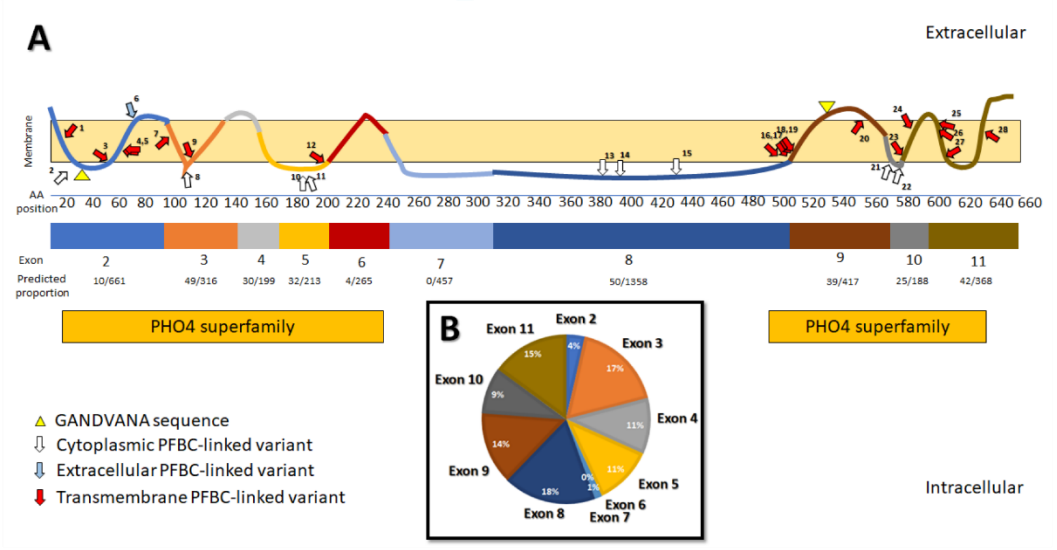


Figure S1

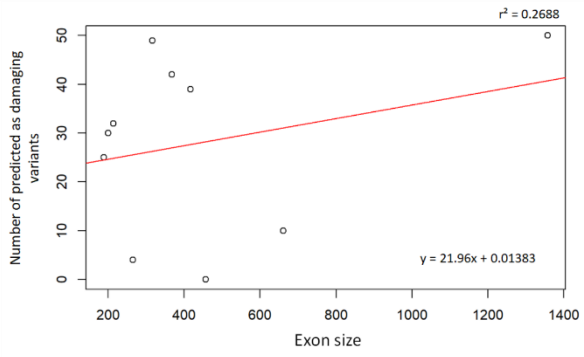
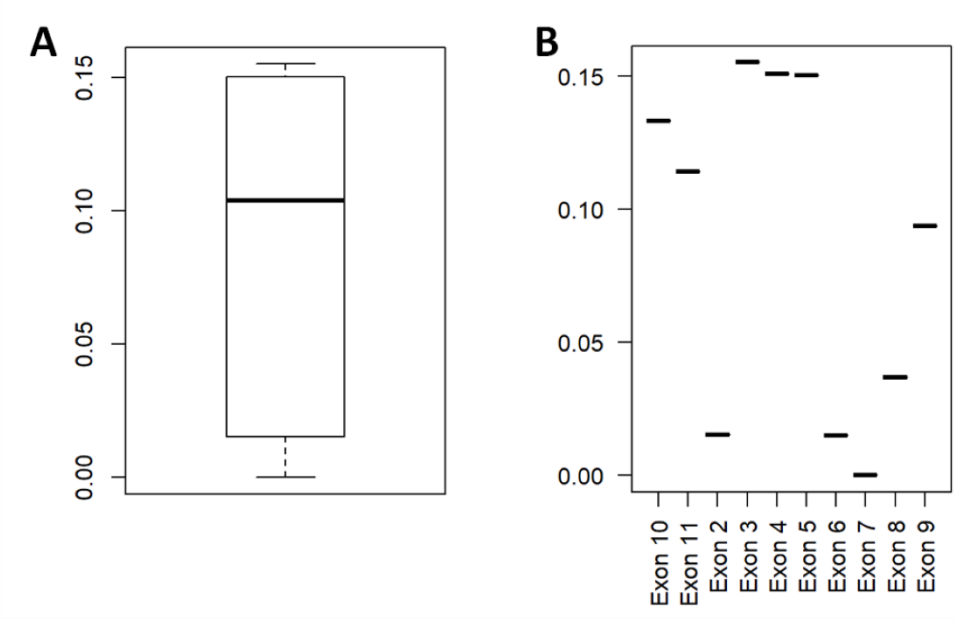

r² = 0.2688
Number of predicted as damaging variants
y = 21.96x + 0.01383
Exon size

Figure S2


A
B
0.00
0.05
0.10
0.15
Exon 10
Exon 11
Exon 2
Exon 3
Exon 4
Exon 5
Exon 6
Exon 7
Exon 8
Exon 9

4.2 *SCRIPT* NA LINGUAGEM DE PROGRAMAÇÃO R

- Carregando os dados do dbSNFP.

```{r}
allslc20a2 = read.csv("C:/Users/Denis/Documents/dbNSFP/slc20a2-2.csv", h = F, sep = ",")
View(allslc20a2)
```

- Determinando as posicoes referentes aos Exons.

```{r}
pit2exon1 = allslc20a2[allslc20a2$V3 %in% c(42501031:42501460),]
pit2exon2 = allslc20a2[allslc20a2$V3 %in% c(42472102:42472654),]
pit2exon3 = allslc20a2[allslc20a2$V3 %in% c(42465777:42465918),]
pit2exon4 = allslc20a2[allslc20a2$V3 %in% c(42463005:42463095),]
pit2exon5 = allslc20a2[allslc20a2$V3 %in% c(42459896:42459991),]
pit2exon6 = allslc20a2[allslc20a2$V3 %in% c(42444646:42444762),]
pit2exon7 = allslc20a2[allslc20a2$V3 %in% c(42439450:42439653),]
pit2exon8 = allslc20a2[allslc20a2$V3 %in% c(42436989:42437577),]
pit2exon9 = allslc20a2[allslc20a2$V3 %in% c(42430064:42430249),]
pit2exon10 = allslc20a2[allslc20a2$V3 %in% c(42428758:42428842),]
pit2exon11 = allslc20a2[allslc20a2$V3 %in% c(42416475:42417967),]
```

Visualizando os resultados das predicoes por exon utilizando o SIFT.

```{r}
slc20a2 = read.csv("C:/Users/Denis/Documents/dbNSFP/slc20a2-2.csv", h = F, sep = ",")
slc20a2[] <- lapply(slc20a2, function(x) {
  if(is.factor(x)) as.numeric(as.character(x)) else x
}) #Transforma a coluna em valores numéricos e mantém os decimais.
exon1 = slc20a2[slc20a2$V3 %in% c(42501031:42501460),]
exon2 = slc20a2[slc20a2$V3 %in% c(42472102:42472654),]
exon3 = slc20a2[slc20a2$V3 %in% c(42465777:42465918),]
exon4 = slc20a2[slc20a2$V3 %in% c(42463005:42463095),]
exon5 = slc20a2[slc20a2$V3 %in% c(42459896:42459991),]
exon6 = slc20a2[slc20a2$V3 %in% c(42444646:42444762),]
exon7 = slc20a2[slc20a2$V3 %in% c(42439450:42439653),]
exon8 = slc20a2[slc20a2$V3 %in% c(42436989:42437577),]
exon9 = slc20a2[slc20a2$V3 %in% c(42430064:42430249),]
exon10 = slc20a2[slc20a2$V3 %in% c(42428758:42428842),]
exon11 = slc20a2[slc20a2$V3 %in% c(42416475:42417967),]
library(ggplot2)
qplot(exon2$V3, exon2$V26, main = "SIFT values for SLC20A2 exon 2", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon3$V3, exon3$V26, main = "SIFT values for SLC20A2 exon 3", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon4$V3, exon4$V26, main = "SIFT values for SLC20A2 exon 4", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon5$V3, exon5$V26, main = "SIFT values for SLC20A2 exon 5", xlab = "Chr position", ylab = "SIFT
value", geom = c("point", "smooth"))
qplot(exon6$V3, exon6$V26, main = "SIFT values for SLC20A2 exon 6", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon7$V3, exon7$V26, main = "SIFT values for SLC20A2 exon 7", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon8$V3, exon8$V26, main = "SIFT values for SLC20A2 exon 8", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon9$V3, exon9$V26, main = "SIFT values for SLC20A2 exon 9", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
qplot(exon10$V3, exon10$V26, main = "SIFT values for SLC20A2 exon 10", xlab = "Chr position", ylab = "SIFT
value",geom = c("point", "smooth"))
```

```
qplot(exon11$V3, exon11$V26, main = "SIFT values for SLC20A2 exon 11", xlab = "Chr position", ylab = "SIFT value",geom = c("point", "smooth"))
```

- Determinando posicoes consideradas danosas por todos os metodos de predicao.

```{r}
#SIFT
danoso = allslc20a2[allslc20a2$V27 == "D",]
#Polyphen2 HDIV
danoso = danoso[allslc20a2$V33 == "D",]
#Polyphen2 HVAR
danoso = danoso[danoso$V36 == "D",]
#LRT
danoso = danoso[danoso$V39 == "D",]
#MutationTASTER
danoso = danoso[danoso$V43 == "D;D;D",]
#FATHHM
danoso = danoso[(danoso$V53 == "D;D;D"),]
#PROVEAN
danoso = danoso[danoso$V56 == "D",]
#MetaSVM
danoso = danoso[danoso$V63 == "D",]
#MetaLR
danoso = danoso[danoso$V66 == "D",]
#FATHHM-MKL
danoso = danoso[danoso$V75 == "D",]
```

