



UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE CIÊNCIAS MÉDICAS

HELENA TADIELLO DE MORAES

CARACTERIZAÇÃO GENÉTICA DE UMA COORTE LATINO-AMERICANA DE
PACIENTES COM ENCEFALOPATIAS EPILÉPTICAS E DO DESENVOLVIMENTO.

CAMPINAS

2022

HELENA TADIELLO DE MORAES

CARACTERIZAÇÃO GENÉTICA DE UMA COORTE LATINO-AMERICANA DE
PACIENTES COM ENCEFALOPATIAS EPILÉPTICAS E DO DESENVOLVIMENTO.

Tese apresentada à Faculdade de Ciências Médicas da Universidade
Estadual de Campinas como parte dos requisitos exigidos para a
obtenção do título de Doutora em Ciências.

ORIENTADORA: PROFA. DRA. ÍSCIA TERESINHA LOPES CENDES

ESTE TRABALHO CORRESPONDE À VERSÃO
FINAL DA TESE DEFENDIDA PELA ALUNA HELENA
TADIELLO DE MORAES, E ORIENTADA PELA
PROFA. DRA. ÍSCIA TERESINHA LOPES CENDES.

CAMPINAS

2022

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Faculdade de Ciências Médicas
Ana Paula de Moraes e Oliveira - CRB 8/8985

M791c Moraes, Helena Tadiello de, 1987-
Caracterização genética de uma coorte latino-americana de pacientes com encefalopatias epiléticas e do desenvolvimento / Helena Tadiello de Moraes. – Campinas, SP : [s.n.], 2022.

Orientador: Iscia Teresinha Lopes-Cendes.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Ciências Médicas.

1. Encefalopatia epilética e do desenvolvimento. 2. Sequenciamento completo de exoma. 3. Microarranjos de DNA. 4. Diagnóstico molecular. 5. Herança complexa. I. Lopes-Cendes, Iscia Teresinha, 1964-. II. Universidade Estadual de Campinas. Faculdade de Ciências Médicas. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Genetic characterization of a Latin American cohort of patients with developmental and epileptic encephalopathies

Palavras-chave em inglês:

Developmental and epileptic encephalopathy

Whole exome sequencing

DNA microarray

Molecular diagnosis

Complex inheritance

Área de concentração: Fisiopatologia Médica

Titulação: Doutora em Ciências

Banca examinadora:

Iscia Teresinha Lopes-Cendes [Orientador]

Ana Carolina Coan

Antonia Paula Marques de Faria

Kette Dualibi Ramos Valente

Magda Lahorgue Nunes

Data de defesa: 07-06-2022

Programa de Pós-Graduação: Fisiopatologia Médica

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-1285-2687>

- Currículo Lattes do autor: <http://lattes.cnpq.br/9774925559884301>

COMISSÃO EXAMINADORA DA DEFESA DE DOUTORADO

HELENA TADIELLO DE MORAES

ORIENTADORA: PROFA. DRA. ÍSCIA TERESINHA LOPES CENDES

MEMBROS TITULARES:

- 1. PROFA. DRA. ÍSCIA TERESINHA LOPES CENDES**
- 2. PROFA. DRA. ANA CAROLINA COAN**
- 3. PROFA. DRA. ANTONIA PAULA MARQUES DE FARIA**
- 4. PROFA. DRA. KETTE DUALIBI RAMOS VALENTE**
- 5. PROFA. DRA. MAGDA LAHORGUE NUNES**

Programa de Pós-Graduação em Fisiopatologia Médica da Faculdade de Ciências Médicas da Universidade Estadual de Campinas.

A ata de defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Tese e na Secretaria do Programa da FCM.

Data de Defesa: 07/06/2022

DEDICATÓRIA

Dedico este trabalho a todas as crianças e familiares que sofrem com o diagnóstico de algum tipo de epilepsia, especialmente para aquelas que participaram de forma tão esperançosa dessa pesquisa. O contato que tive com essas crianças e famílias durante o tempo do doutorado me fez ver a vida de uma outra maneira. Espero ter contribuído, nem que seja um pouquinho, para o avanço do conhecimento nessa área e que isso possa trazer alguma melhoria na vida dessas pessoas.

AGRADECIMENTOS

Primeiramente gostaria de agradecer a minha família e amigos que sempre estiveram ao meu lado me apoiando em qualquer decisão. Principalmente meu marido Felipe pelo apoio e amor incondicionais. Não conseguiria nada sem você!

Agradecimentos especiais a Profa. Dra Íscia Lopes Cendes por ter aberto tantas portas para mim e por ter me dado todo o apoio necessário durante o desenvolvimento deste trabalho. Agradeço também a todos do Laboratório de Genética Molecular por nunca medirem esforços para me ajudar, especialmente Paty Aline e Lu Bonadia. Agradeço às meninas que sempre me ajudaram tanto no trabalho, quanto no dia a dia, nas conversas, nos almoços, nos cafés, nos bolos, nas risadas...Donatti, Canto, Jaque, Tânia, Dani Bruno, Diana, Estela e, em especial, Vanessa, amiga de muitos anos que tem um papel fundamental em toda a minha vida acadêmica e pessoal. Levarei sempre todos vocês em meu coração!

Agradeço imensamente a toda a equipe do ambulatório de epilepsias da infância do HC-UNICAMP. Principalmente a Profa. Dra. Ana Carolina Coan, Profa. Dra. Maria Augusta Montenegro e Malu.

Agradecimentos ao Prof. Dr. André Reis, meu orientador durante o estágio no exterior, e a todos do Instituto de Genética Humana da Universidade Friedrich-Alexander Erlangen-Nürnberg que me receberam tão bem e com muito profissionalismo.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001 (processo nº 88887.508863/2020-00) e da Fundação de Amparo à Pesquisa do Estado de São Paulo (processo nº 2017/00648-1).

RESUMO

As encefalopatias epilépticas e do desenvolvimento (EEDs) formam um grupo de epilepsias graves, geralmente resistentes ao tratamento medicamentoso e associadas a atraso do desenvolvimento neuropsíquico-motor e cognitivo. Com o avanço de estudos moleculares nos últimos anos, novas mutações têm sido descritas para as EEDs, porém, uma parcela significativa dos pacientes ainda permanece sem diagnóstico molecular, assumindo modelo monogênico de herança. Assim, o objetivo principal deste trabalho foi aplicar novos paradigmas analíticos para identificar e investigar alterações genéticas em uma grande coorte de pacientes com EEDs da América Latina. Dos 275 pacientes com diferentes tipos de EEDs que tivemos acesso, foi possível obter amostras de DNA em 239 deles. Primeiramente, utilizamos dados de sequenciamento completo do exoma (WES) de 234 desses pacientes e dados de microarranjos cromossômicos (CMA) de 236 desses pacientes para determinar a taxa de diagnóstico molecular dessas técnicas assumindo um modelo monogênico. Após análise individual de cada paciente, a taxa de diagnóstico foi de 38% para WES, de 4,6% para CMA e de 38,9% para as duas técnicas combinadas, o que representa uma alta taxa de diagnóstico genético se comparado com a literatura. Em segundo lugar, no grupo de pacientes em que não foi possível encontrar uma causa genética, utilizamos duas abordagens exploratórias tendo em mente modelos complexos de herança: regressão logística, que leva em conta a presença de variantes genéticas comuns, que apresentam frequência alélica, $MAF > 0,01$; e SKAT-O (do inglês, *optimized sequence kernel association test*) que analisa variantes raras ($MAF = 0,01$) e ultrarraras ($MAF = 0,0005$). Na análise de variantes comuns encontramos uma diferença entre as variantes dos grupos com e sem diagnóstico genético, sendo que o último apresentou uma variante exclusiva (rs9374755) em um gene de RNA (*LOC101927314*) com índice de associação ($p = 0,000005$). Na análise com variantes ultrarraras também observamos uma diferença entre os grupos com e sem diagnóstico, sendo que para o último encontramos sete genes candidatos que podem conter variantes de predisposição às EEDs: *ASPM* ($p = 0,045517718$); *CREBBP* ($p = 0,032086627$); *FASN* ($p = 0,009504475$); *LAMC3* ($p = 0,001417445$); *RELN* ($p = 0,017928247$); *RYR3* ($p = 0,047068168$); e *SPTAN1* ($p = 0,045613193$). O nosso estudo demonstrou que utilizando estratégias tradicionais de diagnóstico molecular foi possível encontrar uma causa monogênica para as EEDs em quase 40% dos

pacientes investigados, sendo que o WES teve o maior rendimento. Porém, o uso de CMA conseguiu adicionar quase cinco pontos percentuais no rendimento diagnóstico e, portanto, deve ser também considerado na investigação da etiologia genética das EEDs. Além disso, encontramos evidências de que pelo menos uma parte do efeito genético presente na etiologia das EEDs deve seguir um modelo complexo de herança. Apesar de não ser ainda possível aplicar essas análises ao diagnóstico genético de rotina, nossos resultados indicam que uma etiologia genética deve ser considerada em pacientes com EED mesmo que resultados negativos sejam obtidos utilizando as técnicas genômicas de rotina. A medida em que a possibilidade de diagnóstico molecular para doenças genéticas de herança complexa e multifatorial avança, acreditamos que as EEDs poderão se beneficiar desse novo conhecimento e novos métodos de análise dos dados genéticos.

Palavras-chave: Encefalopatia epilética e do desenvolvimento, sequenciamento completo de exoma, microarranjos de DNA, diagnóstico molecular, herança complexa.

ABSTRACT

Developmental and epileptic encephalopathies (DEEs) are a group of severe epilepsies, usually resistant to drug treatment and associated with delayed neuropsychomotor and cognitive development. With the advance of molecular genetic studies, mutations in many genes have been described in patients with DEEs; however, a significant proportion of these patients remains undiagnosed, assuming a monogenic model of inheritance. Thus, the main objective of this work was to apply new analytical paradigms to identify and investigate genetic variants in a large cohort of patients with DEEs from Latin America. Of the 275 patients with different types of DEEs that were initially ascertained, we obtained DNA samples from 239 of them. First, we used whole exome sequencing (WES) data from 234 of these patients and chromosomal microarray analysis (CMA) data from 236 of these patients to determine the molecular diagnostic yield of these techniques assuming a monogenic model. After individual analysis, the diagnostic yield was 38% for WES, 4.6% for CMA, and 38.9% for both techniques combined, representing a high diagnostic yield compared to the literature. Next, in the group of patients in whom a genetic cause could not be found, we used two exploratory approaches considering complex models of inheritance: logistic regression, which takes into account the presence of common genetic variants, which have allele frequency, $MAF > 0.01$; and SKAT-O (optimized sequence kernel association test) which analyzes rare ($MAF = 0.01$) and ultra-rare ($MAF = 0.0005$) variants. We found a difference between the groups with and without a genetic diagnosis in the analysis of common variants. The latter showed a unique variant (rs9374755) with evidence of association ($p = 0.000005$) located in an RNA gene (*LOC101927314*). In the analysis with ultra-rare variants, we also observed a difference between the groups with and without genetic diagnosis, and for the latter we found seven candidate genes containing variants predisposing to the DEEs: *ASPM* ($p = 0.045517718$); *CREBBP* ($p = 0.032086627$); *FASN* ($p = 0.009504475$); *LAMC3* ($p = 0.001417445$); *RELN* ($p = 0.017928247$); *RYR3* ($p = 0.047068168$); and *SPTAN1* ($p = 0.045613193$). Our study showed that using traditional molecular diagnostic strategies, it was possible to find a monogenic cause in almost 40% of the patients with DEE investigated, with WES having the highest yield. The use of CMA was able to add almost five percentage points in the diagnostic yield and, therefore, should also be considered when investigating the genetic etiology of DEEs. Furthermore, we found

evidence that at least part of the genetic effect present in the etiology of DEEs may follow a complex model of inheritance. Although it is not yet possible to apply these analyses to the routine genetic diagnosis of patients, our results indicate that a genetic etiology should be considered in patients with DEEs even in the absence of an identifiable monogenic cause. As the possibility of molecular diagnosis for genetic diseases of complex and multifactorial inheritance advances, we believe that the DEEs may benefit from this new knowledge and new methods of analyzing genetic data.

Key words: Developmental and epileptic encephalopathy, whole exome sequencing, DNA microarray, molecular diagnosis, complex inheritance.

LISTA DE ILUSTRAÇÕES

| | |
|--|----|
| Figura 1: Linha do tempo da descoberta de genes em epilepsias..... | 18 |
| Figura 2: Representação de algumas das EED bem conhecidas em relação às diferentes idades de início das crises..... | 21 |
| Figura 3: Classificação e distribuição dos pacientes com EED quanto ao fenótipo de acordo com a hipótese diagnóstica..... | 27 |
| Figura 4: Passos da técnica de CMA utilizando protocolo <i>Genome-Wide Human SNP Array 6.0 (Affymetrix™)</i> | 29 |
| Figura 5: Ancestralidade dos pacientes de EED da América Latina (BRA) calculada a partir da análise de componentes principais utilizando os dados de WES..... | 37 |
| Figura 6: (A) Distribuição do número de variantes em relação ao gene em que estão localizadas. (B) Tipo de mudança na sequência causada pela variante. (C) Classificação das variantes segundo a ACMG. (D) Padrão de herança da doença relacionada ao gene em que a variante se encontra..... | 30 |
| Figura 7: Distribuição dos pacientes com variantes causais no gene <i>SCN1A</i> e em outros genes mostrando que o gene <i>SCN1A</i> é o mais frequentes em todos os grupos..... | 41 |
| Figura 8: Cariótipo do paciente 67 com resultado 47,XY,+idic(15)(q11.2). A seta evidencia o cromossomo adicional idic(15)..... | 48 |
| Figura 9: (A) Cariótipo do paciente 144 com resultado 46,XY ish der(22)t(16;22)(p13.3;q13)(TUPLE1+,ARSA-,121I4+). (B) e (C) Análise de FISH em metáfase do paciente..... | 49 |
| Figura 10: (A) Cariótipo do pai do paciente 144 com resultado 46,XY ish t(16;22)(p13.3;q13)(ARSA+;TUPLE1+,ARSA-). (B) Análise de FISH em metáfase do pai do paciente 144. (C) Cariótipo normal da mãe do paciente 144. (D) Análise de FISH em metáfase normal da mãe do paciente 144..... | 50 |
| Figura 11: Gráficos QQ: O eixo Y mostra os valores de $-\log_{10} P$ observados, e o eixo X mostra os valores de $-\log_{10} P$ esperados. (A) Gráfico QQ do grupo de pacientes totais com λ de 1,108. (B) Gráfico QQ de pacientes com diagnóstico genético (DG) e λ de 1,109. (C) Gráfico QQ de pacientes sem diagnóstico genético e λ de 1,109..... | 59 |
| Figura 12: Gráficos de Manhattan obtidos a partir da análise de regressão logística. (A) Gráfico de Manhattan para o grupo de pacientes totais mostrando os três SNPs de interesse que ultrapassaram o primeiro valor de referência. (B) Gráfico de Manhattan para o grupo de pacientes com DG mostrando o único SNP de interesse que ultrapassou o primeiro valor de referência. (C) Gráfico de Manhattan para o grupo de pacientes sem DG mostrando o único SNP de interesse que ultrapassou o primeiro valor de referência..... | 61 |

LISTA DE TABELAS

| | |
|---|----|
| Tabela 1: Parâmetros de qualidade e número de variantes dos WES..... | 38 |
| Tabela 2: Variantes selecionadas após análise de WES..... | 42 |
| Tabela 3: CNVs causais selecionadas a partir da análise de WES..... | 45 |
| Tabela 4: CNVs patogênicas causais selecionadas a partir da análise de CMA..... | 47 |
| Tabela 5: Pacientes que apresentaram grandes regiões de homozigose na análise de CMA..... | 52 |
| Tabela 6: Taxa de diagnóstico genético dos 239 pacientes com EEDs de acordo com a técnica utilizando a análise otimizada..... | 53 |
| Tabela 7: Taxas de diagnóstico genético (DG) em diferentes estudos com pacientes com epilepsias de início precoce ou da infância..... | 55 |
| Tabela 8: Taxas de diagnóstico genético (DG) dos pacientes em relação ao país de origem..... | 57 |
| Tabela 9: SNPs associados a cada grupo mostrando suas posições, grupo em que estão associados e genes em que estão localizados..... | 60 |
| Tabela 10: Genes cujos grupos de variantes ultrarraras apresentaram p-valor significativo ($p \leq 0,05$) no grupo de pacientes totais..... | 62 |
| Tabela 11: Genes cujos grupos de variantes ultrarraras apresentaram p-valor significativo ($p \leq 0,05$) no grupo de pacientes com DG..... | 63 |
| Tabela 12: Genes cujos grupos de variantes ultrarraras apresentaram p-valor significativo ($p \leq 0,05$) no grupo de pacientes sem DG..... | 63 |
| Tabela 13: Lista EEDG contendo genes relacionados às EEDs..... | 80 |
| Tabela 14: Lista EG contendo genes relacionados às epilepsias em geral..... | 81 |
| Tabela 15: Lista PG contendo genes potenciais para epilepsias..... | 82 |
| Tabela 16: Lista de todos os pacientes analisados no estudo..... | 86 |