E variacoes preditas como nao danosas.

```{r}
length(allslc20a2$V1)
head(allslc20a2$V1)
#Subtraindo as variacoes preditas como danosas do total.
naodanoso = allslc20a2[!(danoso$V1 %in% allslc20a2$V1),]
```

Contabilizando variacoes danosas e nao danosas por exon.

```{r}
#Para o Exon 2
exon2dam=pit2exon2[(pit2exon2$V1 %in% danoso$V1),]
exon2tol=pit2exon2[!(pit2exon2$V1 %in% danoso$V1),]
#Exon 3
exon3dam=pit2exon3[(pit2exon3$V1 %in% danoso$V1),]
exon3tol=pit2exon3[!(pit2exon3$V1 %in% danoso$V1),]
#Exon 4
exon4dam=pit2exon4[(pit2exon4$V1 %in% danoso$V1),]
exon4tol=pit2exon4[!(pit2exon4$V1 %in% danoso$V1),]
#Exon 5
exon5dam=pit2exon5[(pit2exon5$V1 %in% danoso$V1),]
exon5tol=pit2exon5[!(pit2exon5$V1 %in% danoso$V1),]
#Exon 6
exon6dam=pit2exon6[(pit2exon6$V1 %in% danoso$V1),]
exon6tol=pit2exon6[!(pit2exon6$V1 %in% danoso$V1),]
#Exon 7
exon7dam=pit2exon7[(pit2exon7$V1 %in% danoso$V1),]
exon7tol=pit2exon7[!(pit2exon7$V1 %in% danoso$V1),]
#Exon 8
exon8dam=pit2exon8[(pit2exon8$V1 %in% danoso$V1),]
exon8tol=pit2exon8[!(pit2exon8$V1 %in% danoso$V1),]
```

```
#Exon 9
exon9dam=pit2exon9[(pit2exon9$V1 %in% danoso$V1),]
exon9tol=pit2exon9[!(pit2exon9$V1 %in% danoso$V1),]
#Exon 10
exon10dam=pit2exon10[(pit2exon10$V1 %in% danoso$V1),]
exon10tol=pit2exon10[!(pit2exon10$V1 %in% danoso$V1),]
#Exon 11
exon11dam=pit2exon11[(pit2exon11$V1 %in% danoso$V1),]
exon11tol=pit2exon11[!(pit2exon11$V1 %in% danoso$V1),]
```

Comparando variacoes entre exons.

```{r}
#Calculando proporcoes de variacoes danosas por Exon.
propdam2 = length(exon2dam$V1)/length(pit2exon2$V1)
propdam3 = length(exon3dam$V1)/length(pit2exon3$V1)
propdam4 = length(exon4dam$V1)/length(pit2exon4$V1)
propdam5 = length(exon5dam$V1)/length(pit2exon5$V1)
propdam6 = length(exon6dam$V1)/length(pit2exon6$V1)
propdam7 = length(exon7dam$V1)/length(pit2exon7$V1)
propdam8 = length(exon8dam$V1)/length(pit2exon8$V1)
propdam9 = length(exon9dam$V1)/length(pit2exon9$V1)
propdam10 = length(exon10dam$V1)/length(pit2exon10$V1)
propdam11 = length(exon11dam$V1)/length(pit2exon11$V1)
propexons =
c(propdam2,propdam3,propdam4,propdam5,propdam6,propdam7,propdam8,propdam9,propdam10,propdam11)
nomes = c("Exon 2", "Exon 3", "Exon 4", "Exon 5", "Exon 6", "Exon 7", "Exon 8", "Exon 9", "Exon 10", "Exon
11")
names(propexons) = nomes
propexons
dataprop = data.frame(propexons)
nomes = c("Exon 2","Exon 3", "Exon 4", "Exon 5", "Exon 6", "Exon 7", "Exon 8", "Exon 9", "Exon 10", "Exon
11")
#Anotando os valores totais por Exon.
total = c(length(pit2exon2$V1), length(pit2exon3$V1), length(pit2exon4$V1), length(pit2exon5$V1),
length(pit2exon6$V1), length(pit2exon7$V1), length(pit2exon8$V1), length(pit2exon9$V1),
length(pit2exon10$V1), length(pit2exon11$V1))
names(total) = nomes
total
#Anotando os valores danosos por Exon.
danos =
c(length(exon2dam$V1),length(exon3dam$V1),length(exon4dam$V1),length(exon5dam$V1),length(exon6dam$V1
),length(exon7dam$V1),length(exon8dam$V1),length(exon9dam$V1),length(exon10dam$V1),length(exon11dam$
V1))
names(danos) = nomes
danos
#Anotando os valores toleraveis por Exon.
toleravel =
c(length(exon2tol$V1),length(exon3tol$V1),length(exon4tol$V1),length(exon5tol$V1),length(exon6tol$V1),length(
exon7tol$V1),length(exon8tol$V1),length(exon9tol$V1),length(exon10tol$V1),length(exon11tol$V1))
names(toleravel) = nomes
toleravel
#Juntando todos os dados em um dataframe.
tudo = data.frame(total,danos,toleravel)
tudo
tudo$proporcao = propexons
View(tudo)
```

```
#Anotando os valores para o gene inteiro
danosslc20a2 = sum(tudo$danos)
danosslc20a2
toleravelslc20a2 = sum(tudo$toleravel)
#Grafico de pizza para proporcoes de anotacoes de variantes danosas por Exon para o gene SLC20A2
pie(danos)
#Grafico de pizza para tamanho total de Exons do gene SLC20A2.
pie(total)
#Visualizando os dados de predicao em um grafico de dispersao.
plot(propexons,xlab="Exons 2-11",main="Proporcao de variantes preditas danosas/total para o gene
SLC20A2",ylab="Proporcao")
#Comparando os dados reais com a proporcao.
plot(propexons,danos,xlab="Proporcao",ylab="Numero de variantes consideradas danosas",main="Comparacao do
numero de variantes danosas com a proporcao")
#Comparando variantes totais por proporcao das consideradas danosas.
plot(tudo$proporcao,tudo$total)
plot(tudo$danos,tudo$total)
#Testando se esses dados seguem uma distribuicao linear e calculando a correlacao entre tamanho do exon e numero
de variantes danosas.
linear = lm(danos~total,data=tudo)
linear #Aqui percebemos a intercessao de 14.34 e inclinacao de 0.005983 (quanto que o numero de variantes
consideradas danosas aumenta para cada numero de variantes anotadas)
plot(tudo$total,tudo$danos)
abline(linear,col="red")
cor(tudo$danos,tudo$total) #Uma correlacao de 0.1617 que pode ser traduzida como fraca ou nao correlacionada.
#Visualizando os dados de predicao em um histograma
par(mfrow=c(1,2))
hist(danos,main="Variantes consideradas danosas por Exon",xlab="Variantes danosas",ylab="Frequencia")
hist(propexons,main="Variantes danosas/Total de anotacoes",xlab="Proporcoes",ylab="Frequencia")
#Visulaizando os dados de predicao de danos para o gene SLC20A2 em um boxplot.
boxplot(propexons, main="Proporcao de variantes danosas por exon para o gene SLC20A2",ylab="Proporcao
Danoso/Total") #O boxplot mostra que a mediana se encontra em ~5% (metade dos exons tem 5% de anotacoes
consideradas muito danosas)
#Visualizando os dados de predicao em graficos.
tudo$Exon = c("Exon 2", "Exon 3", "Exon 4", "Exon 5", "Exon 6", "Exon 7", "Exon 8", "Exon 9", "Exon 10",
"Exon 11")
boxplot(tudo$danos~tudo$total,main="Comparacao das anotacoes, por exon, para o gene
SLC20A2",ylab="Anotacoes consideradas danosas",xlab=("Todas as anotacoes"))
boxplot(tudo$danos~tudo$Exon,las=2,main="Comparacao das anotacoes danosas por exon para o gene
SLC20A2",ylab="Anotacoes consideradas danosas")
dataprop = data.frame(nomes,propexons)
dataproptest = data.frame(nomes,(propexons*100))
dataprop
plot(dataprop,las=2,main="Proporcoes de anotacoes para o gene SLC20A2",ylab="Anotacoes
danosas/total",xlab=NA)
dataproptest
names(dataproptest) = c("Exon", "Proporcao")
dataprop
#Testando se a media das variantes danosas e diferente de 50%.
t.test(dataprop$propexons,mu=0.1,alternative="less") #Com esse P-value, a proporcao de variantes consideradas
danosas encontrada pros Exons  e inferior a 10%.
#Testando se o numero de variacoes danosas estao igualmente distribuidas entre os Exons do gene SLC20A2
media = length(danoso$V1)/10
media
#Teste de Student comparando os valores reais de variantes danosas por Exon e a proporcao de variantes danosas
pelo total de anotacoes por exon.
```

```
t.test(danos) #P-value confirma que os numeros de variantes e estatisticamente diferente entre os exons.
t.test(propexons) #P-value mostra que as proporcoes sao muito diferentes entre os exons.
```

5 CONCLUSÕES

Durante os 24 meses referentes ao mestrado, entre Março de 2016 e Fevereiro de 2018, foi possível realizar uma grande quantidade de pesquisas científicas, acumulando uma enorme quantidade de conhecimento e construindo uma formação acadêmica sólida. Foram concluídos 24 créditos (360 horas) em disciplinas pelo Programa de Pós-Graduação em Biologia Aplicada à Saúde, obtendo nota máxima em todas as disciplinas cursadas, finalizando o curso com média A.