LISTA DE ABREVIATURAS E SIGLAS

ACMG – Colégio Americano de Genética Médica
ADNPM – Atraso do desenvolvimento neuropsicomotor
CAAE – Certificado de Apresentação e Apreciação Ética
CEP – Comitê de Ética em Pesquisa
CMA – Análise de microarranjos cromossômicos
CNV – Variantes de número de cópias
CTCG – Crises tônico-clônicas generalizadas
DG – Diagnóstico genético
EED – Encefalopatias epilépticas e do desenvolvimento
EEDG – Painel de genes ligados às EEDs
EEG – Eletroencefalograma
EEMES – Epilepsia com estado de mal elétrico do sono
EG – Painel de genes ligados às epilepsias
FCM – Faculdade de Ciências Médicas
FISH – Hibridização *in situ* por Fluorescência
GEFS+ - Epilepsia genética com crises febris plus
HC – Hospital de Clínicas
IBD – Identidade por descendência
ILAE – Liga Internacional contra Epilepsia
IMECC – Instituto de Matemática, Estatística e Computação Científica
Indel – Inserção/deleção
LD – Desequilíbrio de Ligação
LGS – Síndrome de Lennox-Gastaut
LKS – Síndrome de Landau-Kleffner
MAF – Frequência do alelo menor
NGS – Sequenciamento de nova geração]
NGSVA – Plataforma *NGS Variant Analyzer*
PCA – Análise de Componentes Principais
PCR – Reação em Cadeia da Polimerase

PG – Painel de genes candidatos/potenciais para epilepsias

POCS – ponta-onda contínua durante o sono

qPCR – PCR quantitativo em tempo real

QQ – Quantil-Quantil

RM – Ressonância magnética

ROH – Regiões de homozigose

SKAT-O – do inglês *optimized sequence kernel association test*

SNP – Polimorfismos de nucleotídeo único

SNV – Variantes de nucleotídeo único

TCLE – Termo de consentimento livre e esclarecido

UNICAMP – Universidade Estadual de Campinas

VUS – Variante de significado incerto

WES – Sequenciamento completo do exoma

WGS – Sequenciamento completo do genoma

WS – Síndrome de West

SUMÁRIO

| | |
|--|----|
| Introdução..... | 16 |
| Objetivos..... | 24 |
| Material e Métodos | |
| Seleção de Pacientes..... | 25 |
| Casuística..... | 26 |
| Extração de DNA e Processamento das amostras..... | 28 |
| Análise de WES Utilizando Plataforma <i>NGS Variant Analyzer</i> | 29 |
| Análise de WES Utilizando Plataforma <i>Varstation</i> | 32 |
| Análise de CMA Utilizando Plataforma <i>ChAS</i> | 32 |
| Análise de Regressão Logística..... | 34 |
| Análise de SKAT-O..... | 36 |
| Resultados e Discussão | |
| Sequenciamento Completo de Exoma..... | 37 |
| Análise de Microarranjos Cromossômicos..... | 46 |
| Taxa de Diagnóstico Genético..... | 51 |
| Análise de Regressão Logística..... | 58 |
| Análise de SKAT-O..... | 62 |
| Conclusão..... | 65 |
| Referências..... | 66 |
| Anexos | |
| I – Parecer do CEP..... | 74 |
| II – Termo de Consentimento..... | 77 |
| III – Listas de Genes Usadas nas Análises..... | 80 |
| IV – Artigo de Revisão Publicado como Primeira Autora..... | 84 |
| V – Artigo Publicado como Coautora..... | 85 |
| VI – Lista de Pacientes | 86 |

INTRODUÇÃO

Encefalopatias epiléticas e do desenvolvimento (EEDs) formam um grupo heterogêneo de epilepsias graves de início precoce, caracterizadas por diferentes tipos de crises de difícil controle e intensa atividade interictal no eletroencefalograma (EEG), geralmente resistentes ao tratamento medicamentoso e que levam ao risco de deterioração cognitiva e neuropsicomotora progressiva¹. A incidência é idade-dependente e é estimada em cerca de 45 casos a cada 100.000 crianças²⁻⁷, podendo chegar a mais de 60 casos a cada 100.000 crianças a nível populacional. Uma porção significativa dessas crianças (cerca de 35%) com epilepsias nos primeiros três anos de vida apresentam refratariedade e alto risco de morte precoce e de comorbidades cognitivas e comportamentais⁸. A taxa de mortalidade é alta, podendo chegar até 25%^{9,10}.

Segundo a Liga Internacional contra Epilepsia (ILAE, do inglês *International League Against Epilepsy*), existem três definições quando tratamos desse assunto: encefalopatias do desenvolvimento, onde ocorre apenas comprometimento do desenvolvimento, mas sem atividade epileptiforme frequente associada à regressão ou a um maior atraso do desenvolvimento; encefalopatia epilética, quando a própria atividade epileptiforme contribui para as alterações cognitivas e comportamentais, para além do esperado apenas pelo mecanismo da doença em si, onde não há atraso de desenvolvimento preexistente e não é provável que a mutação genética cause atraso por si só (malformação cortical, por exemplo); e encefalopatias epiléticas e do desenvolvimento, onde há atuação de ambos os fatores (atividade epileptiforme e comprometimento do desenvolvimento) e, por muitas vezes, não é possível distinguir qual dos dois contribui mais para a apresentação clínica do paciente¹¹⁻¹³.

É amplamente conhecido que as EEDs possuem uma vasta gama de etiologias, incluindo, por exemplo, lesões cerebrais estruturais congênitas ou adquiridas, distúrbios metabólicos, e causas genéticas^{1,14}. O diagnóstico das EEDs ainda é baseado inicialmente e predominantemente nas características clínicas e de EEG e sua etiologia permanece desconhecida na maioria dos pacientes¹⁵. Os avanços nos estudos de genética molecular demonstraram que uma proporção desses pacientes tem mutações em genes, muitos deles associados à função de canais iônicos voltagem dependentes expressos no sistema nervoso central¹⁶.

As evidências de que há predisposição genética em síndromes epiléticas existem desde a década de 1950^{17,18}. A partir de estudos tradicionais de análise de ligação em grandes famílias onde o gene segregava junto com a doença começaram a relacionar genes envolvidos com canais iônicos às epilepsias, o que ficou conhecido como a era das canalopatias.

A primeira mutação relacionada com uma forma de epilepsia idiopática foi descrita em 1995 por Steinlein e colaboradores. Eles identificaram uma mutação no gene *CHRNA4* que codifica uma subunidade de um receptor acetilcolinérgico em uma forma de epilepsia focal¹⁹. Em 1998, Charlier e colaboradores e Singh e colaboradores identificaram as primeiras mutações em genes que codificam subunidades de canal de potássio voltagem dependente (*KNCQ2* e *KCNQ3*)^{20,21}, enquanto que Wallace e colaboradores identificaram mutações no gene de canal de sódio *SCN1B* em uma família com epilepsia genética com crises febris plus (GEFS+ do inglês *genetic epilepsy with febrile seizures plus*)²². Em meados de 2000 Escayg e colaboradores identificaram, em duas famílias também com GEFS+, mutações no gene *SCN1A* que, mais tarde foi o primeiro gene associado a uma EED^{23,24}.

Após um período de transição, a partir de meados de 2010, o número de mutações associadas a epilepsias vem crescendo muito rapidamente, principalmente para síndromes epiléticas monogênicas raras^{16,25}. Esse crescimento se deu, principalmente, pelo avanço de técnicas na área da Genética Molecular, como o sequenciamento de nova geração (NGS, do inglês *next generation sequencing*), o que ficou conhecido como a era do NGS. A tecnologia de NGS permite sequenciar porções significativas do genoma de forma rápida, fácil e com alto custo-benefício^{26,27}. A linha do tempo com os genes separados dentro de cada uma dessas eras pode ser visualizada na Figura 1²⁸.

O sequenciamento de apenas um gene pela técnica de Sanger ainda pode ser utilizado para identificar variantes em genes que já estão associados a doenças monogênicas. O sequenciamento do gene *SCN1A*, por exemplo, é um ótimo candidato para a síndrome de Dravet, uma vez que mutações nesse gene podem explicar cerca de 80% dos casos²⁹⁻³³. Apesar de o sequenciamento de Sanger ser uma técnica eficaz e ainda muito utilizada para identificar mutações em um único gene³⁴, o NGS é uma forma muito mais rápida e custo-efetiva para isso.

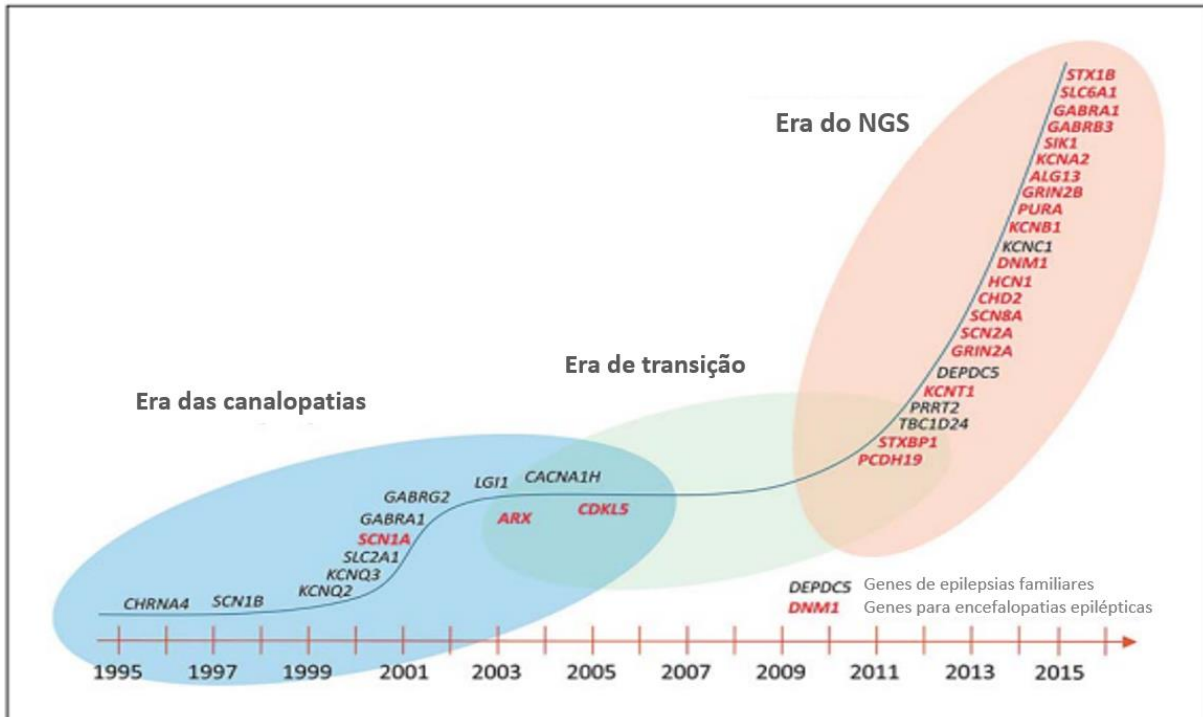


Figura 1: Linha do tempo da descoberta de genes em epilepsias. Os genes em vermelho correspondem às epilepsias eifépticas e os genes em preto às epilepsias. NGS: sequenciamento de nova geração. Adaptado de I. Helbig, A.N.A. Tayoun / Mol Syndromol 2016;7:172–181.

Outra alternativa quando falamos em diagnóstico molecular é o painel de NGS, onde vários genes podem ser sequenciados de uma só vez^{34,35}. Os painéis podem ser específicos para cada doença e são customizáveis. Atualmente existem muitos painéis comerciais para epilepsias que contêm de 100 a mais de 300 genes³⁶.

A técnica mais utilizada atualmente para diagnóstico de doenças mendelianas raras causadas por variantes de alta penetrância é o sequenciamento completo de exoma (WES, do inglês *whole exome sequencing*)^{27,37,38}, que tem o poder de capturar e sequenciar os éxons da maior parte dos 20.000 genes codificantes do genoma humano. Essa opção possui uma alta taxa de diagnóstico para doenças neurológicas^{34,39,40}.

Por último, temos a técnica de sequenciamento completo do genoma (WGS, do inglês *whole genome sequencing*). O WGS sequencia todo o genoma do paciente, incluindo regiões codificantes, não codificantes, e regulatórias e também permite uma alta taxa de diagnóstico para doenças genéticas. Por apresentar uma cobertura mais uniforme, permite até a identificação de variantes em região codificante que não estariam presentes em dados de WES. Porém, o alto custo dessa técnica e a alta

complexidade das análises de bioinformática devido ao grande volume de dados, fazem com que o WES seja mais utilizado como teste genético ainda hoje. Com o avanço dos estudos utilizando WGS e com a queda do custo, essa técnica será, no futuro, a mais aconselhada para uma avaliação mais abrangente em pacientes com doenças neurológicas não diagnosticadas³⁴.

Utilizando as estratégias de NGS é possível identificar diferentes tipos de variantes genéticas, como variantes de nucleotídeo único (SNV, do inglês *single nucleotide variants*), inserções e deleções (indels) e variantes de número de cópias (CNV, do inglês *copy number variants*). As CNVs são importantes causas de doenças do neurodesenvolvimento³⁴ e, apesar de poderem ser detectadas também através de NGS (preferencialmente no WGS), a principal técnica hoje em dia para identificação desse tipo de variante é a análise de microarranjos cromossômicos (CMA, do inglês *chromosomal microarray analysis*).

Estudos recentes mostram que as estratégias de NGS possuem uma taxa de diagnóstico entre 15 e 50% para as epilepsias, sendo essa taxa ainda mais alta (40 a 80%) para pacientes com idade de início antes dos dois meses, ou em certos grupos de doenças, como a síndrome de Ohtahara^{29,34,41}. Para as EEDs a taxa de diagnóstico molecular utilizando WES varia de 11 a 70%⁴²⁻⁴⁵ e pode chegar a cerca 80% nos casos de Dravet com o gene *SCN1A*, como dito anteriormente. No caso das CNVs, a taxa de diagnóstico molecular nas EEDs é de 5 a 19%⁴⁶⁻⁴⁸. A maioria dos pacientes possui variantes patogênicas *de novo* em heterozigose, mas também são observadas variantes com herança autossômica recessiva e ligada ao cromossomo X^{1,49}.

Apesar de a grande maioria das causas de EEDs já descritas ser monogênica, alguns genes modificadores já foram propostos, mas ainda é um assunto controverso. São genes que não são a causa primária da doença, mas podem modificar o fenótipo e por isso devem ser levados também em consideração. Variantes em *CACNB4*, por exemplo, aumentam a liberação de neurotransmissores em neurônios excitatórios na condição de insuficiência de neurônios inibitórios causada por *SCN1A*, melhorando o fenótipo⁵⁰. Variantes sozinhas em *SCN9A* podem ser assintomáticas ou causar crises febris infrequentes devido a penetrância incompleta e expressividade variável, mas contribui de forma multifatorial para a síndrome de Dravet⁵¹. Variantes em *scn2a*, *scn8a* e *kcnq2* podem influenciar o fenótipo em modelos animais com GEFS+ com a mutação *Scn1a*-R16648H⁵². A perda de função do gene *Hlf* modifica o fenótipo em

modelos animais *Scn2a*Q54 e *Scn1a*+/- de epilepsia⁵³. Já a diminuição da expressão de *Cacna1g* atua como modificador aumentando a sobrevivência e diminuindo a frequência de crises espontâneas em modelos animais com variantes em *Scn2a*Q54 e *Scn1a*+/-^{54,55}.

Um diagnóstico molecular conclusivo deve ser baseado em diversos parâmetros, incluindo o conhecimento prévio da literatura sobre a relação entre o suposto gene causal e o fenótipo do paciente. Essa relação entre o genótipo e fenótipo é muito complexa para a maioria das EED que já possuem genes associados. Nesses casos podemos observar uma alta heterogeneidade tanto fenotípica (onde mutações em um mesmo gene podem causar diferentes fenótipos), quanto genética (onde mutações em diferentes genes podem causar o mesmo fenótipo)¹². Essas duas características das EED representam um grande desafio para a prática clínica e para identificação das síndromes específicas.

Contudo, é inquestionável que a identificação de variantes patogênicas em pacientes com EED proporciona muitos benefícios para esses e para suas famílias, ainda mais quando o teste genético é realizado de forma precoce, reduzindo o tempo e o custo associados à odisséia diagnóstica. A confirmação da natureza genética da doença, além do impacto emocional positivo significativo, permite a busca de aconselhamento genético, o estabelecimento mais acurado do prognóstico, oportunidades de pesquisa na área, permite a conexão de famílias com o mesmo diagnóstico através de grupos de apoio e, mais importante, permite a busca por terapias personalizadas e mais adequadas⁵⁶. Como exemplo disso podemos citar o caso de mutações com perda de função no gene *SCN1A*, onde bloqueadores de canais de sódio, como a carbamazepina, oxcarbazepina e fenitoína, podem piorar o controle das crises e, portanto, devem ser evitados⁵⁷. No caso de mutações com ganho de função em *SCN8A* ocorre o oposto e há boa resposta aos fármacos anticrises bloqueadores de canal de sódio⁵⁸. Já para pacientes com mutações nos genes *SLC2A* e *PDHA1* a melhor opção de tratamento é a dieta cetogênica^{59,60}.

As EEDs compreendem uma série de síndromes com diferentes características clínicas e de EEG. Como podemos observar na Figura 2, elas são relacionadas à idade de início das crises que, por muitas vezes se sobrepõe, assim como a semiologia das crises e os fenótipos. Como já foi dito, essa classificação clínica do tipo de síndrome do paciente é um desafio, permanecendo, na maioria das vezes,

como uma EED indeterminada. A seguir descrevemos brevemente, com informações relevantes ao que será discutido, as síndromes que foram identificadas e incluídas nesse trabalho.

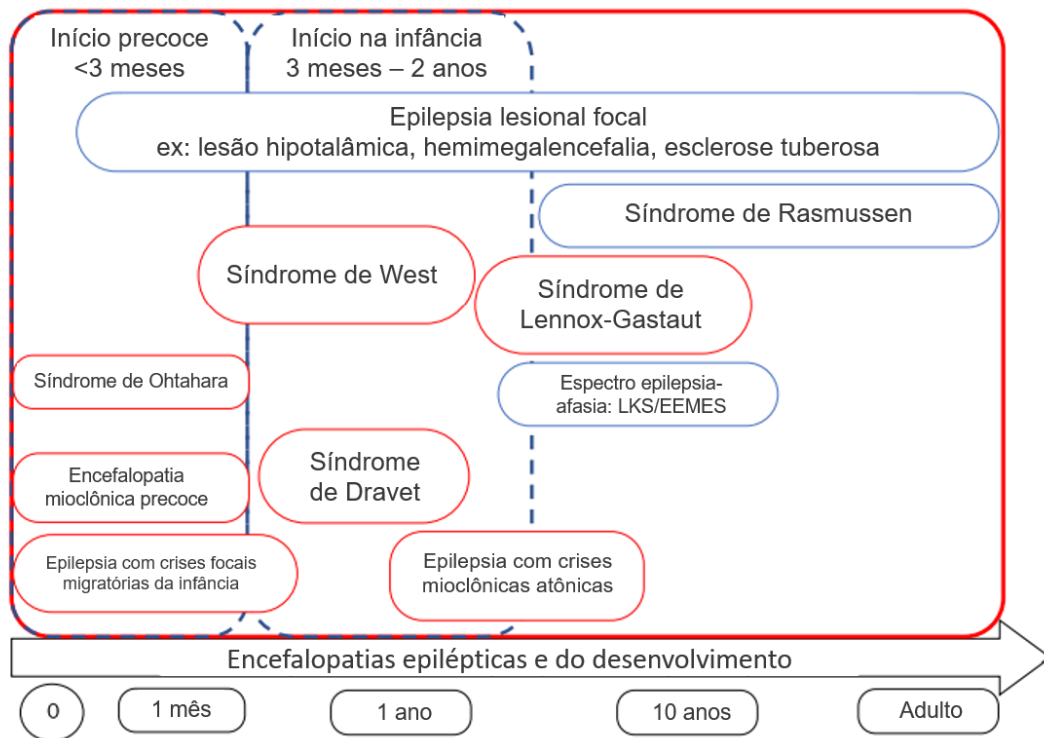


Figura 2: Representação de algumas das EED bem conhecidas em relação às diferentes idades de início das crises (eixo X). O tamanho das caixas reflete as frequências das síndromes. Adaptado de I.E. Scheffer, J. Liao / *European Journal of Paediatric Neurology* 24 (2020) 11e14.

A síndrome de espasmos epilépticos infantis é um termo usado para englobar tanto pacientes com síndrome de West (WS, do inglês *West syndrome*), quanto pacientes que apresentam espasmos epilépticos mas que não preenchem critérios para WS⁶¹. Aqui foram incluídos somente pacientes com WS que apresentaram atraso do desenvolvimento neuropsicomotor (ADNPM). A WS tem início entre um e 24 meses de idade, sendo que o pico é entre três e 12 meses. Inícios mais tardios são raros, mas podem ocorrer⁶². A incidência é de cerca de 30 casos a cada 100.000 crianças e é maior no sexo masculino, porém, ambos os sexos são afetados^{8,61}. Quanto à etiologia, metade a dois terços das crianças possuem alterações estruturais que podem ser congênicas ou adquiridas. Após resultado de ressonância magnética (RM) normal, uma etiologia genética deve ser considerada. A taxa de diagnóstico genético (DG) pode chegar até 41% e são observadas variantes patogênicas *de novo* na

maioria dos casos. Dentre as causas genéticas já conhecidas temos: *CDKL5*, *DEPDC5*, *STXBP1*, *TSC1*, *TSC2*, alterações cromossômicas estruturais, CNVs, entre outros. Etiologias metabólicas são raras, mas também são importantes e devem ser consideradas se causas genéticas e estruturais forem descartadas^{8,61,63}.

A síndrome de Dravet tem início dentro do primeiro ano de vida^{64,65} e a incidência é de aproximadamente 6,5 casos a cada 100.000 crianças^{8,63}. É importante ressaltar que, na grande maioria dos casos, a etiologia é genética e cerca de 80% dos pacientes possuem variantes patogênicas *de novo* no gene *SCN1A*²⁹⁻³³. Entretanto, o diagnóstico de síndrome de Dravet envolve diversas características clínicas e não deve ser feito baseado somente na variante genética encontrada, já que variantes em *SCN1A* podem ser observadas em outras síndromes epiléticas⁶⁶. Outros genes já associados a essa síndrome são: *GABRG2*, *GABRA1*, *STXBP1*, *SCN1B*, entre outros⁶⁷.

A síndrome de Doose, também conhecida por epilepsia com crises mioclônica-atônicas, possui incidência de um caso a cada 10.000 crianças e representa aproximadamente 2% de todas as epilepsias da infância⁶⁸. O início das crises ocorre entre seis meses a oito anos de idade (pico de dois a seis anos) e indivíduos do sexo masculino são mais afetados⁶⁹. Os exames de imagem nesses pacientes são normais e, na maioria das crianças, há envolvimento de herança complexa com padrão poligênico. Em um terço dos casos existe histórico familiar de epilepsia ou crises febris, o que favorece o prognóstico. Alguns dos genes já associados a essa síndrome são: *SCN1A*, *SCN1B*, *SCN2A*, *STXBP1*, *SLC6A1*, *CHD2*, *SYNGAP1*, *NEXMIF*, *KIAA2022* e, em 5% dos casos, *SLC2A1* que está associado à deficiência de GLUT1⁷⁰.

Na síndrome de Lennox-Gastaut (LGS, do inglês *Lennox-Gastaut syndrome*) as crises tem início principalmente entre 18 meses a 8 anos de idade (pico entre três e cinco anos), porém, somente em 0,6% dos casos o diagnóstico ocorre no início das crises. Essa síndrome é responsável por um a 2% de todas as pessoas com epilepsias e é ligeiramente mais comum no sexo masculino. Cerca de 20% dos casos de LGS evolui de síndrome dos espasmos epiléticos infantis⁷⁰. Além disso, cerca de 3,6% de todas as crianças com epilepsias e 19% das crianças com início das crises na infância, também evoluem para LGS^{70,71}. Etiologias estruturais são mais comuns e exame de RM no início das crises é o mais recomendado, podendo ter um grande impacto na

tomada de decisões e tratamentos⁷². Após exame de RM normal, testes genéticos devem ser considerados. São observadas variantes patogênicas, geralmente, *de novo*^{42,73}, alterações cromossômicas estruturais e CNVs. O teste genético também pode ser recomendado para detectar a causa por trás da alteração estrutural encontrada no exame de imagem. Raramente a LGS é associada a uma doença metabólica que deve ser considerada após as etiologias estrutural e genética serem descartadas⁷⁰.

Por último, temos as síndromes epiléticas associadas a atividade ponta-onda contínua durante o sono (POCS). Nesse grupo estão duas síndromes muito raras conhecidas como epilepsia com estado de mal elétrico do sono (EEMES) e síndrome de Landau-Kleffner (LKS, do inglês *Landau-Kleffner syndrome*). O início é entre dois e 12 anos de idade (pico entre quatro e cinco anos) e ambos os sexos são igualmente afetados⁷⁰. A etiologia pode ser estrutural com alterações adquiridas ou relacionadas ao desenvolvimento. Em alguns casos existe uma base genética para a doença que pode seguir modelo de herança monogênica ou complexa, sendo o gene *GRIN2A* a causa monogênica mais comum^{70,74}. Histórico familiar é observado em até 50% dos pacientes⁷⁵.

Além dessas síndromes existem as síndromes com etiologias específicas, como a EED por *KCNQ2*, EED por *CDKL5* e a EED dependente de piridoxina (relacionada ao gene *ALDH7A1*), por exemplo. Nesses casos existe uma etiologia específica para a epilepsia que está associada a um fenótipo clínico claramente definido, relativamente uniforme e diferenciado na maioria dos indivíduos afetados. Existe uma tendência de aumento do número dessas síndromes com o tempo, à medida que novas etiologias, principalmente genéticas forem sendo descobertas⁷⁰.

É evidente que, apesar de uma grande parcela dos pacientes com EED possuir mutações de efeito monogênico e causas conhecidas, uma proporção significativa fica sem diagnóstico etiológico, mesmo com a obtenção de dados genômicos. Por outro lado, a disponibilidade dos dados de WES em um número crescente de pacientes com EEDs cria a oportunidade de aplicar novas análises considerando, também, outros modelos mais complexos de herança (modelo poligênico, por exemplo, onde variantes de efeito menor podem se combinar para produzir o efeito fenotípico final). E foi nesse contexto que esse trabalho foi conduzido.

OBJETIVOS

Objetivo geral

Caracterizar, do ponto de vista genético, uma grande coorte de pacientes latino-americanos com EEDs utilizando modelo monogênico de herança e investigar o possível envolvimento de modelos mais complexos de herança.

Objetivos específicos

- Obter amostras de pacientes com EEDs com caracterização clínica adequada provenientes de diferentes países da América Latina;
- Realizar as técnicas de sequenciamento completo de exoma e de microarranjos cromossômicos a partir das amostras de DNA desses pacientes e obter dados genômicos a partir de diferentes tipos de processamentos;
- Utilizar os dados genômicos gerados para investigar, em todos os pacientes, a presença de variantes de sequência (SNV e indel) utilizando dois tipos de análises para sequenciamento completo de exoma; e de variantes cromossômicas estruturais (CNV) utilizando ambas as análises de exoma e microarranjos, visando obter o maior rendimento diagnóstico possível seguindo um modelo monogênico, ou seja, um gene ou variante de efeito maior causando o fenótipo;
- Explorar evidências de envolvimento de modelos complexos de herança nos pacientes em que não foi possível encontrar uma causa monogênica através de análises complexas de regressão logística e de SKAT-O (do inglês, *optimized sequence kernel association test*), envolvendo variantes comuns, raras e ultrarraras.

MATERIAL E MÉTODOS

Seleção e Coleta dos Pacientes

Foram selecionados para esse estudo pacientes com diagnóstico clínico de EED de acordo com os critérios clínicos propostos pela ILAE, incluindo: atraso do desenvolvimento neuropsicomotor após o surgimento das crises, epilepsia farmacorresistente, ausência de alterações nos exames de imagens que pudesse explicar o quadro, ausência de suspeita de doenças metabólicas e não houve limitação de idade. A maior parte dos pacientes foi recrutada no ambulatório de epilepsias da infância do Hospital de Clínicas (HC) da Faculdade de Ciências Médicas da Universidade Estadual de Campinas (FCM-UNICAMP). Recebemos também amostras de pacientes de centros colaboradores de outras regiões do Brasil e de outros países da América Latina. Para as amostras provenientes de outros países recebemos o apoio logístico da ILAE pela Comissão Regional para a América Latina. Essa é a primeira vez que é realizado um estudo em larga escala com EED na América Latina.

Os pacientes do HC-UNICAMP foram acompanhados prospectivamente no ambulatório de epilepsias da infância e os dados clínicos foram revisados e discutidos pela equipe clínica responsável: Dra. Ana Carolina Coan, neurologista com especialidade em epilepsia da infância; Dra. Maria Augusta, neuropediatra; Dra. Íscia Lopes Cendes, médica geneticista; Dr. Fernando Cendes, neurologista, especialista em epilepsia e neuroimagem. As discussões tiveram como objetivo chegar a um consenso sobre a inclusão ou exclusão dos pacientes, bem como classificar a EED a qual o paciente pertence de acordo com os padrões clínicos e eletroencefalográficos apresentados. No caso das amostras enviadas, os médicos responsáveis pelos pacientes responderam questionários clínicos padronizados e os dados também foram revisados e discutidos por nossa equipe clínica do HC-UNICAMP.

É importante ressaltar que a preocupação com a qualidade da informação clínica em nosso estudo, o que garante a homogeneidade fenotípica dos pacientes incluídos, é um dos pontos fortes do trabalho, e só foi atingida pelo meu grande envolvimento com a equipe clínica, e é claro, pela excelência do grupo de pesquisa em epilepsia infantil do nosso centro chefiado pelas Profas. Dras. Ana Carolina Coan

e Maria Augusta Montenegro, ambas com mais de uma década de experiência na avaliação de crianças com epilepsia grave.

Esse trabalho teve aprovação do Comitê de Ética em Pesquisa (CEP) da UNICAMP e o Certificado de Apresentação e Apreciação Ética (CAAE: 12112913.3.0000.5404) se encontra em anexo. Todos os responsáveis pelos pacientes concordaram em participar da pesquisa e assinaram o termo de consentimento livre e esclarecido (TCLE) que também se encontra em anexo.

Para as amostras do HC-UNICAMP, a coleta procedeu-se de acordo com a disponibilidade e possibilidade do paciente. Foram coletadas amostras de sangue periférico quando possível e, em caso de impossibilidade, foram coletadas amostras de células de mucosa bucal. No caso dos centros colaboradores recebemos amostras de sangue periférico e/ou amostras de DNA já extraído. Quando possível, foram realizadas também as coletas dos pais, irmãos afetados ou não e/ou outros familiares. Não foram realizadas análises de trio para as técnicas avaliadas e essas amostras de familiares foram utilizadas nas etapas de validações para determinação da segregação e do padrão de herança das variantes genéticas selecionadas.

Casuística

Reunimos um total de 275 pacientes diagnosticados clinicamente com EEDs. Em casos de baixa quantidade e qualidade do DNA, os pacientes eram recoletados para nova extração, porém, isso não foi possível para 36 destes pacientes, que foram excluídos da casuística. Dos 239 pacientes com amostras de DNA (Anexo VI), 130 são do sexo masculino (54,4%) e 109 são do sexo feminino (45,6%). Em relação ao local de coleta, 124 deles foram coletados no HC-UNICAMP, 22 foram enviados de outros centros colaboradores no Brasil, 57 enviados da Argentina, 15 do Chile, 14 do Peru, quatro do Uruguai e três de Honduras. É importante ressaltar a dificuldade em se obter amostras provenientes de outros centros colaboradores, principalmente de outros países. Isso foi possível devido ao envolvimento do nosso grupo com diferentes serviços clínicos que possibilitou o contato com outros médicos que se disponibilizaram para colaborar com a pesquisa e se comprometeram com a coleta, extração, armazenamento e transporte adequados das amostras. Além disso, trata-se

do primeiro estudo com EEDs realizado em larga escala com pacientes provenientes de diferentes regiões da América Latina.