O projeto proposto foi finalizado com êxito, resultando em uma submissão para publicação em periódico internacional, o *Proteins: Structure, Function, and Bioinformatics* (fator de impacto 2,289) e um *script* para análises de perfil de genes e suas respectivas proteínas em relação a áreas de maior ou menor susceptibilidade a variações genéticas.

As atividades complementares realizadas durante o período do mestrado permitiram uma vivência acadêmica considerável, resultando em publicações em periódicos com alto fator de impacto - incluindo uma carta à prestigiosa revista *Science*; uma co-autoria em artigo completo na revista *Neurology: Genetics*, com autores de referência na neurogenética da CCFP; e um artigo como primeiro autor na revista *Journal of Molecular Neuroscience*. Além disso, foram obtidas também premiações em anos consecutivos como segunda melhor apresentação oral no simpósio internacional anual promovido pelo Laboratório de Imunopatologia Keizo Asami, o Simpósio Internacional em Diagnóstico e Terapêutica (SINATER), edições de 2016 e 2017.

O conjunto de atividades realizadas durante o mestrado forneceram uma sólida base acadêmica para realização de um projeto maior durante o doutorado. Projeto de doutorado já aprovado para início em Março de 2018, pelo Programa de Pós-Graduação em Biologia Aplicada à Saúde (Laboratório de Imunopatologia Keizo Asami) da Universidade Federal de Pernambuco.

REFERÊNCIAS

Collins, J. F., Bai, L. & Ghishan, F. K. The SLC20 family of proteins: dual functions as sodium-phosphate cotransporters and viral receptors. Pflugers Arch. 447, 647–652, 2004.

Ferreira, J. B. et al. First report of a de novo mutation at SLC20A2 in a patient with brain calcification. J. Mol. Neurosci. 54, 748–751, 2014.

Fujioka, Y. et al. FUS-regulated region- and cell-type-specific transcriptome is associated with cell selectivity in ALS/FTLD. Sci. Rep. 3, 2388, 2013.

Gianchelli, C. M.; Jono, S.; Shioi, A.; Nishizawa, Y.; Mori, K.; e Morii, H. Vascular calcification and inorganic phosphate. American Journal of Kidney Diseases, 38(4), S34-37, 2001.

Griffiths, A. J. F. Introduction to genetic analysis (Tenth Edition) W.H. Freeman and Co., New York, 2012.

Hsu, P. D., Lander, E. S., Zhang, F. Development and applications of CRISPR-Cas9 for genome engineering. Cell, 157, pp. 1262–1278, 2014.

Hsu, S. C. et al. Mutations in SLC20A2 are a major cause of familial idiopathic basal ganglia calcification. Neurogenetics 14, 11–22, 2013.

Kavanaugh, M. P. & Kabat, D. Identification and characterization of a widely expressed phosphate transporter/retrovirus receptor family. Kidney Int. 49, 959–963, 1996.

Keasey, M. P., Lemos, R. R., Hagg, T. & Oliveira, J. R. M. Vitamin-D receptor agonist calcitriol reduces calcification in vitro through selective upregulation of SLC20A2 but not SLC20A1 or XPR1. Sci. Rep. 6, 25802, 2016.

Keller, A.; Westenberger, A.; Sobrido, M. J.; et al. Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. Nature Genetics, 45, 1077–1082, 2013.

Legati, A; Giovannini, D; Nicolas, G; et al. Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. Nature Genetics, 2015.

Lemos, R. R.; Ferreira, J. B. M. M.; Keasey, M. P.; e Oliveira, J. R. M. An update on primary familial brain calcification. Metal Related Neurodegenerative Disease, cap. 14, 2013.

Liu X, Jian X, Boerwinkle E. dbNSFP: a lightweight database of human nonsynonymous SNPs and their functional predictions. Hum Mutat 2011; 32: 894–899.

Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Hum Mutat 2016; 37: 235–241.

Livingston, J. H., Stivaros, S., van der Knaap, M. S. & Crow, Y. J. Recognizable phenotypes associated with intracranial calcification. Dev. Med. Child Neurol. 55, 46–57, 2013.

Lowe WL, Reddy TE. Genomic approaches for understanding the genetics of complex disease. Genome Res 2015; 25: 1432–1441.

Mackenzie, N. C. W. et al. MOVAS-1 cell line: a new in vitro model of vascular calcification. Int. J. Mol. Med. 27, 663–668, 2011.

Mahy, N., Prats, A., Riveros, A., Andres, N. & Bernal, F. Basal ganglia calcification induced by excitotoxicity: an experimental model characterised by electron microscopy and X-ray microanalysis. Acta Neuropathol. 98, 217–225, 1999.

Manavalan, A. et al. Brain site-specific proteome changes in aging-related dementia. Exp. Mol. Med. 45, e39, 2013.

Middleton FA, Rosenow C, Vailaya A, Kuchinsky A, Pato MT, Pato CN. Integrating genetic, functional genomic, and bioinformatics data in a systems biology approach to complex diseases: application to schizophrenia. Methods Mol Biol 401: 337–364, 2007.

Moura, D. A. P. & Oliveira, J. R. M. XPR1: a Gene Linked to Primary Familial Brain Calcification Might Help Explain a Spectrum of Neuropsychiatric Disorders. J. Mol. Neurosci. 57, 519–521, 2015.

Moura, D. A. P., Lemos, R., Oliveira, J. R. M. New Data from Pdfgbret/ret Mutant Mice Might Lead to a Paradoxical Association Between Brain Calcification, Pericytes Recruitment and BBB Integrity. J. Mol. Neurosci. Online only. 2017. DOI: 10.1007/s12031-017-0992-z

Nicolas, G.; Pottier, C.; Maltête, D.; et al. Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. Neurology, 2013.

NITSCHKE, Y.; e RUTSCH, F. Genetics in arterial calcification: lessons learned from rare diseases. Trends in cardiovascular medicine, 2012.

Oliveira, D. F., de Lemos, R. R. & de Oliveira, J. R. M. Mutations at the SLC20A2 gene and brain resilience in families with idiopathic basal ganglia calcification (‘Fahr’s disease’). Front. Hum. Neurosci. 7, 420, 2013.

Oliveira, J. R. M. & Steinberg, S. S. Far from rare: revisiting the relevance of idiopathic basal ganglia calcifications. Neurological sciences : official journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology 31, 679, 2010.

Oliveira, J. R. M. Managing idiopathic basal ganglia calcification (“Fahr’s disease”), primeira edição. Nova Publishing, New York, 2011.

Platt, R. J. et al. CRISPR-Cas9 Knockin Mice for Genome Editing and Cancer Modeling. Cell 159, 440–455, 2014.

Rodriguez, M. J., Pugliese, M. & Mahy, N. Drug abuse, brain calcification and glutamate-induced neurodegeneration. Curr. Drug Abuse Rev. 2, 99–112, 2009.

Schlachetzki JCM, Saliba SW, Oliveira ACP de. Studying neurodegenerative diseases in culture models. Rev. Bras. Psiquiatr. 35: S92–S100, 2013.

Sobrido, M. J.; Coppola, G.; Oliveira, J. R. M.; Hopfer, S.; e Geschwind, D. H. Primary Familial Brain Calcification. 2004 Apr 18 [Updated 2014 Oct 16]. In: Pagon RA, Adam MP, Ardinger HH, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2015.

Vanlandewijck, M. et al. Functional Characterization of Germline Mutations in PDGFB and PDGFRB in Primary Familial Brain Calcification. PLoS One 10, e0143407, 2015.

Villasenor, R. et al. Region-specific permeability of the blood-brain barrier upon pericyte loss. J. Cereb. Blood Flow Metab. 271678X17697340, 2017.

Virkki, L. V, Biber, J., Murer, H. & Forster, I. C. Phosphate transporters: a tale of two solute carrier families. Am. J. Physiol. Renal Physiol. 293, F643-54, 2007.

Wang, C.; Li, Y.; Shi, L.; Ren, J.; Patti, M.; et al. Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis. Nature Genetics, 44:254–256, 2012.

Xu, L. M. et al. AutismKB: an evidence-based knowledgebase of autism genetics. Nucleic Acids Res. 40, 2012.

Yamada, M. et al. High frequency of calcification in basal ganglia on brain computed tomography images in Japanese older adults. Geriatr. Gerontol. Int. 13, 706–710, 2013.

Zhang, X. et al. Human intracellular ISG15 prevents interferon-alpha/beta over-amplification and auto-inflammation. Nature 517, 89–93, 2015.

APÊNDICE A – CARTA PUBLICADA EM PERIÓDICO: ''Clarifying samples in Zika analyses'

PERIÓDICO *SCIENCE*

Fator de impacto (JCR 2011): 31,201
Qualis CAPES Ciências Biológicas I: A1

Autores: João R. M. de Oliveira* and Denis A. P. Moura Laboratório de Imunopatologia Keizo Asami, Federal University of Pernambuco, Recife, Pernambuco, 50670-901, Brazil. *Corresponding author. Email: joao.ricardo@ufpe.br

Título: Clarifying samples in Zika analyses

LETTERS

Edited by **Jennifer Sills**

Clarifying samples in Zika analyses

IN RESPONSE TO the Zika outbreak and putatively related microcephaly cases in Brazil, many research groups in Brazil, North America, and Europe are studying the virus ("Evidence grows for Zika virus as pregnancy danger," G. Vogel, In Depth, 11 March, p. 1123). The simultaneous efforts to address this urgent matter have led to the use of some samples in multiple studies. There is no clear leadership or protocol to regulate access to patients, some of whom are also participants in studies of potential microcephaly causes unrelated to Zika, such as cytomegalovirus and genetic inbreeding (*1*). Some studies are reporting different results for the same set of samples. For example, Calvet *et al.* (*2*) and Oliveira Melo *et al.* (*3*) acknowledge overlap in the methods section. It is difficult to assess studies if we cannot determine whether, or to what degree, the cohorts are independent. Such confusion has already necessitated clarification of similar studies from different groups (*4*).