Foi possível obter informações sobre as idades no momento da coleta de 159 pacientes e sobre as idades de início das crises de 175 pacientes. A idade dos pacientes no momento da coleta em meses variou de 1 até 444 meses (média=104,7 \pm 72,7) e a de idade de início das crises variou de zero (dentro do primeiro mês de vida) a 168 meses (média=24,7 \pm 31,3).

Quanto ao fenótipo, os pacientes foram classificados de acordo com a Figura 3. Após análise clínica detalhada de nossos médicos colaboradores chegamos a oito grupos classificados de acordo com as hipóteses diagnósticas.

O grupo das EEDs indeterminadas reúne aqueles pacientes que cumpriram todos os nossos critérios de inclusão, mas que não foi possível identificar a síndrome específica, achado que ilustra bem a dificuldade da classificação clínica da síndrome específica. Como consta na introdução deste trabalho, as EEDs compreendem uma grande proporção das síndromes epiléticas de início precoce e esses fenótipos se sobrepõem e evoluem no tempo; portanto, após novas consultas testes o diagnóstico clínico ainda pode ser alterado, fato que ressalta a importância da avaliação clínica prospectiva e cuidadosa desses pacientes. Os pacientes com fenótipo “síndrome de Doose vs. Síndrome de Lennox-Gastaut” não foram especificamente categorizados devido a sobreposição que ocorre na semiologia das crises nessas duas síndromes⁷⁶.

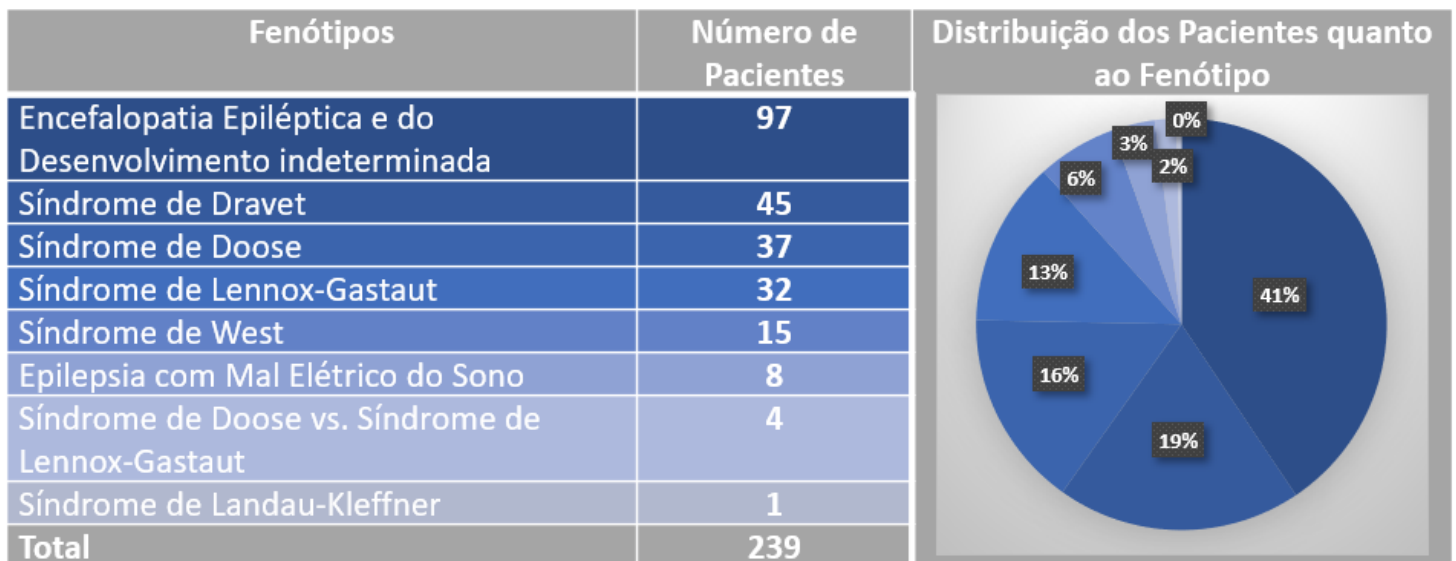


Figura 3: Classificação e distribuição dos pacientes com EED quanto ao fenótipo de acordo com a hipótese diagnóstica.

Extração de DNA e Processamento das Amostras

Para as amostras de sangue periférico, a extração de DNA genômico de linfócitos foi realizada utilizando protocolo fenol-clorofórmio-álcool-isoamílico. Para as amostras de células da mucosa bucal foi utilizado o kit *OraCollect* da marca *DNA Genotek* e a extração seguiu protocolo indicado pelo fabricante.

Foram realizadas as verificações da integridade e da qualidade e a quantificação do DNA extraído. A integridade do DNA foi testada por eletroforese em gel de agarose 1% para verificação do grau de fragmentação. Os padrões de pureza foram medidos com um espectrofotômetro modelo Epoch da marca *BioTek*, onde os valores de absorbância 260/280nm entre 1.8 e 2.0 e taxa de composto fenólico 320 Raw menor de 0.1, foram considerados como ideais para qualidade do DNA. A quantificação também foi realizada pelo Epoch e os DNAs foram diluídos até atingirem concentração entre 50 e 60ng/μL para serem processados utilizando as técnicas de WES e CMA.

Para o WES, as amostras foram aliquotadas e enviadas à empresa *Macrogen, Inc.* para preparação das bibliotecas e sequenciamento. Foram enviados cerca de 70μL de cada amostra a uma concentração de 50ng/μL e, essa quantificação final foi realizada utilizando sistema *Qubit*. A preparação das bibliotecas seguiu protocolo recomendado pelo fabricante utilizando o kit *SureSelectXT Human All Exons v6 (Agilent Technologies)* que é otimizado para 3 μg de DNA genômico total. Já o sequenciamento foi realizado parte em plataforma *Illumina Hiseq 4000* e parte em plataforma *Illumina Novaseq 6000*. Os dados brutos dos sequenciamentos (arquivos FASTQ) foram enviados e armazenados nos servidores do laboratório de Biologia Computacional e Bioestatística da FCM-UNICAMP e do Instituto de Matemática, Estatística e Computação Científica (IMECC-UNICAMP), chefiados pela Profa. Íscia Lopes Cendes e pelo Prof. Benilton Carvalho.

Para CMA, utilizamos também as amostras já quantificadas para realização da técnica em nosso laboratório. Foram utilizados chips *Genome-Wide Human SNP Array 6.0 (Affymetrix™)*, que possuem mais de 1.8 milhões de marcadores genéticos, incluindo 906600 polimorfismos de base única (SNP, do inglês *single nucleotide polymorphisms*) e 946000 sondas para detecção de CNVs. Os passos dessa técnica seguiram protocolo recomendado pelo fabricante e podem ser observados na Figura

4. Após escaneamento dos chips utilizando o *GeneChip® Scanner 3000 7G* (*Affymetrix®*), foram obtidos os arquivos brutos para análise (arquivos CEL) que também se encontram armazenados em nossos servidores.

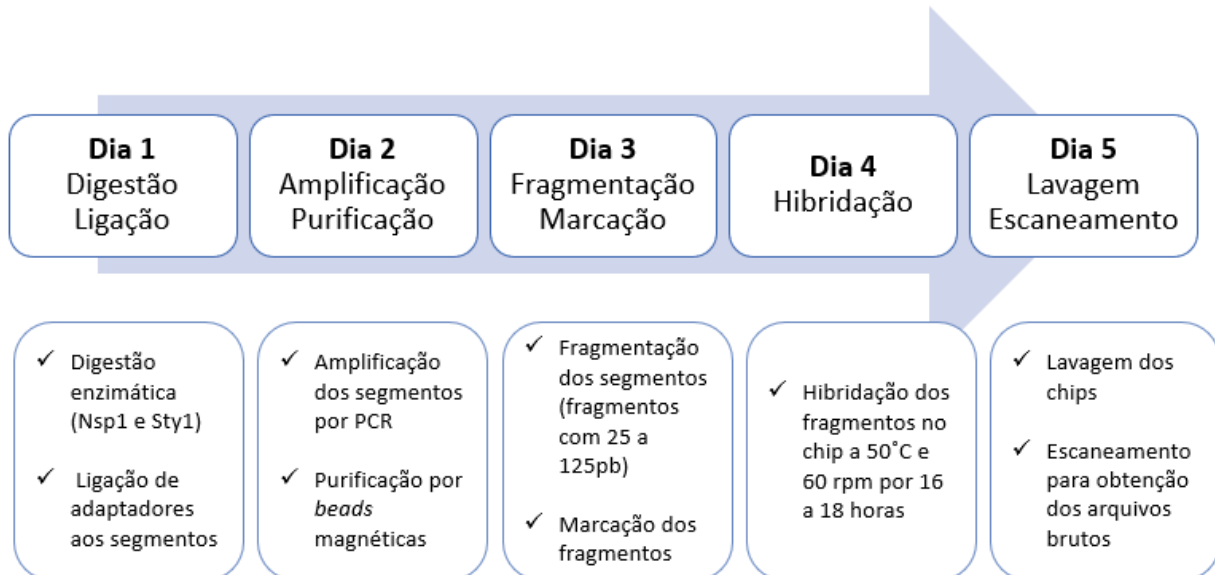


Figura 4: Passos da técnica de CMA utilizando protocolo *Genome-Wide Human SNP Array 6.0* (*Affymetrix™*).

Os arquivos brutos resultantes de cada uma dessas duas técnicas individualmente foram analisados em relação a presença ou ausência de variantes com potencial patogênico de diversas formas diferentes que serão descritas nas seções seguintes. Os dados também foram analisados em conjunto (WES+CMA) de duas formas diferentes (regressão logística e SKAT-O) para testar a hipótese do envolvimento de herança complexa em dados de pacientes em que não foi possível encontrar uma causa genética.

Análise de WES Utilizando Plataforma NGS *Variant Analyzer*

Essa etapa do trabalho foi desenvolvida durante o estágio no exterior realizado entre setembro de 2019 e fevereiro de 2021 na Universidade Friedrich-Alexander, Erlangen-Nürnberg, Alemanha. O estágio foi realizado no Instituto de Genética Humana do Hospital de Clínicas da Universidade sob supervisão do Prof. Dr. André Reis com o intuito de analisar os arquivos brutos de WES e CMA para fins de diagnóstico genético (DG).

Os arquivos brutos provenientes dos WES (arquivos FASTQ) foram mapeados ao genoma de referência (*GRCh37/hg19*) utilizando o programa *BWA MEM* (versão 0.7.17). Em seguida, as duplicações foram removidas utilizando *MarkDuplicates* (*Picard*; versão 2.18.21) e as regiões de inserções/deleções foram realinhadas utilizando *GATK* (versão 3.8-1). A qualidade de mapeamento de todas as leituras que tinham MQ igual a zero foi reajustada para 60 manualmente usando *AWK* e *SAMtools*. Para a chamada de variantes, cinco programas foram utilizados: *GATK HaplotypeCaller* (versão 3.8-1); *GATK UnifiedGenotyper*; *Freebayes* (versão 1.2.0); *Platypus* (versão 0.8.1); e *SNVer*.

Em seguida, as chamadas do *Freebayes* foram filtradas para ter uma qualidade mínima de 0,1. Todas as variantes foram então transferidas para a base de dados de análises do Instituto e anotadas utilizando *AnnoVar*. O processamento foi realizado pela equipe de bioinformática do grupo do Prof. Reis.

Para visualização, filtragem e análise das variantes foi utilizado o programa *NGS Variant Analyzer* (NGSVA), desenvolvido também pelo grupo do Prof. Dr. André Reis. Os WES foram carregados na plataforma um por um utilizando os seguintes parâmetros para filtragem das variantes:

- Variantes com frequência na população geral menor ou igual a 0,01 (ou 1%) em bases de dados;
- Variantes presentes em quatro ou menos controles internos (cerca de 960 indivíduos normais da população alemã, constituída de alemães e turcos);
- Variantes exônicas e/ou de sítios de *splicing*;
- Variantes com função exônica não sinônima;
- Variantes com profundidade igual ou maior que cinco;
- Novos alelos de variantes heterozigotas com profundidade igual ou maior que três;
- Variantes chamadas pelos programas *GATK HaplotypeCaller* e *GATK UnifiedGenotyper*.

Após filtragem das variantes foram aplicados dois painéis de genes para priorização das variantes. O primeiro painel é formado por 339 genes conhecidos relacionados epilepsias, incluindo as EED (EG) e o segundo painel é formado por 848

genes potenciais para epilepsia (PG). As listas dos genes contidos em cada painel se encontram nas tabelas do anexo III. As variantes localizadas nesses genes foram então analisadas uma por uma utilizando os seguintes parâmetros:

- Cobertura e profundidade;
- Frequências interna (utilizando os controles internos) e na população geral (utilizando bases de dados como *gnomeAD*, *1000 Genomes* e *ExAC*);
- Padrão de herança (se a zigosidade condiz ou não com o padrão de herança da doença e dos genes);
- Potencial patogênico dado pelos algoritmos preditores e *CAAD score*;
- Presença e classificação em bancos de dados como *dbSNP* e *ClinVar*;
- Correlação do gene com os fenótipos já observados em bancos de dados como *OMIM*.

Após análise, as variantes de interesse selecionada foram classificadas de acordo com os critérios estabelecidos pelo Colégio Americano de Genética Médica (ACMG, do inglês *American College of Medical Genetics*) com o auxílio da plataforma *Franklin/Genoox* (<https://franklin.genoox.com/>). As variantes consideradas causais foram então validadas nos pacientes e nos pais (quando coletados) pelo método sequenciamento de Sanger utilizando oligonucleotídeos iniciadores (*primers*) que foram desenhados especificamente para cada variante.

Ainda utilizando os dados de WES, foi realizada também a chamada de CNVs em ambiente R através do pacote *ExomeDepth* (versão 1.1.15) utilizando os alvos do *SureSelect* como exons de referência. Em seguida as CNVs foram transferidas para a base de dados e visualizadas e analisadas através do programa NGSVA. Para cada paciente as CNVs foram filtradas utilizando filtro de tamanho (≥ 1000 pb) e filtro de frequência (CNVs presentes em menos de 10 amostras). Após filtragem as CNVs foram priorizadas utilizando um painel de 127 genes associados a EEDs (EEDG; Anexo III). As CNVs consideradas causais foram validadas por PCR em tempo real (do inglês, *polymerase chain reaction*), também conhecido como PCR quantitativo (qPCR). Para isso foram desenhadas sondas específicas para os genes contidos na CNVs e foi utilizado *SYBR Green Master Mix (Applied Biosystems™)* para a preparação das reações.

Análise de WES Utilizando Plataforma *Varstation*

A plataforma *Varstation*⁷⁷ permite a análise a partir dos arquivos de dados brutos (arquivos FASTQ) e integra todos os passos para obtenção da lista de variantes de interesse para cada paciente (arquivos VCF). Essa etapa foi realizada para complementar a análise anterior já que permite a análise de variantes com uma maior liberdade em relação aos filtros. Aqui foram analisados com maior atenção os pacientes em que não foi possível encontrar uma variante causal na primeira análise.

Os arquivos FASTQ foram transferidos para a plataforma e foram processados utilizando o genoma de referência *GRCh37/hg19* com *pipeline* padrão específico para o protocolo de preparação da biblioteca (*Exome_Sureselect_all_exon_v6*). Após esse processamento inicial foram realizadas as etapas de identificação e priorização de variantes. As listas de variantes encontradas foram ordenadas por grupos de classificação e as consideradas patogênicas ou provavelmente patogênicas pelos critérios da ACMG foram priorizadas e analisadas sem filtros. O gene *SCN1A* também foi analisado sem filtro, devido a sua alta relação com as EEDs. Em seguida, as variantes foram filtradas em três etapas. Primeiramente priorizamos as variantes com profundidade de no mínimo 5x, com frequência populacional menor que 1% nas bases de dados (*gnomeAD*, *1000 Genomes*, *ExAC*, *ABRaOM*, etc.) e com consequência exônica não sinônima. Nas duas etapas seguintes as variantes já filtradas foram analisadas utilizando os dois painéis de genes (EG e PG) e os mesmos parâmetros descritos na seção anterior.

Após análise, as variantes de interesse selecionadas foram também classificadas de acordo com os critérios da ACMG com o auxílio da plataforma *Franklin/Genoox* e também validadas utilizando *primers* específicos e sequenciamento de Sanger.