To address this problem, the World Health Organization could create a universal code for each baby with confirmed microcephaly and Zika infection, to be used by all research groups working with these data. The codes should include unique identifiers, generated by a competent government agency, that indicate the country and institution of diagnosis as well as a serial number for each patient. The World Health Organization could also provide a public database for Zika cases that would include the Zika codes as well as epidemiological information. These steps would allow individual cases to be identified in multiple studies while protecting privacy.

João R. M. de Oliveira* and Denis A. P. Moura

Laboratório de Imunopatologia Keizo Asami, Federal University of Pernambuco, Recife, Pernambuco, 50670-901, Brazil.

*Corresponding author. Email: joao.ricardo@ufpe.br

REFERENCES

1. G. França *et al.*, *The Lancet* (2016); http://dx.doi.org/10.1016/S0140-6736(16)30902-3.
2. G. Calvet *et al.*, *Lancet Infect. Dis.* **16**, 653 (2016).
3. A. S. Oliveira Melo *et al.*, *Ultrasound Obstet. Gynecol.* **47**, 6 (2016).
4. M. de Fatima Vasco Aragao, *BMJ* **353**, i2444 (2016).

10.1126/science.aah3733

Researchers studying Zika and microcephaly are using overlapping sets of data.

Animal-based antibodies: Obsolete

THE GLOBAL ANTIBODY industry produces an indispensable resource for biological, molecular, and cell scientists. Antibodies are harvested from immunized animals. The animals suffer side effects from the immunizations (*1*) and are, in some cases, mistreated (*2*). It is no longer necessary to compromise animal welfare: Since the mid-1990s, animals have not been required for antibody production (*3*). It is long past time to replace the use of animal-generated antibodies with nonimmunized recombinant antibodies.

Animal-Friendly Affinity reagents (AFAs) are antibodies that are generated by using recombinant technology in viruses or yeasts. The technology allows cloning of immunoglobulin gene segments, to produce antibody libraries with high diversity from which antibodies with desired specificities can be chosen. These are translated on the surfaces of cells or phage particles, and exposed to the target antigen, which selects a highly specific antibody, after which production can be scaled up within cell culture. AFAs are commercially available and can also be developed in individual laboratories. They have wide-ranging applicability as well as specificity and affinity—equal or greater to their animal-generated counterparts—to a huge repertoire of antigens. They also give researchers greater control over antibody properties, generation time, and cost (*4*). Thanks to AFAs, the use of animals has become obsolete.

EU Directive 2010/63/EU (*5*) requires the replacement of animals used in scientific procedures when alternatives exist. Yet, despite the maturation of a growing number of techniques to produce AFAs, antibody production using animals continues to be authorized. Twenty years ago, EU Member States were advised that "in the near future," antibody production "without prior immunization of [animals would] avoid the need to use living animals" (*6*). That prediction was correct. It is incomprehensible that such needless animal use continues.

There is little clarity about how many animals are used to produce antibodies. Only two EU Member State countries have published antibody production numbers. In 2013, the United Kingdom reported use of 9522 animals (*7*), and The Netherlands reported the use of 25,697 animals (*8*). The numbers do not include animals used to produce antibodies that were imported into those countries. Neither country has published the number of animals used for this purpose since 2013.

We recommend the following actions: Antibody production methods that use animal immunization should be replaced in EU Member States. Manufacturers outside the European Union should be required to adhere to European standards to qualify for import to Member States. An expert working group should be established to set up a roadmap for replacement. Programs should be implemented to ensure that animal-friendly antibody producers are fully supported. Subsequent reports from the Commission to the Council and the European Parliament on the statistics on the number of animals used for experimental and other scientific purposes should include data on the use of animals for antibody production as an independent category. These actions should be reinforced through international cooperation and national agencies that can execute government regulation and prevent outsourcing to regions where animal welfare is less well regulated.

A. C. Gray,[1]* S. S. Sidhu,[2] P. C. Chandrasekera,[3] C. F. M. Hendriksen,[4] C. A. K. Borrebaeck[5]

[1]AFABILITY, 74110, Morzine, France. [2]The Donnelly Centre, University of Toronto, Toronto, ON M5S 3E1, Canada. [3]CM Scientia—Life Sciences Consulting, Harrow, ON NOR 1GO, Canada. [4]Faculty of Veterinary Medicine, Utrecht University, 3584 CL Utrecht, Netherlands. [5]CREATE Health Cancer Center, Department of Immunotechnology, Lund University, S-223 81, Lund, Sweden.

*Corresponding author. Email: draligray@live.com

REFERENCES

1. P. P. Leenaars *et al.*, *Lab Anim.* **32**, 387 (1998).
2. S. Reardon, *Nature News* 10.1038/nature.2016.19958 (2016).
3. C. Borebaeck, Ed., *Antibody Engineering, 2nd Edition* (Oxford Univ. Press, 1995).
4. S. Dübel *et al.*, *Trends Biotechnol.* **28**, 333 (2010).
5. Directive 2010/63/EU (http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32010L0063).
6. U. Marx *et al.*, *ATLA* **25**, 121 (1997).
7. Home Office, Annual Statistics of Scientific Procedures on

PHOTO: RAFAEL FABRES/PICTURE-ALLIANCE/DPA/AP IMAGES

APÊNDICE B - CO-AUTORIA EM ARTIGO PUBLICADO EM PERIÓDICO: 'Brain calcifications and PCDH12 variants'

Fator de impacto (JCR 2011): ausente
Qualis CAPES Ciências Biológicas I: ausente

Autores: Nicolas G, Sanchez-Contreras M, Ramos EM, Lemos RR1, Ferreira J, Moura D, Sobrido MJ, Richard AC, Lopez AR, Legati A, Deleuze JF, Boland A, Quenez O, Krystkowiak P, Favrole P, Geschwind DH, Aran A, Segel R, Levy-Lahad E, Dickson DW, Coppola G, Rademakers R, de Oliveira JRM.

Título: Brain calcifications and PCDH12 variants

Brain calcifications and *PCDH12* variants

OPEN

Gaël Nicolas, MD, PhD
Monica Sanchez-Contreras, MD, PhD
Eliana Marisa Ramos, PhD
Roberta R. Lemos, PhD
Joana Ferreira, PhD
Denis Moura, BSc
Maria J. Sobrido, MD, PhD
Anne-Claire Richard, BSc
Alma Rosa Lopez, BSc
Andrea Legati, PhD
Jean-François Deleuze, MD, PhD
Anne Boland, PharmD, PhD
Olivier Quenez, MSc
Pierre Krystkowiak, MD, PhD
Pascal Favrole, MD
Daniel H. Geschwind, MD, PhD
Adi Aran, MD
Reeval Segel, MD
Ephrat Levy-Lahad, MD
Dennis W. Dickson, MD
Giovanni Coppola, MD
Rosa Rademakers, PhD
João R.M. de Oliveira, MD, PhD

Correspondence to Dr. de Oliveira: Joao.ricardo@ufpe.br

Supplemental data at Neurology.org/ng

ABSTRACT

Objective: To assess the potential connection between PCDH12 and brain calcifications in a patient carrying a homozygous nonsense variant in *PCDH12* and in adult patients with brain calcifications.

Methods: We performed a CT scan in 1 child with a homozygous *PCDH12* nonsense variant. We screened DNA samples from 53 patients with primary familial brain calcification (PFBC) and 26 patients with brain calcification of unknown cause (BCUC).

Results: We identified brain calcifications in subcortical and perithalamic regions in the patient with a homozygous *PCDH12* nonsense variant. The calcification pattern was different from what has been observed in PFBC and more similar to what is described in in utero infections. In patients with PFBC or BCUC, we found no protein-truncating variant and 3 rare (minor allele frequency <0.001) *PCDH12* predicted damaging missense heterozygous variants in 3 unrelated patients, albeit with no segregation data available.

Conclusions: Brain calcifications should be added to the phenotypic spectrum associated with *PCDH12* biallelic loss of function, in the context of severe cerebral developmental abnormalities. A putative role for *PCDH12* variants remains to be determined in PFBC. ***Neurol Genet* 2017;3: e166; doi: 10.1212/NXG.0000000000000166**

GLOSSARY

BCUC = brain calcification of unknown cause; **ExAC** = Exome Aggregation Consortium; **PFBC** = primary familial brain calcification.