Análise de CMA Utilizando Plataforma *ChAS*

Essa etapa também foi realizada durante o estágio no exterior. Os arquivos brutos (arquivos CEL) gerados pelo scanner ao final do protocolo foram analisados pelo programa *Genotyping Console 4.2 (Affymetrix™)* para determinação da qualidade dos dados. Em seguida, foram feitas as análises de genotipagem (chamada

de SNP), de número de cópias (chamada de CNV) e de regiões de homozigose (ROH, do inglês *regions of homozygosity*). ROHs são longas regiões em homozigose consecutivas no genoma, provavelmente devido a segmentos que foram compartilhados de forma idêntica por descendência. Esses segmentos podem ser resultado de consanguinidade, dissomia uniparental, redução do tamanho da população ou seleção natural e são regiões de interesse na análise do WES para identificação de alterações genéticas com herança autossômica recessiva⁷⁸. Para chamadas de SNP, o programa utiliza os algoritmos *Birdseed V1* e *Birdseed V2*, já para as chamadas de CNV e ROH o programa utiliza os algoritmos *CNV5* e *BRLMM-P+*. Para ambas as análises foi usado genoma de referência *GRCh37/hg19*.

Após chamada de CNV, os arquivos gerados (arquivos CNCHP) foram importados para o programa *Chromosome Analysis Suite 4.0 (ChAS; Thermo Fisher Scientific)* para visualização e análise. As CNVs e ROH de cada paciente foram filtradas da seguinte forma:

- Para duplicações (ganhos) foram analisadas CNVs com 50 ou mais marcadores na região;
- Para deleções (perda) foram analisadas CNVs com 25 ou mais marcadores na região;
- Para duplicações e deleções foram analisadas CNVs com tamanho mínimo de 100kb para pacientes que passaram nos dois parâmetros de qualidade ($MAPD \leq 0,35$ e $Waviness \leq 0,12$) e CNVs com tamanho mínimo de 300kb para pacientes que falharam em apenas um dos parâmetros de qualidade. Pacientes que falharam nos dois parâmetros de qualidade foram excluídos das análises;
- Para ROH foram analisadas regiões com 500 ou mais marcadores e com no mínimo 1500kb de tamanho.

Cada CNV e ROH foi analisada individualmente para determinação da qualidade e do nível de patogenicidade. Para determinação da frequência foram utilizados os seguintes conjuntos de dados de microarranjos como controles:

- Cerca de 2700 indivíduos da população geral disponibilizados pela *Affymetrix* (protocolo *Cytoscan HD*);
- Cerca de 800 controles internos provenientes da população alemã (protocolo *SNP Array 6.0*);

- Cerca de 270 indivíduos brasileiros (protocolo *SNP Array 6.0*) disponíveis no BIPMed⁷⁹ (www.bipmed.org);
- Bases de dados de frequência de CNVs como *DGV* e *aDGV*.

As CNVs raras (que foram encontradas em menos de 1% da população utilizando todos os conjuntos de controles) foram então selecionadas e classificadas quanto a patogenicidade utilizando também os critérios estabelecidos pela ACMG com a ajuda da plataforma *Franklin/Genoox*. Após seleção das variantes de interesse, essas foram relacionadas ao fenótipo observado a partir de seu potencial patogênico.

As CNVs foram também validadas pelo método de qPCR. Para a confirmação de rearranjos cromossômicos, utilizamos duas técnicas em colaboração com o Prof. Dr. Táris Antônio Paiva Vieira do Laboratório de Citogenética Humana e Citogenômica da FCM-UNICAMP: cariótipo com bandamento G e resolução de 400 a 550 bandas; e hibridização *in situ* por fluorescência (FISH, do inglês *fluorescence in situ hybridization*).

As ROH filtradas foram analisadas quanto à quantidade e distribuição no genoma e aquelas com tamanho maior que 10Mb foram selecionadas e analisadas quanto à presença de variantes patogênicas em genes já ligados às epilepsias no WES. Para isso foi utilizada a plataforma *Varstation* onde as variantes foram ordenadas pelo número do cromossomo e filtradas para priorizar variantes em homozigose nos genes relacionados às epilepsias.

Análise de Regressão Logística

As análises complexas utilizadas nesse trabalho (análises de regressão logística e SKAT-O) testam associações de variantes (SNPs) aos grupos de pacientes aqui descritos e foram realizadas com a ajuda da Dr^a. Thais Crippa de Oliveira e de nossa equipe de bioinformática chefiada pelo Prof. Dr. Benilton S. Carvalho. Foram utilizados aqui os pacientes que possuíam dados pareados, ou seja, pacientes que possuíam dados tanto de WES quanto de CMA.

Em dados de WES, o processamento foi feito utilizando o programa de automação *Espresso-Caller* (versão 1.2.0). Os arquivos brutos (FASTQ) foram mapeados contra o genoma de referência humano (*GRCh37/hg19*) utilizando o

programa *BWA* (versão 0.7.17) e a chamada de variantes das amostras individualmente foi feita utilizando o programa *GATK* (versão 4.1.0.0), com o uso do manual de melhores práticas do *Broad Institute*. A chamada de variantes foi realizada considerando toda a coorte, utilizando *JointGenotyping*. A calibração da genotipagem foi realizada utilizando os arquivos das amostras do banco de dados BIPMed, que são de alta qualidade⁷⁹. Ao final, foi gerado um único arquivo VCF para todos os pacientes que possui os dados de variantes presentes na coorte e a anotação dos SNPs foi gerada utilizando o software *ClinEff* (disponível em www.dnaminer.com/clineff.html).

Para dados de CMA, utilizamos os dados de genotipagem a partir dos arquivos CEL gerados para esse tipo de técnica. A primeira parte da conversão de arquivo CEL em arquivo CHP foi feita utilizando o *apt-probeset-genotype* com os parâmetros para *GenomeWideSNP_6*. Para a segunda conversão de CHP para VCF, foram utilizados os *plug-ins gtc2vcf* e *affy2vcf* para *bcftools*⁸⁰. Um filtro de percentil de 90% por indivíduo foi utilizado como parâmetro de qualidade dos SNPs tipados.

Os VCFs de ambas as técnicas foram então unificados por indivíduo, de forma que todos os indivíduos nos grupos caso e controle possuíssem dados de ambas as técnicas (dados pareados). *Pipelines* próprios foram utilizados para combinar os dados e, no caso específico de SNPs em posições idênticas em ambas as técnicas foi realizada a imputação dos dados para remoção de lacunas em branco/faltantes (*missing calls*). O grupo de SNPs caso-controle unificado foi então submetido a dois fluxos distintos de filtragem de SNPs e de amostras utilizando o programa *PLINK* (versão 2.0)⁸¹: um para variantes comuns e outro para variantes raras. O fluxo de filtragem para variantes comuns se sucedeu da seguinte forma:

- (A) Frequência do alelo menos (MAF, do inglês *minor allele frequency*) maior que 0,01, parâmetro PASS e qualidade maior que 30;
- (B) Teste *Fisher* em variantes que possuem *missing calls* de caso/controlado;
- (C) Exclusão de amostras com mais de 10% de *missing calls* e SNPs com mais de 5% de *missing calls*;
- (D) Teste de heterozigosidade;
- (E) Equilíbrio de *Hardy-Weinberg*, com valores de $1e-6$ para controles e $1e-10$ para casos;

- (F) Teste de desequilíbrio de ligação (LD, do inglês *linkage disequilibrium*) com *cut-off* de $R^2 = 0,2$, com janelas de 50kb e passos de 5kb;
- (G) IBD (identidade por descendência) $> 0,125$;
- (H) Inspeção visual com uso de análise de componentes principais (PCA, do inglês *Principal Component Analysis*).

Para o fluxo de variantes raras, a duas únicas mudanças ocorreram no item (A), em que não há definição de MAF e um último item é adicionado com a separação entre os bancos de dados de variantes raras (MAF = 0,01) e ultrarraras (MAF = 0,0005). A análise de regressão logística, que foi realizada utilizando também o software *PLINK* (versão 2.0)⁸¹, verifica se há relação entre a variável dependente e as variáveis independentes de um conjunto de dados, ou seja, se há associação entre as variantes testadas e o conjunto de pacientes. Para essa análise, separamos os pacientes em três grupos pós filtragem: conjunto total de pacientes + controles; conjunto dos pacientes com DG + controles; e conjunto dos pacientes sem DG + controles. Definindo que em casos com DG foi possível encontrar uma causa genética de maior efeito, e que em casos sem DG não foi possível encontrar uma causa genética para a doença.

Análise de SKAT-O

A análise de SKAT-O é uma análise baseada em genes e testa a associação entre o grupo de variantes em um dado gene utilizando grupos de casos e controles, assumindo que todas essas variantes possuem efeito cumulativo no fenótipo dos pacientes⁸².

Essa análise foi realizada utilizando o banco de dados de variantes raras e ultrarraras criado na seção anterior, localizadas apenas nos genes descritos no painel EG (Anexo III), não sendo utilizados os genes localizados no cromossomo X. Foram realizadas três análises utilizando o pacote *SKAT*⁸² no programa *R* e os resultados de p-valores por gene ajustados $\leq 0,05$ foram avaliados. Assim como os dados de variantes comuns, foram utilizados os mesmo três conjuntos de indivíduos também utilizando somente os dados pareados.

RESULTADOS E DISCUSSÃO

Sequenciamento Completo de Exoma

Após análise de qualificação e quantificação de DNA, todas as 239 amostras foram enviadas para o sequenciamento e obtivemos resultado de 234 delas. As cinco amostras restantes, apesar de apresentarem boa qualidade na primeira triagem, não apresentaram boa qualidade para a preparação da biblioteca e foram excluídas em triagens posteriores. Os resultados apresentados na Figura 5 e nas Tabelas 1 e 3 foram obtidos a partir da análise utilizando plataforma *NGS Variant Analyzer*. Já os resultados das variantes selecionadas (Tabela 2 e Figuras 6 e 7) e de taxa de DG foram descritos utilizando as duas análises combinadas (*NGS Variant Analyzer* e *Varstation*).

Após processamento dos WES, foi calculada a ancestralidade através de análise de componentes principais (PCA; Figura 5). É possível observar um padrão de miscigenação entre a população europeia e ambas populações africana e nativa, com possível presença de alguns alelos do leste asiático, como já observado em outros estudos⁸³. Essa proximidade da população analisada com a população europeia justifica também o uso de controles europeus nas análises.

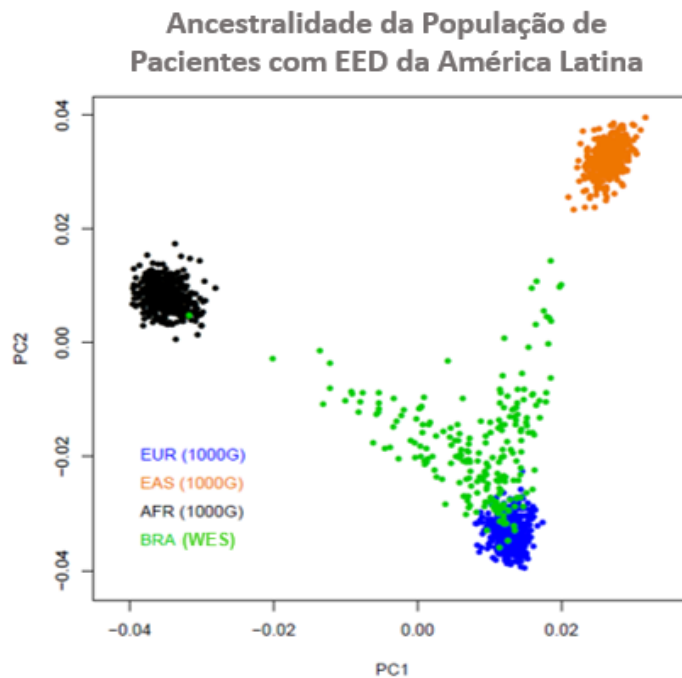


Figura 5: Ancestralidade dos pacientes de EED da América Latina (BRA) calculada a partir da análise de componentes principais utilizando os dados de WES. EUR – população europeia; EAS – população do leste asiático; AFR – população africana

Os resultados de WES foram separados em dois grupos de acordo com a profundidade dos sequenciamentos e seus parâmetros podem ser observados na Tabela 1. O primeiro grupo com 118 pacientes foi sequenciado desejando-se atingir uma profundidade média de 30x e o segundo grupo, com 116 pacientes foi sequenciado visando atingir uma profundidade média de 100x. Podemos observar uma grande diferença entre as qualidades dos sequenciamentos, principalmente na porcentagem do genoma que foi coberta com no mínimo 20x de profundidade (% cobertura 20x), que é quando a chamada de variantes é mais confiável. Observamos também um número maior de variantes após aplicação dos filtros no grupo de 30x, devido à maior ocorrência de variantes não reais e artefatos. Foi também devido a esse fato que validamos todas as variantes selecionadas utilizando outras técnicas.

Após análise de todos os pacientes, utilizando os dois *pipelines* e as duas plataformas de análise (*NGS Variant Analyzer* e *Varstation*) selecionamos 100 variantes de interesse em 96 pacientes em 47 genes diferentes (Tabela 2). Sendo que, dois pacientes não relacionados apresentaram a mesma variante em *GNAO1* (pacientes 175 e 186), um par de gêmeas apresentou a mesma variante em *SCN1A* (pacientes 52 e 53) e outro par de gêmeas apresentou a mesma variante em *NRXN2* (pacientes 36 e 37).

Tabela 1: Parâmetros de qualidade e número de variantes dos WES. Valores correspondem às médias.

| Parâmetro | Grupo 1 – 30X | Grupo 2 – 100X |
|--------------------------------|---------------|----------------|
| Número de pacientes | 118 | 116 |
| Profundidade | 29 | 73 |
| % cobertura 10X | 85,9 | 97 |
| % cobertura 20X | 59,9 | 90,5 |
| Nº total de variantes | 186738 | 271546,4 |
| Nº de variantes após filtro | 1044,27 | 789,7 |
| Nº de variantes após painel PG | 52,7 | 37,9 |
| Nº de variantes após painel EG | 18,4 | 14,7 |

Foi possível realizar o diagnóstico genético em 79 pacientes, até o momento avaliando somente variantes de sequência (SNV e indel). Algumas variantes provavelmente patogênicas ainda aguardam validação pois apresentam baixa profundidade no sequenciamento. As variantes de significado incerto (VUS) selecionadas também estão aguardando validação e análise de segregação, pois suas classificações serão alteradas se forem variantes não herdadas de nenhum dos progenitores (*de novo*).

A frequência dos genes e as distribuições quanto ao tipo de variante, classificação da ACMG e padrão de herança nesses pacientes com diagnóstico podem ser observadas na Figura 6. Essa a primeira vez que um estudo desse tipo é realizado em uma grande coorte de população da América Latina. Em relação às distribuições de tipo de variante e padrão de herança, nossa população não apresentou padrões diferentes do esperado para pacientes com epilepsia e doenças do neurodesenvolvimento⁸⁴. Além disso, observamos um grande número de variantes do tipo SNV, sendo essas as mais frequentes nas epilepsias genéticas⁸⁵.

Observamos uma maior frequência do gene *SCN1A*, sendo responsável por 35,5% dos casos com diagnóstico. Esse alto número pode ser devido ao grande número de pacientes com síndrome de Dravet, que somam 72,7% de todos os pacientes com variantes nesse gene. O que já era esperado, já que o gene *SCN1A* é responsável por cerca de 80% dos casos de Dravet, como já explicado anteriormente^{29,30}. Porém, excluindo os pacientes com síndrome de Dravet de nossa casuística, o gene *SCN1A* continua sendo o mais frequente, somando 13,8% dos casos com DG, e aparece em pacientes pertencentes ao grupo das EEDs indeterminadas e síndrome de Doose. As distribuições dos diagnósticos moleculares envolvendo o gene *SCN1A* no grupo total de pacientes e nos grupos de pacientes com e sem diagnóstico clínico de síndrome de Dravet podem ser observadas na Figura 7.

O gene *SCN1A* é o mais frequente para as EEDs e para outros tipos de epilepsias, principalmente nas de início precoce, também em outras populações⁸⁵. Para os outros genes, observamos frequências ligeiramente diferentes do apresentado por coortes de outros países, como Austrália⁸⁶, Canadá⁸⁷, Estados Unidos^{84,88}, Itália⁴⁵, Coreia do Sul⁸⁹ e Índia⁹⁰. Em um estudo prospectivo de pacientes com epilepsias da infância da população escocesa, o gene mais frequente foi o

PRRT2, o que indica que estudos a níveis populacionais são muito importantes na avaliação de uma coorte^{8,63}. Entre os outros genes que aparecem com uma variante causal cada estão: *AHDC1*, *ALG13*, *ANKRD11*, *ASXL3*, *ATP1A3*, *BSCL2*, *CDK19*, *CLN5*, *EEF1A2*, *GAMT*, *GNAO1*, *GRIN2B*, *IQSEC2*, *KCNQ2*, *MEF2C*, *NEXMIF*, *NPRL2*, *NRAS*, *NSUN2*, *PROK2*, *PRRT2*, *RARS2*, *SLC13A5*, *SLC6A1*, *SLC9A6*, *SMC1A* e *TBCD*. As variantes causais presente nos 79 pacientes estão localizadas em 36 genes, o que ilustra bem a heterogeneidade genética característica das EEDs.

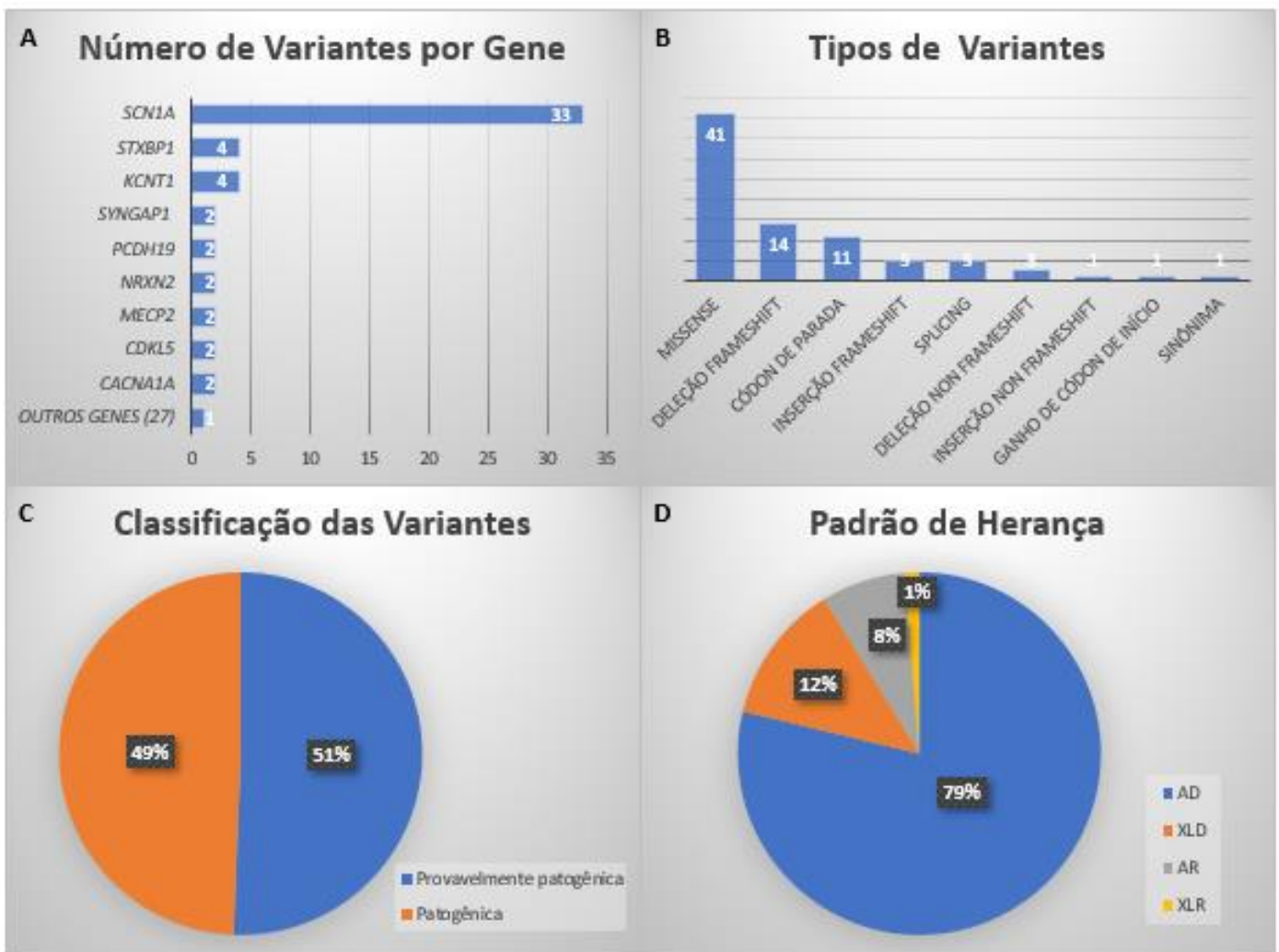


Figura 6: (A) Distribuição do número de variantes em relação ao gene em que estão localizadas. (B) Tipo de mudança na sequência causada pela variante. (C) Classificação das variantes segundo a ACMG. (D) Padrão de herança da doença relacionada ao gene em que a variante se encontra. AD – Autossômica Dominante; XLD – Ligada ao X Dominante; AR – Autossômica Recessiva; XLR – Ligada ao X Recessiva.

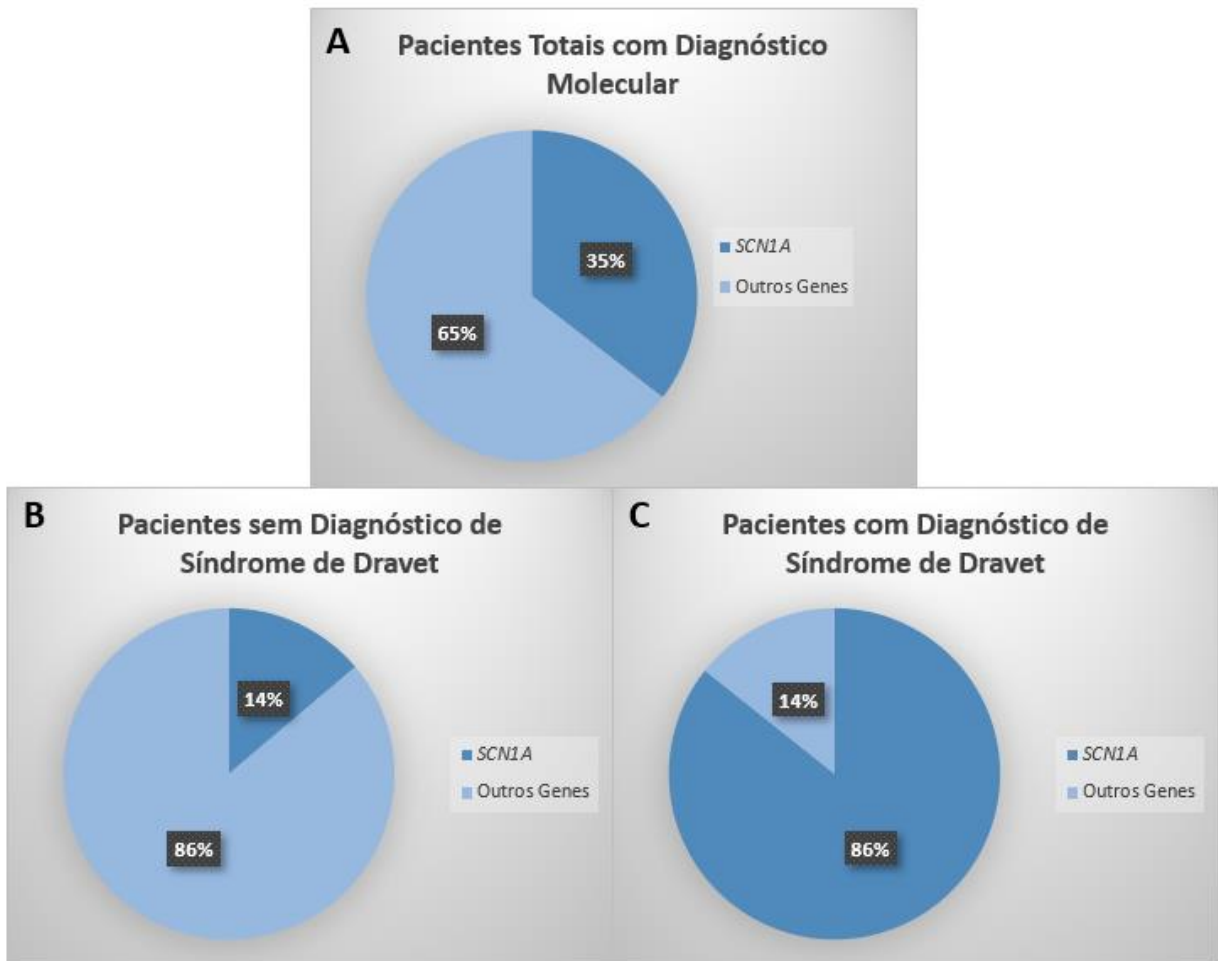


Figura 7: Distribuição dos pacientes com variantes causais no gene *SCN1A* e em outros genes mostrando que o gene *SCN1A* é o mais frequente em todos os grupos: (A) Grupo de pacientes totais onde foi possível realizar o diagnóstico genético. (B) Grupo de pacientes onde foi possível realizar o diagnóstico genético excluindo os pacientes com diagnóstico clínico de síndrome de Dravet. (C) Grupo de pacientes com diagnóstico clínico de síndrome de Dravet em que foi possível realizar o diagnóstico genético.

Para os 33 pacientes com variantes causais no gene *SCN1A* foi possível realizar a análise de segregação com validação nos progenitores em 20 deles e em três casos essas variantes foram herdadas de um dos pais. Já é conhecido que as EEDs são doenças geneticamente heterogêneas⁹¹ e que variantes no gene *SCN1A* podem causar um amplo espectro de condições, desde de crises febris benignas até casos de morte súbita^{66,92}. No caso da síndrome de Dravet, em cerca de 10% dos pacientes com mutações *de novo* no gene *SCN1A*, um dos progenitores é mosaico, e isso deve ser levado em consideração para o aconselhamento genético⁶⁶. Além disso, a síndrome de Dravet ainda pode ocorrer em um membro de uma família com GEFS+ e variantes patogênicas em *SCN1A* podem apresentar penetrância incompleta e podem ocorrer em outras síndromes epiléticas⁶⁶.

Tabela 2: Variantes selecionadas após análise de WES. ID - número de identificação do paciente; DG - diagnóstico genético; PH - padrão de herança da doença relacionada ao gene; VUS – variante de significado incerto.

| ID | DG | Gene | Variante | Tipo da variante | Consequência | Classificação | PH | Fenótipo |
|-----|---------|---------|---------------------------------|---------------------------------|---------------------------|--------------------------|-----|----------------------------|
| 178 | OK | AHDC1 | c.92_93insGG | Inserção frameshift | p.Gly31fs | Provavelmente patogênica | AD | Síndrome de Doose |
| 213 | OK | ALG13 | c.320A>G | Missense | p.Asn107Ser | Patogênica | XLD | Síndrome de Dravet |
| 229 | OK | ANKRD11 | c.7814T>G | Missense | p.Leu2605Arg | Provavelmente patogênica | AD | Síndrome de Dravet |
| 12 | OK | ASXL3 | c.4215_4216del | Deleção frameshift | p.Leu1407fs | Patogênica | AD | Síndrome de Lennox-Gastaut |
| 240 | OK | ATP1A3 | c.2443G>A | Missense | p.Glu815Lys | Patogênica | AD | Síndrome de Dravet |
| 62 | Validar | ATP6V1A | c.1157A>T | Missense | p.Tyr386Phe | VUS | AD | Síndrome de Doose |
| 100 | OK | CACNA1A | c.6525_6537dupCAC ATCGGTGTAG | Inserção frameshift | p.Asp2180Leufs*94 | Provavelmente patogênica | AD | Síndrome de Dravet |
| 257 | OK | CACNA1A | c.2986G>T | Códon de parada | p.Glu996* | Provavelmente patogênica | AD | Síndrome de Dravet |
| 43 | OK | CDK19 | c.1066C>T | Códon de parada | p.Arg356* | Patogênica | AD | Síndrome de Dravet |
| 160 | OK | CDKL5 | c.419A>G | Missense | p.Asn140Ser | Patogênica | XLD | Síndrome de Dravet |
| 201 | OK | CDKL5 | c.118G>A | Missense | p.Ala40Thr | Provavelmente patogênica | XLD | Síndrome de Dravet |
| 222 | OK | CLN5 | c.659delA | Deleção frameshift | p.Lys220fs | Provavelmente patogênica | AR | Síndrome de Dravet |
| 167 | Validar | CLTC | c.3603_3604insTTTTT | Inserção frameshift | p.Gly1202fs | Provavelmente patogênica | AD | Síndrome de Dravet |
| 81 | Validar | DEPDC5 | c.1174C>T | Missense | p.Arg392Cys | VUS | AD | Síndrome de Dravet |
| 228 | OK | EEF1A2 | c.370G>A | Missense | p.Glu124Lys | Provavelmente patogênica | AD | EED indeterminada |
| 82 | Validar | GABBR2 | c.2743_2744insCCAG AGGGTCA | Inserção frameshift | p.Val915fs | VUS | AD | EED indeterminada |
| 102 | OK | GAMT | c.506G>A c.491dupG | Missense Inserção frameshift | p.Cys169Tyr p.Val165fs | Patogênica | AR | Síndrome de Dravet |
| 175 | Validar | GNAO1 | c.448A>C | Missense | p.Asn150His | Provavelmente patogênica | AD | Síndrome de Dravet |
| 186 | Validar | GNAO1 | c.448A>C | Missense | p.Asn150His | Provavelmente patogênica | AD | Síndrome de Dravet |
| 219 | OK | GNAO1 | c.736G>A | Missense | p.Glu246Lys | Patogênica | AD | Síndrome de Dravet |
| 225 | OK | GRIN2B | c.1906G>A | Missense | p.Ala636Thr | Provavelmente patogênica | AD | Síndrome de Doose |
| 133 | Validar | HIVEP2 | c.7337delA | Deleção frameshift | p.His2446fs | VUS | AD | Síndrome de Dravet |
| 194 | OK | IQSEC2 | c.23delC | Deleção frameshift | p.Pro8fs | Provavelmente patogênica | XLD | Síndrome de Dravet |
| 207 | Validar | KCNB1 | c.574G>C | Missense | p.Ala192Pro | VUS | AD | Síndrome de Dravet |
| 177 | Validar | KCNQ2 | c.2518A>G | Missense | p.Arg840Gly | VUS | AD | Síndrome de Dravet |
| 251 | OK | KCNQ2 | c.790T>C | Missense | p.Tyr264His | Patogênica | AD | EED indeterminada |
| 2 | Validar | KCNQ3 | c.105_107dupGGC | Inserção frameshift | p.Ala36dup | VUS | AD | Síndrome de Dravet |
| 78 | OK | KCNT1 | c.1066C>T | Missense | p.Arg356Trp | Provavelmente patogênica | AD | Síndrome de Dravet |
| 149 | OK | KCNT1 | c.2944-2A>C | Splicing | - | Provavelmente patogênica | AD | EED indeterminada |
| 161 | OK | KCNT1 | c.2800G>A | Missense | p.Ala934Thr | Provavelmente patogênica | AD | EED indeterminada |
| 212 | OK | KCNT1 | c.1193G>A | Missense | p.Arg398Gln | Provavelmente patogênica | AD | EED indeterminada |
| 72 | OK | MECP2 | c.916C>T | Missense | p.Arg306Cys | Patogênica | XLD | EED indeterminada |
| 221 | OK | MECP2 | c.473C>T | Missense | p.Thr158Met | Patogênica | XLD | EED indeterminada |