A homozygous nonsense *PCDH12* variant has recently been reported in consanguineous families, where the affected children had congenital microcephaly, epilepsy, and profound global developmental disability.[1] Fetal MRI and USG showed dysplastic elongated masses in the midbrain-hypothalamus-optic tract area and hyperechogenic perithalamic foci. *PCDH12* encodes a protocadherin associated with membrane physical stability, adhesion, and vasculature maintenance and has recently been pointed out as a candidate gene for primary familial brain calcification (PFBC). PFBC is characterized by the presence of calcifications affecting primarily the basal ganglia, in the absence of secondary cause.[2] Clinical manifestations include movement disorders, cognitive impairment, psychiatric disturbances, and headache, most frequently beginning during adulthood.[2,3] Heterozygous variants causing autosomal dominant PFBC in up to 50% of the families were identified in 4 genes: *SLC20A2*, *PDGFRB*, *PDGFB*, and *XPR1*.[4–8] We previously searched for genes with a cerebral expression pattern similar to the PFBC major

From the Department of Genetics and CNR-MAJ (G.N., A.-C.R., O.Q.), Normandie Univ, UNIROUEN, Inserm U1245, Rouen University Hospital, Normandy Center for Genomic and Personalized Medicine, France; Department of Human Genetics (G.N.), Genome Research, Radboud UMC, Nijmegen, The Netherlands; Department of Neuroscience (M.S.-C., D.W.D., R.R.), Mayo Clinic, Jacksonville, FL; Department of Psychiatry (E.M.R., A.R.L., A.L., G.C., D.H.G.), Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California Los Angeles; Keizo Asami Laboratory (R.R.L., J.F., D.M., J.R.M.d.O), Federal University of Pernambuco, Recife, Brazil; Fundación Pública Galega de Medicina Xenómica (M.J.S.), Clinical University Hospital of Santiago de Compostela-SERGAS, Spain; Centre National de Recherche en Génomique Humaine (CNRGH) (J.-F.D., A.B.), Institut de Biologie François Jacob, CEA, Evry; Department of Neurology (P.K.), Amiens University Hospital; Department of Neurology (P.F.), Tenon Hospital, AP-HP, Paris, France; Medical Genetics MRI Unit (R.S., E.L.-L.), Shaare Zedek Medical Center; Hebrew University-Hadassah School of Medicine (R.S., E.L.-L.); and Neuropsychiatry Department (J.R.M.d.O), Universidade Federal de Pernambuco, Recife, Brazil.

Funding information and disclosures are provided at the end of the article. Go to Neurology.org/ng for full disclosure forms. The Article Processing Charge was funded by the authors.

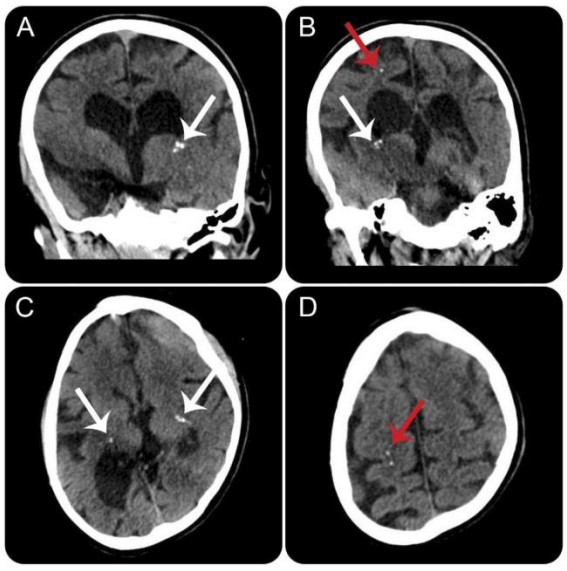
causative gene *SLC20A2* using the Allen Brain Atlas (brain-map.org/),[9,10] observing a higher *SLC20A2* expression in regions affected by calcifications in PFBC. *PCDH12* was singled out with the highest significant correlation,[10] and a follow-up analysis with additional brains still shows *PCDH12* as the most similar pattern to *SLC20A2*, even when compared with the other known PFBC causative genes (table e-1 at Neurology.org/ng).

To evaluate the potential link between *PCDH12* and brain calcifications, (1) we performed a CT scan in a patient reported to carry a homozygous nonsense *PCDH12* variant and (2) we screened DNA samples from patients with PFBC or brain calcifications of unknown cause (BCUC).

METHODS

CT imaging in *PCDH12* homozygous variant carriers. In the original report, patients with symmetric intrauterine growth retardation, severe microcephaly, visual impairment, dystonia, epilepsy, and profound developmental disability were shown to carry a *PCDH12* c.995T>A, p.R839X homozygous variant.[1] This variant is considered to be pathogenic when carried at the homozygous state following the American College of Medical Genetics and Genomics and the Association for Molecular Pathology recommendations.[11] Brain imaging revealed midbrain hypothalamus dysplasia and significant periventricular and/or periventricular hyperechogenicity. Fetal USG and MRI did not enable to determine whether these foci are eventually calcifications. Therefore, we performed a brain CT scan in individual III-1, family B from the original pedigree.[1]

Figure Brain CT imaging of a patient carrying the PCDH12 c.995T>A, p.R839X homozygous variant



(A, B) Coronal sections. (C, D) Transversal sections. Spot calcifications affecting perithalamic regions (white arrows, A–C) and subcortical regions (red arrows, B, D).

***PCDH12* screening in patients with brain calcification.** We included a total of 79 worldwide adult cases with brain calcifications that were referred to 5 centers of expertise, negatively screened for the known PFBC causative genes (supplemental data). Of these, 53 cases matched the clinical inclusion criteria for PFBC (detailed previously in reference 3). Briefly, these cases exhibited at least bilateral basal ganglia calcifications and no secondary cause. The remaining 26 patients were included on a neuropathologic basis if they presented moderate-to-severe basal ganglia calcifications. Note that calcifications also involved other brain regions in almost all cases and that other causes of brain calcifications could not be excluded in these patients, thereafter referred as having BCUC. All patients were screened for pathogenic variants by sequencing all coding exons of *PCDH12* (reference transcript: NM_016085.3). Bioinformatics predictions were performed using direct access to Polyphen2 HumDiv,[12] SIFT,[13] and Mutation Taster[14] tools, and the minor allele frequency (MAF) was checked at the Exome Aggregation Consortium (ExAC) website accessed in August 2016 (exac.broadinstitute.org/).[15] Detailed inclusion criteria and sequencing methods are provided in supplemental data.

Standard protocol approvals, registrations, and patient consents. All patients provided written informed consent for genetic analyses.

RESULTS

CT of a *PCDH12* homozygous variant carrier. CT is the reference imaging to identify brain calcification, so we used it to determine the nature of the hyperechogenic foci identified in a patient with a homozygous nonsense p.R839X *PCDH12* variant.[1] We identified spots of perithalamic calcification located in the posterior arms of the internal capsules and in juxtacortical right white matter (figure).

***PCDH12* screening in patients with brain calcification.** As we provided evidence that *PCDH12* biallelic loss of function is associated with brain calcification and given the high level of coexpression with the PFBC major causative gene *SLC20A2*, we next screened this gene in a group of patients with PFBC or BCUC. Among the 79 patients with PFBC or BCUC, we did not identify any protein-truncating variant (nonsense, splice site, or frameshift insertion/deletion). However, we detected 4 rare (MAF <0.001 in ExAC) heterozygous *PCDH12* missense variants in 4 unrelated patients: c.163C>G, p.(R55G); c.440G>T, p.(S147I); c.995T>A, p.(I332N); and c.3271G>A, p.(G1091S) (table 1). Three were predicted damaging by at least 1 in silico tool, while variant p.R55G was predicted benign by all 3 tools.

The c.440G>T, p.(S147I) variant had an MAF of 2.5e-05 in the ExAC database and was exclusively found in 3 individuals with the same ancestry as the patient (classified in ExAC as European non-Finnish). Two of the 3 in silico tools (Mutation Taster and Polyphen2 HumDiv, but not SIFT) predicted a damaging effect for this change to the protein

Table 1 Rare *PCDH12* variants identified in a series of 79 patients with PFBC or BCUC

Location (Ghrc37)	cDNA change[a]	Protein change[a]	ExAC frequency[b]	SIFT prediction	Polyphen2 HumDiv prediction	Mutation Taster prediction	PhyloP
chr5:141337254	c.163C>G	p.(R55G)	6.1e-04	Tolerated	Benign	Polymorphism	−1.01
chr5:141336977	c.440G>T	p.(S147I)	2.5e-05	Tolerated	Possibly damaging[c]	Disease causing[c]	2.14[c]
chr5:141336422	c.995T>A	p.(I332N)	1e-04	Deleterious[c]	Probably damaging[c]	Disease causing[c]	4.48[c]
chr5:141325230	c.3271G>A	p.(G1091S)	3.3e-05	Deleterious[c]	Probably damaging[c]	Disease causing[c]	4.81[c]

Abbreviations: BCUC = brain calcification of unknown cause; cDNA = complementary DNA; ExAC = Exome Aggregation Consortium; PFBC = primary familial brain calcification.
[a] Accession number: NM_016085.3.
[b] ExAC minor allele frequency assessed in August 2016.[15]
[c] Values are above each threshold.

function. DNA from relatives was not available for segregation analysis. This variant is located in the second cadherin tandem repeat domain (EC2) (NCBI accession cd11304) and, therefore, could affect homophilic adhesive behavior and calcium-dependent cell adhesion.[16]

The c.995T>A, p.(I332N) and c.3271G>A, p.(G1091S) variants are both predicted damaging by all 3 in silico tools. The p.I332N variant was reported with an overall MAF of 0.0001 in ExAC, found in 12 individuals of East Asian ancestry (the patient was born in Southeastern Asia) and 1 individual of European non-Finnish ancestry. The p.G1091S variant has an overall MAF of 3.3e-05, found in 1 individual of European non-Finnish ancestry (same as the patient) and 3 individuals of South Asian ancestry. DNA from relatives was not available for segregation analysis of any variant. Variant p.I332N is also located in a cadherin tandem repeat domain, namely EC3. However, p.G1091S variant is located in a highly conserved site in the cytoplasmic domain, which has a unique sequence among the cadherin family. Unlike the other cadherins, the cytoplasmic domain of PCDH12 does not interact with catenins, and it is involved in cellular processes other than cell junction, such as regulation of gene expression and signaling pathways.[17] Clinical details of all 3 predicted damaging variant carriers are provided in the supplemental data.

DISCUSSION We show here that a homozygous nonsense *PCDH12* variant, detected in patients with severe developmental delay and microcephaly,[1] is associated with brain calcifications. This feature should therefore be added to the phenotypic spectrum of this rare disorder. The pattern of calcifications is, however, different from the typical findings in PFBC, where calcifications always affect at least both pallidum,[3] and resembled to those observed in various neuroinfectious prenatal conditions, such as TORCH infections.[18] Brain calcification is a highly informative feature on brain imaging of children with neurodevelopmental disorders.[18] Although CT is the reference imaging tool for detecting and assessing calcifications, MRI is the primary imaging tool for the detection of all other brain abnormalities in the absence of radiation. T2* or susceptibility-weighted images increase the diagnostic performance of MRI for calcification compared with the other sequences. However, they can sometimes miss small calcifications, and they are still complementary with CT to describe precise shape and intensity and to definitely conclude on the differential identification with iron deposits.[19,20] In our patient, neither T2* nor susceptibility-weighted images were available.

In the original report, the efficiency of nonsense-mediated decay has been measured as 84%, suggesting a strong loss of function. The patients carrying the nonsense *PCDH12* variant in a homozygous state may still express little amount of the truncated protein, but no full-length PCDH12. This supports the hypothesis that loss of function of PCDH12 is the mechanism leading to the patient's phenotype, including brain calcification.

In a candidate gene approach, we searched for rare *PCDH12* variants in PFBC and BCUC patients and found no protein-truncating variants. Three heterozygous missense variants, predicted damaging by at least one of the tools, were identified in 2 patients with PFBC and 1 patient with BCUC. Given the fact that biallelic loss of *PCDH12* function leads to a severe neurodevelopmental phenotype, it is unlikely that these variants have a dominant-negative effect. However, as they are missense variants, their putative effect on protein function is hard to predict, and it remains possible that they are responsible for loss of function, gain of function, or have a neutral effect on protein function. The frequencies of theses variants in the patients' respective populations as estimated in ExAC are not inconsistent with a causative effect, as they are in the same frequency ranges as other disease-causing variants in *SLC20A2*.[8] Because neither segregation nor functional data are available, it is not possible to conclude about their pathogenicity at this stage.

Besides PFBC, brain calcifications can be detected in other numerous distinct conditions, such as

systemic phosphocalcic metabolism disorders of inherited or acquired cause, in utero or postnatal infections, interferonopathies, inborn errors of metabolism, and other rare inherited diseases.[21] Calcifications are believed to be related to increased type-I interferon response in both in utero viral infections and interferonopathies.[22] Several of these clinical presentations, including TORCH in utero infections and typical Aicardi-Goutières syndrome, are similar to the ones observed in the *PCDH12* homozygous carriers. In other conditions, mutations in *OCLN* and *JAM3* genes, encoding endothelial cell adhesion proteins, result in microangiopathy associated with calcifications.[18,23,24] Given the known function of PCDH12, we postulate that similar mechanisms could be associated with the calcifications observed in the *PCDH12* homozygous loss-of-function carriers.

PCDH12 is a protocadherin associated with membrane physical stability and adhesion.[25] A *Pcdh12* knockout mouse model revealed several age-independent vessel impairments, such as ramifications of medial elastic lamellae and increased inner diameter and circumferential mid-wall stress.[26] PCDH12 has been widely studied as a key-player cadherin involved in placental maintenance and also a preeclampsia biomarker; however, little is known about its involvement in brain physiology. It is conceivable that mutations in PCDH12 and SLC20A2, which share similar expression patterns in the brain, might lead to similar phenotypes. Of interest, *Slc20a2* knockout mice developed not only brain calcifications but also fetal growth restriction, lower birth viability, and placental calcification associated with thickened basement membranes.[27] In both mouse models, the placental phenotype and the vascular impairment are additional putative links between SLC20A2 and PCDH12, which deserve additional studies on mouse models.

PCDH12 biallelic loss of function causes a severe neurodevelopmental phenotype associated with brain calcifications. Rare predicted damaging heterozygous *PCDH12* variants were identified in patients with PFBC or BCUC here, but whether they are associated with brain calcification or not remains to be determined. To address this question, follow-up studies will be necessary including screening other series, assessing the segregation of rare variants and functional consequences.

AUTHOR CONTRIBUTIONS

Collection and interpretation of data: Gaël Nicolas, Monica Sanchez-Contreras, Eliana Marisa Ramos, Roberta R. Lemos, Joana Ferreira, Denis Moura, Maria J. Sobrido, Anne-Claire Richard, Alma Rosa Lopez, Andrea Legati, Jean-François Deleuze, Anne Boland, Olivier Quenez, Pierre Krystkowiak, Pascal Favrole, Daniel H. Geschwind, Adi Aran, Reeval Segel, Ephrat Levy-Lahad, Dennis W. Dickson, Giovanni Coppola, Rosa Rademakers, and João R.M. de Oliveira. Manuscript draft: Gaël Nicolas, Monica Sanchez-Contreras, Eliana Marisa Ramos, Roberta R. Lemos, and João R.M. de Oliveira. Critical revision: Gaël Nicolas, Monica Sanchez-Contreras, Eliana Marisa Ramos, Roberta R. Lemos, João R.M. de Oliveira, and Giovanni Coppola. Study design and supervision: João R.M. de Oliveira, Gaël Nicolas, Giovanni Coppola, and Rosa Rademakers.

STUDY FUNDING

The French study (G.N.) was funded by Conseil Régional de Haute Normandie—APERC 2014 no. 2014-19 in the context of Appel d'Offres Jeunes Chercheurs (CHU de Rouen). Mayo Clinic group (M.S.-C., D.W.D., and R.R.) was supported by the Morris K. Udall Parkinson's Disease Research Center of Excellence (NINDS P50NS072187). J.R.M.d.O. acknowledges funding from CPNq (480255/2013-0, 440770/2016-5, 470781/2014-9, and 307909/2012-3). Funding sources had no specific roles.

DISCLOSURE

G. Nicolas, M. Sanchez-Contreras, and E.M. Ramos report no disclosures. R.R. Lemos has received research support from FACEPE—Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (Brazil). J. Ferreira has received research support from FACEPE—Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (Brazil). D. Moura has received research support from FACEPE—Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (Brazil). M.J. Sobrido has received speaker honoraria from Actelion Pharmaceuticals; is the CEO of Genomic Consulting; has a private clinical neurology practice and genetic diagnosis of patients with inherited diseases; and has received research support from Actelion Pharmaceuticals, Instituto de Salud Carlos III (Spain), and Asociación Galega de Ataxias (AGA). A.-C. Richard, A.R. Lopez, and A. Legati report no disclosures. J.-F. Deleuze has served on the editorial board of *Human Genetics*; is listed as an inventor on a patent for Quality assessment of induced pluripotent cells by referring to splicing signatures of pluripotency; and has received research support from Labex GenMed from French ANR. A. Boland and O. Quenez report no disclosures. P. Krystkowiak has served on the scientific advisory boards of Lundbeck, Allergan, IPSEN, Novartis Pharma, and Servier Euthérapie; and has received speaker honoraria from Lundbeck, Allergan, IPSEN, Novartis Pharma, and Servier Euthérapie. P. Favrole reports no disclosures. D.H. Geschwind has served on the scientific advisory board of Ovid Therapeutics Inc.; serves on the editorial boards of *Cell*, *Molecular Autism*, *Molecular Neuropsychiatry*, *Human Molecular Genetics*, *Neuron*, *Current Opinion in Genetics and Development*, and *Translational Psychiatry*; has acted as a reviewer for *Neurology®*; is listed as an inventor on the following patents: Peripheral Gene Expression Biomarkers for Autism (3), Genetic Risk Factor for Neurodegenerative Disease, Compositions and Methods for Diagnosing and Treating Brain Cancer and Identifying Neural Stem Cells, Genetic Variants Underlying Human Cognition: Novel Diagnostic and Therapeutic Targets, Brain Gene Expression Changes Associated with Autism Spectrum Disorders, Full Biomarkers in Friedreich's Ataxia (provisional patent application), Signaling Networks Causing Neurodevelopmental Disorders In Human Neurons, Genes Dysregulated in Autism—Potential Biomarkers and Therapeutic Pathways, mTor—A Genetic Target for Treatment of Individuals with Neurocognitive Spectrum Disorders, Neuronal Regeneration, Frataxin Knock-Down Mouse, Jakmip1 Knockout Mouse, and Cyfip1 Transgenic Mouse; receives publishing royalties from Oxford University Press; has been a consultant for OVID Therapeutics Ltd.; has received research support from Takeda Pharmaceutical Company, NIH/NIMH, NIH/NICHD, NIH/NINDS, NIH/NIA, the Simons Foundation, Adelson Medical Research Foundation, and the Tau Consortium; holds stock/stock options in Ovid Therapeutics; and receives license fee payments for a Mouse model of Friedreich Ataxia (licensed by UCLA). A. Aran reports no disclosures. R. Segel has received research support from the Joint Research Fund (Keren Meshutefet) of the Hebrew University and the Hadassah and Shaare Zedek Hospitals. E. Levy-Lahad has received research support from USAID MERC program, Israel Science Foundation (Israel), Breast Cancer Research Foundation (BCRF; New York, NY), and Israel Cancer Research Fund (ICRF; New York, NY). D.W. Dickson has received travel funding and speaker honoraria from Novartis; has served on

the editorial boards of *Acta Neuropathologica*, *Brain*, *Brain Pathology*, *Neuropathology and Applied Neurobiology*, *Annals of Neurology*, *Neuropathology*, *International Journal of Clinical and Experimental Pathology*, and *American Journal of Neurodegenerative Disease*; has received research support from the following grants: P50AG016574 (Core Leader), P50NS072187 (Center Director), P01NS084974 (Project Leader), and P01AG003949 (Core Leader); and has received research support from the Mangurian Foundation. G. Coppola receives publishing royalties from Oxford University Press and receives research support from Takeda Pharmaceutical Company, NIH, the Consortium for Frontotemporal Dementia Research, Adelson Medical Research Foundation, the Tau Consortium, CHDI, and John Douglas French Alzheimer's Foundation. R. Rademakers holds patents for Detecting and Treating Dementia, and for Methods and Materials for Detecting and Treating Dementia and has received research support from NIH and the Consortium for Frontotemporal Dementia. J.R.M. de Oliveira has received research support from the Federal Council for Research and Technology (Brazil) and FACEPE—Fundação de Amparo a Ciência e Tecnologia do Estado de Pernambuco (Brazil). Go to Neurology.org/ng for full disclosure forms.