| | | | | | | | | |
|-----|---------|--------|------------------------------------|-----------------------------|----------------------------------|--|-----|----------------------------|
| 185 | OK | MEF2C | c.565C>T | Códon de parada | p.Arg189* | Patogênica | AD | EED indeterminada |
| 87 | Validar | MTOR | c.2051A>G | Missense | p.Asp684Gly | VUS | AD | EED indeterminada |
| 198 | OK | NEXMIF | c.1882C>T | Códon de parada | p.Arg628* | Patogênica | XLD | EED indeterminada |
| 99 | OK | NPRL2 | c.883C>T | Códon de parada | p.Arg295* + alteração estrutural | Patogênica | AD | EED indeterminada |
| 261 | OK | NRAS | c.179G>A | Missense | p.Gly60Glu | Patogênica | AD | EED indeterminada |
| 36 | OK | NRXN2 | c.808delG | Deleção frameshift | p.Ala270fs | Provavelmente patogênica | AD | EED indeterminada |
| 37 | OK | NRXN2 | c.808delG | Deleção frameshift | p.Ala270fs | Provavelmente patogênica | AD | EED indeterminada |
| 16 | OK | NSUN2 | c.2035G>A | Missense | p.Gly679Arg | Provavelmente patogênica | AR | Síndrome de Dravet |
| 95 | OK | PCDH19 | c.1314delG | Deleção frameshift | p.Phe439fs | Provavelmente patogênica | XLD | EED indeterminada |
| 255 | OK | PCDH19 | c.1019A>G | Missense | p.Asn340Ser | Patogênica | XLD | EED indeterminada |
| 60 | OK | PROK2 | c.163delA | Deleção frameshift | p.Ile55fs | Patogênica | AD | Síndrome de Dravet |
| 223 | OK | PRRT2 | c.641delC | Deleção frameshift | p.Ala214fs | Patogênica | AD | Síndrome de Lennox-Gastaut |
| 202 | OK | RARS2 | c.754T>A c.1A>G | Missense Códon de início | p.Tyr252Asn p.Met1? | Provavelmente patogênica Patogênica | AR | EED indeterminada |
| 15 | OK | SCN1A | c.971A>C | Missense | p.His324Pro | Provavelmente patogênica | AD | EED indeterminada |
| 17 | OK | SCN1A | c.602+1G>A | Splicing | - | Patogênica | AD | Síndrome de Dravet |
| 21 | OK | SCN1A | c.4753C>T | Missense | p.Arg1596Cys | Patogênica | AD | Síndrome de Dravet |
| 22 | OK | SCN1A | c.5146G>T | Missense | p.Asp1727Tyr | Patogênica | AD | EED indeterminada |
| 24 | OK | SCN1A | c.829T>C | Missense | p.Cys277Arg | Provavelmente patogênica | AD | EED indeterminada |
| 28 | OK | SCN1A | c.1170+3G>T | Splicing | - | Provavelmente patogênica | AD | EED indeterminada |
| 29 | OK | SCN1A | c.383+1A>G | Sinônima | p.Ser128Ser | Provavelmente patogênica | AD | EED indeterminada |
| 31 | OK | SCN1A | c.5701C>T | Códon de parada | p.Arg1912* | Patogênica | AD | Síndrome de Dravet |
| 33 | OK | SCN1A | c.5456_5458delAGT | Deleção non Frameshift | p.Gln1819_Phe1820 delinsLeu | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 34 | OK | SCN1A | c.1243delA | Deleção frameshift | p.Ile415fs | Patogênica | AD | EED indeterminada |
| 35 | OK | SCN1A | c.2327T>G | Missense | p.Met787Arg | Provavelmente patogênica | AD | EED indeterminada |
| 38 | OK | SCN1A | c.4060G>T | Missense | p.Gly1354Cys | Patogênica | AD | Síndrome de West |
| 39 | OK | SCN1A | c.2254delC | Deleção frameshift | p.Leu752fs | Provavelmente patogênica | AD | Síndrome de West |
| 40 | OK | SCN1A | c.5144G>A | Códon de parada | p.Trp1715* | Patogênica | AD | EED indeterminada |
| 41 | OK | SCN1A | c.4251+1G>A | Splicing | - | Patogênica | AD | Síndrome de Lennox-Gastaut |
| 46 | OK | SCN1A | c.3656T>C | Missense | p.Leu1219Pro | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 49 | OK | SCN1A | c.4940C>G | Missense | p.Thr1647Arg | Provavelmente patogênica | AD | Síndrome de Doose |
| 52 | OK | SCN1A | c.296_313delTCTTCC GGTTCAGTGCCA | Deleção non Frameshift | p.Ile99_Ala104del | Provavelmente patogênica | AD | EED indeterminada |
| 53 | OK | SCN1A | c.296_313delTCTTCC GGTTCAGTGCCA | Deleção non Frameshift | p.Ile99_Ala104del | Provavelmente patogênica | AD | EED indeterminada |
| 54 | OK | SCN1A | c.1094T>G | Missense | p.Phe365Cys | Provavelmente patogênica | AD | EED indeterminada |
| 56 | OK | SCN1A | c.4402A>T | Missense | p.Ile1468Phe | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 57 | OK | SCN1A | c.2465_2467dupATA | Inserção non frameshift | p.Asn822dup | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 59 | OK | SCN1A | c.1692_1693dupTT | Inserção frameshift | p.Ser565fs | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |

| | | | | | | | | |
|-----|---------|-----------------|--------------------------|--|----------------------------|--------------------------|----------|--|
| 71 | OK | SCN1A | c.5732T>C | Missense | p.Ile1911Thr | Patogênica | AD | Síndrome de Lennox-Gastaut |
| 90 | OK | SCN1A | c.1400C>G | Códon de parada | p.Ser467* | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 108 | OK | SCN1A | c.1007G>A | Missense | p.Cys336Tyr | Patogênica | AD | Síndrome de Doose |
| 110 | OK | SCN1A | c.1177C>T | Missense | p.Arg393Cys | Patogênica | AD | Síndrome de Doose |
| 115 | OK | SCN1A | c.1620delG | Deleção frameshift | p.Asn541fs | Provavelmente patogênica | AD | Síndrome de Doose vs. Síndrome de Lennox-Gastaut |
| 130 | OK | SCN1A | c.5233T>C | Missense | p.Cys1745Arg | Provavelmente patogênica | AD | EED indeterminada |
| 187 | OK | SCN1A | c.2803C>T | Missense | p.Arg935Cys | Patogênica | AD | EED indeterminada |
| 244 | Validar | SCN1A | c.2711T>C | Missense | p.Phe904Ser | Provavelmente patogênica | AD | EED indeterminada |
| 247 | OK | SCN1A | c.1033T>C | Missense | p.Cys345Arg | Patogênica | AD | EED indeterminada |
| 259 | Validar | SCN1A | c.3628G>A | Missense | p.Glu1210Lys | VUS | AD | EED indeterminada |
| 264 | OK | SCN1A | c.5477C>T | Missense | p.Pro1826Leu | Provavelmente patogênica | AD | Síndrome de Lennox-Gastaut |
| 307 | OK | SCN1A | c.2926T>C | Missense | p.Phe976Leu | Provavelmente patogênica | AD | Síndrome de Doose |
| 112 | Validar | SCN2A | c.3013A>G | Missense | p.Ile1005Val | VUS | AD | Síndrome de Doose |
| 269 | Validar | SCN2A | c.23C>A | Missense | p.Pro8Gln | VUS | AD | EED indeterminada |
| 137 | OK | SLC13A5 | c.655G>A | Missense | p.Gly219Arg | Patogênica | AR | EED indeterminada |
| 205 | OK | SLC6A1 | c.813C>A | Códon de parada | p.Tyr271* | Patogênica | AD | EED indeterminada |
| 199 | OK | SLC9A6 | c.220delC | Deleção frameshift | p.Leu74fs | Provavelmente patogênica | XLR | EED indeterminada |
| 105 | OK | SMC1A | c.1771delG | Deleção frameshift | p.Ala591fs | Patogênica | XLD | EED indeterminada |
| 92 | OK | STXBP1 | c.704G>A | Missense | p.Arg235Gln | Patogênica | AD | EED indeterminada |
| 243 | OK | STXBP1 | c.1652G>A | Missense | p.Arg551His | Patogênica | AD | EED indeterminada |
| 262 | OK | STXBP1 | c.1074C>A | Códon de parada | p.Tyr358* | Patogênica | AD | EED indeterminada |
| 113 | OK | STXBP1 BSCL2 | c.685dupC c.412C>T | Inserção frameshift Códon de parada | p.Gln229fs p.Arg138* | Patogênica Patogênica | AD AD | Síndrome de West |
| 75 | OK | SYNGAP1 | c.388-2A>C | Splicing | - | Provavelmente patogênica | AD | Síndrome de West |
| 91 | OK | SYNGAP1 | c.2387delC | Deleção frameshift | p.Pro796fs | Patogênica | AD | EED indeterminada |
| 84 | OK | TBCD | c.1340C>T c.1424C>T | Missense | p.Ala447Val p.Ala475Val | Provavelmente Patogênica | AR | Síndrome de West |
| 267 | Validar | TBL1XR1 | c.821G>A | Códon de parada | p.Trp274* | Provavelmente patogênica | AD | Síndrome de West |
| 58 | Validar | TRIO | c.7355_7360delGGG CCG | Deleção non Frameshift | p.Gly2452_Ala2453 del | VUS | AD | Síndrome de Dravet |

As análises de CNVs a partir dos exomas revelou uma média de 66 CNVs por paciente para o grupo sequenciado com profundidade de 30x e de 120,8 CNVs por paciente para o grupo com 100x. Após filtragem, as médias de CNVs por paciente foram de 6,7 e 17,2 respectivamente. Após aplicação do painel de genes de EEDs as médias de CNVs por paciente foram de 0,25 e 0,55 respectivamente. Após análise de todas essas variantes obtidas após aplicação do painel, 12 delas foram selecionadas

em 11 pacientes, sendo 11 variantes patogênicas e uma provavelmente patogênica (Tabela 3). A maioria das CNVs encontradas no WES também foram observadas na análise de CMA, com exceção das CNVs dos pacientes 255 e 256 que só apareceram no WES devido à má qualidade dos microarranjos. Outras três CNVs (pacientes 132, 239 e 310) foram selecionadas a partir do WES, mas também apareceram no CMA após remoção de filtros.

Tabela 3: CNVs causais selecionadas a partir da análise de WES. Genes OMIM corresponde ao número de genes envolvidos na CNV já associados a doenças presentes no banco de dados do OMIM. ID: número de identificação do paciente.

| ID | Variante | Genes OMIM | Tamanho | Consequência | Classificação | Fenótipo |
|-----|--|------------|------------------|--|--------------------------|----------------------------|
| 63 | ngs[hg19] 15q11.2q13.1(23609191_28525369)x1 | 14 | 4,9Mb | Síndrome de Angelman | Patogênica | EED indeterminada |
| 67 | ngs[hg19] 15q11.2q13.2(23325091_30702538)x4 ngs[hg19] 15q13.2q13.3(31093551_31862472)x3 | 25 5 | 7,4Mb 768,9Kb | Tetrassomia parcial do cromossomo 15 | Patogênica | EED indeterminada |
| 74 | ngs[hg19] 16p11.2(29580502_30215205)x1 | 33 | 634,7Kb | Síndrome de deleção 16p11.2 | Patogênica | EED indeterminada |
| 132 | ngs[hg19] 1q44(244848874_245027664)x1 | 3 | 178,8Kb | Deleção do gene <i>HNRNPU</i> | Patogênica | Síndrome de West |
| 134 | ngs[hg19] 1p36.33p36.32(803773_3816962)x1 | 75 | 3Mb | Síndrome de deleção 1p36 | Patogênica | EED indeterminada |
| 144 | ngs[hg19] 16p13.3(103402_3900998)x3 | 157 | 3,8Mb | Síndrome de duplicação 16p13.3 | Patogênica | Síndrome de Lennox-Gastaut |
| 239 | ngs[hg19] 16p11.2(29547304_30215205)x3 | 33 | 667,9Kb | Síndrome de microduplicação 16p11.2 | Patogênica | Síndrome de West |
| 254 | ngs[hg19] 15q11.2q13.1(23608748_28475084)x3 | 14 | 4,9Mb | Síndrome de duplicação 15q11.2 | Patogênica | EED indeterminada |
| 255 | ngs[hg19]15q11.2(25585161_25599876)x1 | 1 | 14,7Kb | Deleção dos exons 12, 13 e 14 do gene <i>UBE3A</i> | Patogênica | EED indeterminada |
| 256 | ngs[hg19]2q22.3(145182225_145187633)x1 | 1 | 5,4Kb | Deleção dos exons 3 e 4 do gene <i>ZEB2</i> | Provavelmente patogênica | EED indeterminada |
| 310 | ngs[hg19] 2q24.3(166773732_166915261)x1 | 2 | 169,2Kb | Deleção de parte do gene <i>SCN1A</i> | Patogênica | Síndrome de Dravet |

Considerando tanto as variantes de sequência quanto as CNVs, foi possível obter um DG em 89 pacientes utilizando somente dados de WES. Como já dito anteriormente, os WES foram sequenciados utilizando duas profundidades médias diferentes, porém, não houve diferença quanto a taxa de DG nos dois grupos. No grupo 1, que foi sequenciado visando obter uma profundidade média de 30x foi possível obter o DG em 44 de todos os 118 pacientes sequenciados (37,3%). Já no grupo 2, que foi sequenciado visando obter profundidade média de 100x, foi possível obter o DG em 45 de todos os 116 pacientes sequenciados (38,8%). Considerando todos os 89 pacientes com DG, 49,4% pertencem ao grupo 1 (30x) e 50,6% pertencem ao grupo 2 (100x).

Análise de Microarranjos Cromossômicos

Assim como no WES, todas as amostras foram submetidas à técnica de CMA, porém, três das amostras não apresentaram boa qualidade em triagens posteriores e foram excluídas já no início do procedimento. Sendo assim, obtivemos resultados de microarranjos de 236 pacientes. Desses 236 pacientes, 16 foram excluídos após a análise de controle de qualidade. Analisamos então, 220 pacientes quanto a presença de CNVs causais, sendo que 87 passaram em ambos os parâmetros de qualidade ($MAPD \leq 0,35$ e $Waviness \leq 0,12$) e foram analisados com o filtro de tamanho de no mínimo 100kb e 133 passaram em apenas um dos controles de qualidade e foram analisados utilizando filtro de tamanho de no mínimo 300kb.

Na análise utilizando o programa *ChaS* e os filtros descritos nos métodos, observamos 2694 CNVs em todos os 220 pacientes (média de 12,2 CNVs por paciente). Selecionamos 115 CNVs raras em 76 pacientes, sendo que: 25 CNVs foram classificadas como provavelmente benignas por não possuírem genes em suas extensões; 76 classificadas como VUS pois possuem genes com função desconhecida, ainda não relacionados a doenças ou não sensíveis a dosagem; duas classificadas como provavelmente patogênicas, mas não causais, pois possuem genes sensíveis a dosagem mas que são relacionados somente com doenças recessivas; e 15 como patogênicas, sendo que duas são relacionadas a síndrome de Down e 13 possuem relação direta com os fenótipos. Essas 13 CNVs compreendem síndromes de deleção ou duplicação já conhecidas ou possuem genes relacionados a EEDs em sua extensão (Tabela 4).

Após finalizar as duas análises, todas as CNVs encontradas no WES e CMA foram comparadas e, como já dito anteriormente, três CNVs encontradas no WES se confirmaram também no CMA (pacientes 132, 239 e 310). Essas CNVs foram excluídas pelos filtros de tamanho aplicados, sendo assim, foram adicionadas duas análises complementares: análise das CNVs de todos os pacientes sem filtro de tamanho levando em conta a presença de genes relacionados às EEDs; e análise dos 16 pacientes anteriormente excluídos utilizando filtro de tamanho de 300Kb. Porém, em ambos os casos, nenhuma outra CNV além das três detectadas pelo exoma foi selecionada. Isso não quer dizer que não haja variantes causais nesses pacientes, pois não podemos confiar na qualidade desses microarranjos.

As validações das CNVs ainda estão em andamento, e se fazem necessárias, principalmente para as CNVs menores que 1Mb. Os dois casos de CNVs mais interessantes que encontramos e que envolvem rearranjos cromossômicos já foram validados nos pacientes e nos pais e estão descritos a seguir.