Received April 3, 2017. Accepted in final form May 31, 2017.

REFERENCES

1. Aran A, Rosenfeld N, Jaron R, et al. Loss of function of PCDH12 underlies recessive microcephaly mimicking intrauterine infection. Neurology 2016;86:2016–2024.
2. Sobrido MJ, Coppola G, Oliveira J, Hopfer S, Geschwind DH. Primary familial brain calcification. In: Pagon RA, Adam MP, Ardinger HH, et al, editors. GeneReviews. Seattle: University of Washington; 1993.
3. 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. Am J Med Genet B Neuropsychiatr Genet 2015;168:586–594.
4. Legati A, Giovannini D, Nicolas G, et al. Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. Nat Genet 2015; 47:579–581.
5. 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.
6. Keller A, Westenberger A, Sobrido MJ, et al. Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. Nat Genet 2013;45:1077–1082.
7. Wang C, Li Y, Shi L, et al. Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis. Nat Genet 2012;44:254–256.
8. Lemos RR, Ramos EM, Legati A, et al. Update and mutational analysis of SLC20A2: a major cause of primary familial brain calcification. Hum Mutat 2015;36:489–495.
9. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 2012;489:391–399.
10. da Silva RJ, Pereira IC, Oliveira JR. Analysis of gene expression pattern and neuroanatomical correlates for SLC20A2 (PiT-2) shows a molecular network with potential impact in idiopathic basal ganglia calcification ("Fahr's disease"). J Mol Neurosci 2013;50:280–283.
11. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405–424.
12. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013;Chapter 7:Unit 7.20.
13. Ng PC, Henikoff S. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 2003;31: 3812–3814.
14. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 2014;11:361–362.
15. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–291.
16. Bouillot S, Tillet E, Carmona G, et al. Protocadherin-12 cleavage is a regulated process mediated by ADAM10 protein: evidence of shedding up-regulation in pre-eclampsia. J Biol Chem 2011;286:15195–15204.
17. Wu Q, Maniatis T. Large exons encoding multiple ectodomains are a characteristic feature of protocadherin genes. Proc Natl Acad Sci USA 2000;97:3124–3129.
18. Livingston JH, Stivaros S, van der Knaap MS, Crow YJ. Recognizable phenotypes associated with intracranial calcification. Dev Med Child Neurol 2013;55:46–57.
19. Kozic D, Todorovic-Djilas L, Semnic R, Miucin-Vukadinovic I, Lucic M. MR imaging: an unreliable and potentially misleading diagnostic modality in patients with intracerebral calcium depositions. Case report. Neuro Endocrinol Lett 2009;30:553–557.
20. Liu S, Buch S, Chen Y, et al. Susceptibility-weighted imaging: current status and future directions. NMR Biomed 2017:30.
21. Deng H, Zheng W, Jankovic J. Genetics and molecular biology of brain calcification. Ageing Res Rev 2015;22: 20–38.
22. Crow YJ, Manel N. Aicardi-Goutieres syndrome and the type I interferonopathies. Nat Rev Immunol 2015;15:429–440.
23. Rajab A, Aldinger KA, El-Shirbini HA, Dobyns WB, Ross ME. Recessive developmental delay, small stature, microcephaly and brain calcifications with locus on chromosome 2. Am J Med Genet A 2009;149A:129–137.
24. Mochida GH, Ganesh VS, Felie JM, et al. A homozygous mutation in the tight-junction protein JAM3 causes hemorrhagic destruction of the brain, subependymal calcification, and congenital cataracts. Am J Hum Genet 2010;87: 882–889.
25. Rampon C, Bouillot S, Climescu-Haulica A, et al. Protocadherin 12 deficiency alters morphogenesis and transcriptional profile of the placenta. Physiol Genomics 2008;34: 193–204.
26. Philibert C, Bouillot S, Huber P, Faury G. Protocadherin-12 deficiency leads to modifications in the structure and function of arteries in mice. Pathol Biol (Paris) 2012;60: 34–40.
27. Wallingford MC, Gammill HS, Giachelli CM. Slc20a2 deficiency results in fetal growth restriction and placental calcification associated with thickened basement membranes and novel CD13 and lamininalpha1 expressing cells. Reprod Biol 2016;16:13–26.

APÊNDICE C - MANUSCRITO PUBLICADO EM PERIÓDICO: 'New data from Pdgfb ret/ret mice might lead to a paradoxical association between brain calcification, perycites recruitment and BBB integrity'

PERIÓDICO *JOURNAL OF MOLECULAR NEUROSCIENCE*

Fator de impacto (JCR 2011): 2,229
Qualis CAPES Ciências Biológicas I: B1

Autores: Denis A. P. Moura[1]; Roberta R. Lemos[1]; João R. M. de Oliveira[1,2*]
[1]Laboratório de Imunopatologia Keizo Asami, Universidade Federal de Pernambuco.
[2]Departamento de Neuropsiquiatria, Universidade Federal de Pernambuco.
*Corresponding author. Email: joao.ricardo@ufpe.br

Título: New Data from Pdfgb ret/ret Mutant Mice Might Lead to a Paradoxical Association Between Brain Calcification, Pericytes Recruitment and BBB Integrity

J Mol Neurosci
https://doi.org/10.1007/s12031-017-0992-z

New Data from *Pdfgb*$^{ret/ret}$ Mutant Mice Might Lead to a Paradoxical Association Between Brain Calcification, Pericytes Recruitment and BBB Integrity

Denis A. P. Moura[1] · **Roberta R. Lemos**[1] · **João R. M. Oliveira**[1,2]

Received: 27 September 2017 / Accepted: 13 October 2017

Abstract Data of mice with PDGF-B-truncating mutation (*Pdgfb*$^{ret/ret}$) from different research groups indicate that the malfunction of this protein leads to reduced pericyte recruitment, loss of Blood-Brain Barrier (BBB) integrity and bilateral brain calcification. This makes these mice important models for Primary Brain Calcification and pericyte-BBB correlation studies. The global brain pericyte count is reduced in *Pdgfb*$^{ret/ret}$ mice, with higher BBB permeability. We have overlapped the data from other research groups into a figure to further analyze the findings. Calcifications form within midbrain, interbrain, basal forebrain, and pons. Interestingly, these calcification-prone regions have a comparably higher pericyte count and lower BBB leakage in relation to other non-calcifying regions of the *Pdgfb*$^{ret/ret}$ mouse (such as the cortex and striatum). A comparatively higher BBB integrity in regions prone to calcification seems paradoxical and indicates that other region-specific changes are the cause of the calcifications.

Keywords Pdgfb · Animal model · brain calcification · Primary familial brain calcification · Blood-brain barrier · Pericyte

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12031-017-0992-z) contains supplementary material, which is available to authorized users.

✉ João R. M. Oliveira
joao.ricardo@ufpe.br

[1] Laboratório de Imunopatologia Keizo Asami, Universidade Federal de Pernambuco, Recife, Brazil

[2] Departamento de Neuropsiquiatria, Universidade Federal de Pernambuco, Recife, Brazil

In their recent paper, Villaseñor and colleagues discussed the regional heterogeneity of the Blood-Brain Barrier (BBB) in *PDGFB* mutant mice (Villaseñor et al. 2017). This mice model (*Pdgfb*$^{ret/ret}$) is the same used by Keller and collaborators in which they found that mutations in this gene are a cause for Primary Familial Brain Calcification (Keller et al. 2013). This lineage has a homozygous mutation in the *PDGFB* gene, causing its product to be a hypomorphic, absent of C terminus residues, biologically active but less biodisponible and thus compatible with adult life. These mice have been used to model Primary Familial Brain Calcification, for the mutant develops basal ganglia calcifications, and as a model for pericyte and BBB correlation studies. Villaseñor and colleagues used Evans blue staining to investigate BBB permeability. It was found that the cortex, striatum, and hippocampus had a higher BBB permeability while some other regions, such as the midbrain and interbrain, maintained a lower BBB permeability despite the pericyte loss (Villaseñor et al. 2017). Vanlandewijck and colleagues, also using the same mice model, surprisingly found that brain regions more prone to calcifications had higher pericyte coverage and lower BBB leakage than other brain regions (Vanlandewijck et al. 2015).

We overlapped the findings from Villaseñor, Keller and Vanlandewijck and their respective collaborators (Villaseñor et al. 2017; Keller et al. 2013; Vanlandewijck et al. 2015) where regions more prone to calcifications (midbrain, interbrain, pons and the basal forebrain) correlated to regions with lower BBB leakage (midbrain and interbrain) in *Pdgfb*$^{ret/ret}$ mice. We digitally built an approximation model extrapolating from their data, in which we inputted calcification and BBB extravasation data in mouse coronal, sagittal and horizontal brain axes (Fig. 1). A mouse brain framework was retrieved from the Allen Institute Mouse Brain Atlas (Lein et al. 2007, http://mouse.brain-map.org/) and we manually curated, edited and colored it to illustrate BBB integrity and calcification points based on

Published online: 02 November 2017

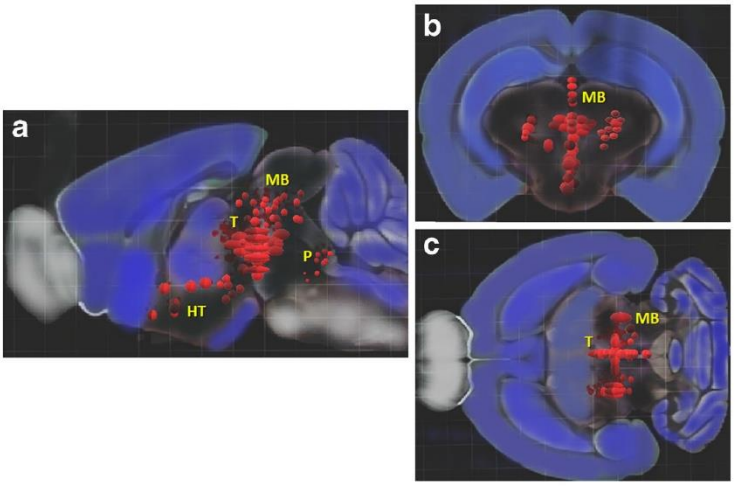


Fig. 1 An illustrative approximation of the *Pdgfb*$^{ret/ret}$ mouse brain. Sagittal (**A**), coronal (**B**) and horizontal (**C**) Allen Mouse Brain Atlas sections (Lein et al. 2007, http://mouse.brain-map.org/) digitally edited to show regions with high BBB leakage (*blue*) and with low BBB leakage (*black*) based on the paper from Villaseñor et al. (2017). Calcification points (*red*) were marked by overlapping figures from Keller et al. (2013, Fig. 3a.) and Vanlandewijck et al. (2015, Fig. 7c.), which concentrated in the midbrain (MB), hypothalamus (HT), thalamus (T) and pons (P). Gray areas represent areas with absent BBB extravasation data

each primary source. We have also created a 3D animated figure illustrating the calcification profundity by extrapolating the 2D data (Supplementary Fig. 1). This data further confirms the findings of Vanlandewijck and colleagues and suggests that other region-specific factors must be involved (Vanlandewijck et al. 2015).

Primary Familial Brain Calcifications (PFBC) is an autosomal dominant neuropathology characterized by calcifications in the basal ganglia, thalamus, cerebellum and brain stem. Affected individuals may present motor, cognitive or psychiatric impairments, while sometimes it remains asymptomatic (Nicolas et al. 2015). Currently, there are four genes confirmed as causative for the disease: *SLC20A2*, which codes the inorganic phosphate transporter PiT2 (Wang et al. 2012); *XPR1*, which codes a membrane receptor for phosphate transport (Legati et al. 2015); *PDGFB*, which codes the subunit beta of the platelet-derived growth factor (Keller et al. 2013); and *PDGFRB*, a membrane receptor for the platelet-derived growth factors (Nicolas et al. 2013).

The identification of mutations in the *PDGFB* and *PDGFRB* genes instigated the hypothesis of the blood brain barrier's (BBB) involvement in the process of calcification. PDGF-B, a product of *PDGFB*, functions in angiogenesis by the recruitment of pericytes (Enge et al. 2002; Bjarnegard et al. 2004). Indeed, mice with partially inactivating mutations in this gene, such as the *Pdgfb*$^{ret/ret}$ and the R26P$^{+/0}$ mice models, present a BBB defect which correlates to pericyte loss and the calcifications (Armulik et al. 2010; Keller et al. 2013). The first hypothesis of this correlation was the formation of a mineral deposition nidus caused by plasma proteins crossing the disrupted BBB (Miklossy et al. 2005). On the other side, mutations in *SLC20A2*, a gene with involvement in most PFBC cases, does not disrupt the BBB permeability. Wallingford (Wallingford et al. 2017) showed that there was no difference in the pericyte count in the brains of Basal Ganglia Calcification-positive *SLC20A2*$^{+/-}$ mice and the controls. The researchers also showed that the BBB permeability remained unchanged in the same mice at the age of 12 months.

What we can hypothesize about those findings is that there may be different ways that the calcifications might form in the case of basal ganglia calcifications and PFBC. One of the ways is the accumulation of inorganic phosphate and deposition of calcium phosphate through the disruption of the phosphate transport channels (PiT2 and XPR1). The other way would be through the PDGF (through PDGFR-B and PDGF-B) malfunction, where globally there is reduction of pericyte coverage and BBB impairment although calcification-prone regions will have a higher pericyte recruitment but less BBB leakage in relation to non-calcification-prone regions.

It is interesting to notice that PFBC patients with *PDGFB*, *SLC20A2* and *XPR1* mutations sometimes present striatum and cortical calcifications (Hayashi et al. 2015; Batla et al. 2017) whereas the mice models *Pdgfb*$^{ret/ret}$ (and R26P$^{+/0}$) have decreased pericyte coverage, lower BBB integrity and no calcifications in these regions (Armulik et al. 2010; Keller et al. 2013). This notes a discrepancy between the mice model and the human phenotypes, and different regional effects might be involved in each case.

Brain calcification formation without the involvement of BBB seems paradoxical. This may indicate a new line of

thought to imagining how the calcifications form within specific brain regions, such as in PFBC. Pericyte coverage and BBB permeability's relationship with the basal ganglia calcifications remain incompletely understood. It is paramount to determinate whether molecular pathways changes correlate to the loss of signalization of the PDGFs or loss of function of the *SLC20A2*/*XPR1* in the models of PFBC. Further pathomolecular studies, especially in the calcification sites, are needed to understand the regional effects that lead to the formation of the calcifications in both genetic scenarios.

Acknowledgements The authors would like to thank the following funding agencies for their financial support: Fundação de Amparo a Ciência e Tecnologia - FACEPE (Award IBPG-0897-2.02/15), Conselho Nacional CNPq (Award 310150/2016-7) and CAPES (Award 8887.131374/2016-00).

Compliance with Ethical Standards

Declaration of Conflicting Interests The authors declare no potential conflict of interest with respect to the research, authorship and publication of this article.

References

Armulik A, Genove G, Mae M, Nisancioglu MH, Wallgard E, Niaudet C et al (2010) Pericytes regulate the blood-brain barrier. Nature 468: 557–561

Batla A, Tai XY, Schottlaender L, Erro R, Balint B, Bhatia KP (2017) Deconstructing Fahr's disease/syndrome of brain calcification in the era of new genes. Parkinsonism Relat Disord 37:1–10

Bjarnegard M, Enge M, Norlin J, Gustafsdottir S, Fredriksson S, Abramsson A et al (2004) Endothelium-specific ablation of PDGFB leads to pericyte loss and glomerular, cardiac and placental abnormalities. Development 131:1847–1857

Enge M, Bjarnegard M, Gerhardt H, Gustafsson E, Kalen M, Asker N et al (2002) Endothelium-specific platelet-derived growth factor-B ablation mimics diabetic retinopathy. EMBO J 21:4307–4316

Hayashi T, Legati A, Nishikawa T, Coppola G (2015) First Japanese family with primary familial brain calcification due to a mutation in the PDGFB gene: an exome analysis study. Psychiatry Clin Neurosci 69:77–83

Keller A, Westenberger A, Sobrido MJ, Garcia-Murias M, Domingo A, Sears RL et al (2013) Mutations in the gene encoding PDGF-B cause brain calcifications in humans and mice. Nat Genet 45: 1077–1082

Legati A, Giovannini D, Nicolas G, Lopez-Sanchez U, Quintans B, Oliveira JRM et al (2015) Mutations in XPR1 cause primary familial brain calcification associated with altered phosphate export. Nat Genet 47:579–581

Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A et al (2007) Genome-wide atlas of gene expression in the adult mouse brain. Nature 445:168–176

Miklossy J, Mackenzie IR, Dorovini-Zis K, Calne DB, Wszolek ZK, Klegeris A et al (2005) Severe vascular disturbance in a case of familial brain calcinosis. Acta Neuropathol 109:643–653

Nicolas G, Pottier C, Maltete D, Coutant S, Rovelet-Lecrux A, Legallic S et al (2013) Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. Neurology 80:181–187

Nicolas G, Charbonnier C, de Lemos RR, Richard A-C, Guillin O, Wallon D et al (2015) Brain calcification process and phenotypes according to age and sex: Lessons from SLC20A2, PDGFB, and PDGFRB mutation carriers. Am J Med Genet B Neuropsychiatr Genet 168:586–594

Vanlandewijck M, Lebouvier T, Andaloussi Mae M, Nahar K, Hornemann S, Kenkel D et al (2015) Functional characterization of germline mutations in PDGFB and PDGFRB in primary familial brain calcification. PLoS One 10:e0143407

Villaseñor R, Kuennecke B, Ozmen L, Ammann M, Kugler C, Gruninger F et al. (2017) Region-specific permeability of the blood-brain barrier upon pericyte loss. J Cereb Blood Flow Metab 271678X17697340

Wallingford MC, Chia JJ, Leaf EM, Borgeia S, Chavkin NW, Sawangmake C et al (2017) SLC20A2 deficiency in mice leads to elevated phosphate levels in cerbrospinal fluid and glymphatic pathway-associated arteriolar calcification, and recapitulates human idiopathic basal ganglia calcification. Brain Pathol 27:64–76

Wang C, Li Y, Shi L, Ren J, Patti M, Wang T et al (2012) Mutations in SLC20A2 link familial idiopathic basal ganglia calcification with phosphate homeostasis. Nat Genet 44:254–256