Tabela 4: CNVs patogênicas causais selecionadas a partir da análise de CMA. Genes OMIM corresponde ao número de genes envolvidos na CNV já associados a doenças presentes no banco de dados do OMIM. ID: número de identificação do paciente.

| ID | Variante | Genes OMIM | Tamanho | Consequência | Fenótipo |
|-----|---|----------------|-------------------------|---|----------------------------|
| 63 | arr[hg19] 15q11.2q13.1(23282799_28445135)x1 | 19 | 5,2Mb | Síndrome de Angelman | EED indeterminada |
| 67 | arr[hg19] 15q11.1q13.2(20016316_30764615)x4 arr[hg19] 15q13.2q13.3(30775982_32876972)x3 | 36 | 12,9Mb | Tetrassomia parcial do cromossomo 15 | EED indeterminada |
| 74 | arr[hg19] 16p11.2(29580611_30226931)x1 | 23 | 646,3Kb | Síndrome de deleção 16p11.2 | EED indeterminada |
| 104 | arr[hg19] 13q12.3q13.1(29626643_33600755)x1 | 13 | 4Mb | Síndrome de deleção 13q12.3 | EED indeterminada |
| 134 | arr[hg19] 1p36.33p36.32(564621_4105141)x1 | 53 | 3,5Mb | Síndrome de deleção 1p36 | EED indeterminada |
| 144 | arr[hg19] 16p13.3(85815_4021612)x3 arr[hg19] 22q13.33(50275935_51234443)x1 | 121 29 | 3,9Mb 958,5Kb | Translocação desbalanceada entre os cromossomos 16 e 22 | Síndrome de Lennox-Gastaut |
| 176 | arr[hg19] 16p12.2(21807855_22558797)x1 | 3 | 750,9Kb | Síndrome de microdeleção 16p12.2 | Síndrome de Lennox-Gastaut |
| 195 | arr[hg19]13q32.1q33.2(97908422_105183839)x3 | 22 | 7,3Mb | Associada a doenças que causam atraso do desenvolvimento e epilepsia | Síndrome de Lennox-Gastaut |
| 250 | arr[hg19] 15q11.1q11.2(20577500_23487534)x1 | 6 | 2,9Mb | Região da síndrome de microdeleção 15q11.2 | EED indeterminada |
| 254 | arr[hg19] 15q11.2q13.2(22751668_30668572)x3 arr[hg19] 16p13.12p13.11(14777381_16290495)x3 arr[hg19] 22q11.21(18876416_21567219)x3 | 27 11 40 | 7,9Mb 1,5Mb 2,7Mb | Síndrome de duplicação 15q11.2 Síndrome de microduplicação 16p13 Síndrome de duplicação 22q11 | EED indeterminada |

No caso do paciente 67, havia suspeita de um cromossomo marcador supranumerário devido a presença de duas grandes CNVs na região proximal do cromossomo 15. Trata-se de uma região de *imprinting* instável conhecida pela recorrência de rearranjos devido à presença de DNA repetitivo⁹³. Esse tipo de cromossomo, conhecido como invdup(15) ou idic(15), já é amplamente conhecido na literatura. Foi descrito pela primeira vez em 1977 por Van Dyke *et al.*, e está relacionado a diferentes fenótipos que incluem hipotonia, ADNPM, deficiência intelectual, epilepsia e comportamento autista^{94,95}. Nosso paciente apresenta epilepsia refratária com CTCG, Miastenia Gravis Ocular, micronefrolitíase bilateral, ADNPM com piora após início das crises, deficiência intelectual e transtorno do espectro autista, entre outras características clínicas. O cariótipo do paciente apresentou 47 cromossomos, sendo o cromossomo adicional um isocromossomo dicêntrico derivado do cromossomo 15, com ponto de quebra em q11.2 (Figura 8).

Este cromossomo derivado contém dois braços curtos do cromossomo 15 e duas cópias da região proximal do braço longo do 15, resultando em tetrassomia parcial do braço longo do 15 (47,XY,+idic(15)(q11.2)). A mãe do paciente relatou que sofreu três abortos e o óbito de uma filha de um ano e quatro meses antes do paciente aqui descrito, porém, os pais apresentaram cariótipos normais e esse cromossomo adicional, muito provavelmente, foi gerado durante a formação de gametas.

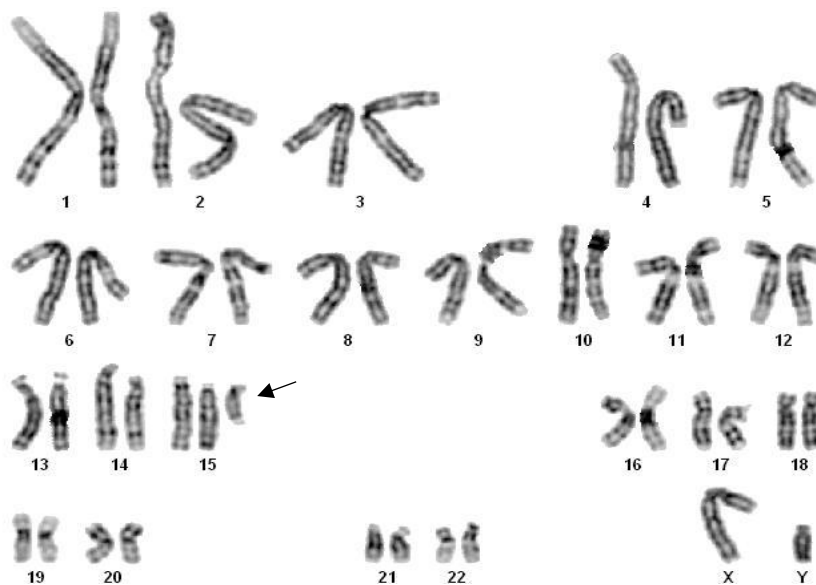


Figura 8: Cariótipo do paciente 67 com resultado 47,XY,+idic(15)(q11.2). A seta evidencia o cromossomo adicional idic(15).

No caso do paciente 144, devido às duas alterações encontradas serem nas porções terminais dos cromossomos, havia a suspeita de uma translocação desbalanceada entre os cromossomos 16 e 22. Essa translocação foi confirmada por cariótipo e FISH no paciente e nos pais onde foram analisadas 20 metáfases em ambos os métodos. No cariótipo do paciente é possível observar um excesso de material genético no braço longo do cromossomo 22. O método de FISH do paciente revelou um cromossomo 22 derivado de uma translocação entre os cromossomos 16 e 22, com pontos de quebra em p13.3 e q13 respectivamente. No cromossomo derivado 22 foi observado um sinal para a sonda da região 22q11.2 (TUPLE1, Abbott®), ausência de sinal para a sonda da região terminal 22q13 (ARSA, Abbott®) e um sinal para a sonda da região terminal 16p13.3 (12114, Cytocell Aquarius®). Concluiu-se que existe um rearranjo cromossômico desequilibrado causado pela translocação, resultando em monossomia parcial 22q13 e trissomia parcial 16p13.3

[46,XY.ish der(22)t(16;22)(p13.3;q13)(TUPLE1+,ARSA-,121I4+); Figura 9]. O método de FISH no pai revelou uma translocação entre os cromossomos 16 e 22, com pontos de quebra em p13.3 e q13 respectivamente. No cromossomo derivado 16 foi observado um sinal para a sonda da região terminal 22q13 (ARSA). No cromossomo derivado 22 foi observado um sinal para a sonda da região 22q11.2 (TUPLE1) e ausência de sinal para a sonda da região terminal 22q13 (ARSA), resultando em uma translocação cromossômica equilibrada entre os cromossomos 16 e 22 [46,XY.ish t(16;22)(p13.3;q13)(ARSA+;TUPLE1+,ARSA-); Figura 9]. A mãe do paciente apresentou cariótipo normal e o FISH revelou dois sinais para a sonda da região 22q11.2 e dois sinais para a sonda da região 22q13 em ambos os cromossomos 22 [46,XX.ish 22q11.2(TUPLE1×2)22q13(ARSA×2), Figura 10].

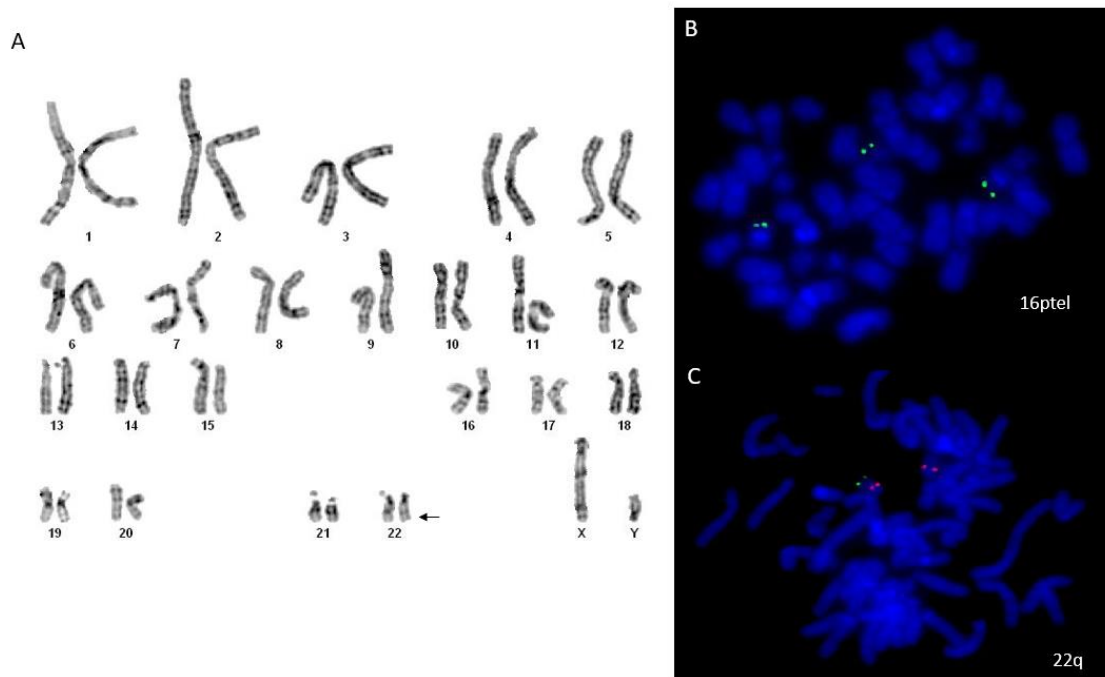


Figura 9: (A) Cariótipo do paciente 144 com resultado 46,XY.ish der(22)t(16;22)(p13.3;q13)(TUPLE1+,ARSA-,121I4+). A seta mostra o material adicional em 22q. (B) Análise de FISH em metáfase do paciente mostrando três sinais para 16p13.3 (verde). (C) Análise de FISH em metáfase do paciente mostrando um sinal para a região 16p13.3 (verde), dois sinais para região 22q11.2 (vermelho) e nenhum sinal para a região 22q13.

A deleção na região terminal do cromossomo 22 está relacionada com a síndrome de Phelan McDermid que pode ser causada por deleção terminal ou subtelo mérica, translocações balanceadas ou não balanceadas, cromossomo em anel, ou mosaicismo dessas anormalidades. De acordo com a base de dados do OMIM, essa síndrome está associada a hipotonia congênita, distúrbios de deglutição, hiporreflexia, alterações de marcha, dismorfismos, atraso e/ou involução do

desenvolvimento neuropsicomotor global, deficiência intelectual, transtorno do espectro autista, crises epilépticas, insensibilidade a dor, alterações renais e cardiovasculares, etc. Nosso paciente apresenta epilepsia de difícil controle com CTG e foi diagnosticado com LGS precedida de WS. Além disso, apresenta hipotonia global com coréia, hiperelasticidade, ADNPM, distúrbio de sucção e deglutição, hipotireoidismo, entre outras características clínicas sendo, o fenótipo do paciente, compatível com a síndrome de Phelan McDermid e com esse tipo de rearranjo cromossômico complexo. Uma translocação especificamente entre os cromossomos 16 e 22 já foi descrita em 2015 por Fontes *et al.*, coincidentemente também em um paciente do HC-UNICAMP, porém, a paciente do estudo apresentava uma duplicação no cromossomo 16 maior e uma deleção no cromossomo 22 menor do que as apresentadas por nosso paciente. Essa paciente apresentou microcefalia, dismorfismos, problemas gastrointestinais, anomalias vasculares, ADNPM, atraso de fala e problemas cardíacos e renais⁹⁶.

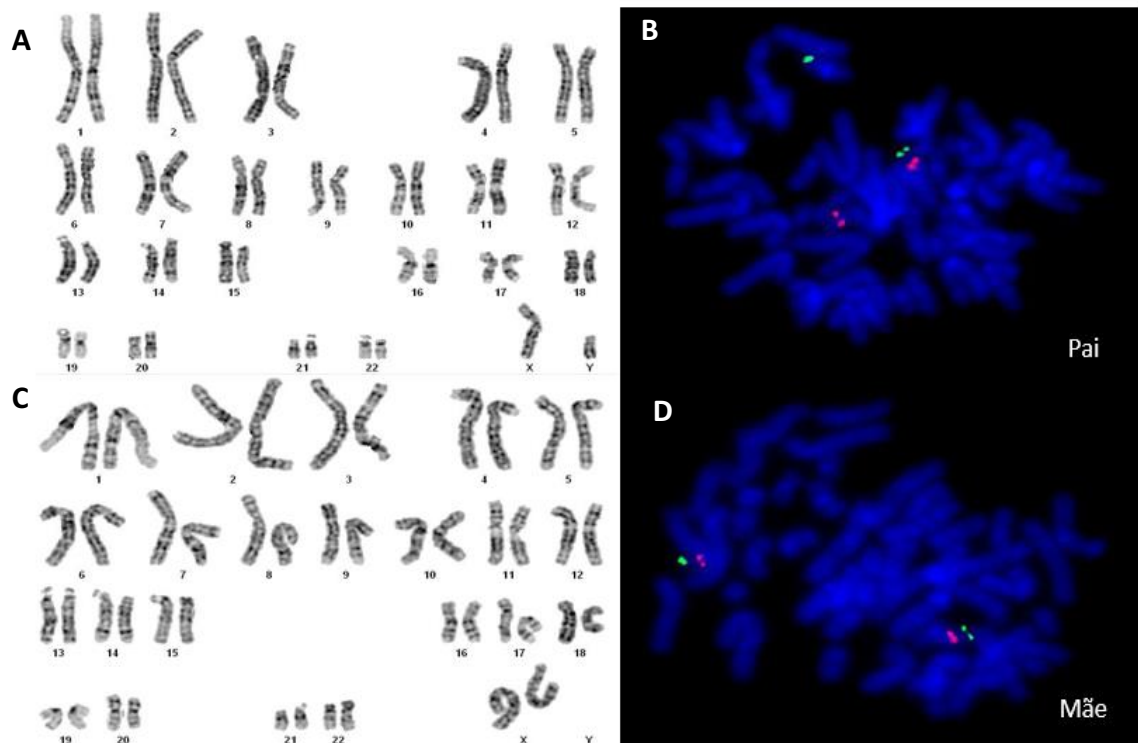


Figura 10: (A) Cariótipo do pai do paciente 144 com resultado 46,XY.ish t(16;22)(p13.3;q13)(ARSA+;TUPLE1+,ARSA-). (B) FISH do pai mostrando um sinal para a região 22q13 no cromossomo derivado 16 (verde) e um sinal para a região 22q11.2 (vermelho) e ausência de sinal para a região 22q13 no cromossomo derivado. (C) Cariótipo normal da mãe com resultado 46,XX. (D) FISH normal da mãe mostrando dois sinais para a região 22q11.2 (vermelho) e dois sinais para a região 22q13 (verde) em ambos os cromossomos 22.

A análise ROH revelou nove pacientes com regiões de homozigose maiores que 10Mb (Tabela 5). Sete pacientes possuem mais de 1% do genoma autossomo composto por regiões de homozigose com 10Mb de tamanho ou mais, e quatro deles possuem essas ROHs distribuídas por todo o genoma (pacientes 16, 128, 248 e 251), o que é característica de consanguinidade ou identidade por descendência. A maior porcentagem (26,06%), presente no paciente 251 é representativa de uma família com consanguinidade frequente ou entre parentes de primeiro grau⁹⁷. Nos outros três casos (pacientes 58, 76 e 97), apesar de possuírem porcentagem maior que 1%, não há indicação de consanguinidade ou identidade por descendência, pois as ROH não estão distribuídas por todo o genoma e são concentradas em apenas alguns cromossomos, o que pode ser resultado de dissomia uniparental segmentar⁹⁸. A paciente 31 do sexo feminino possui uma grande ROH no cromossomo X, porém, é comum encontrar um aumento de ROH nesse cromossomo devido a sua taxa de recombinação limitada. Além disso, pessoas do sexo masculino são sempre hemizigotos⁹⁸.

Essas regiões foram analisadas no WES para identificação de variantes em homozigose para doenças autossômicas recessivas e isso foi possível em apenas um paciente. O paciente 16 possui 4,36 de seu genoma autossomo em homozigose, sendo que a consanguinidade de seus progenitores também foi declarada pela família. Neste caso, na análise de WES através da plataforma *Varstation*, encontramos uma variante em homozigose no gene *NSUN2* que está relacionado a deficiência intelectual que pode ser acompanhada de epilepsia. Essa variante está localizada em uma dessas ROHs no cromossomo cinco. Em outros seis pacientes foi possível realizar o diagnóstico genético, mas nenhuma das variantes está localizada na ROH, estando todas em heterozigose. Nos cinco pacientes restantes não foram encontradas variantes causais, mesmo após análise dessas ROHs.

Taxa de Diagnóstico Genético

Para fins comparativos, separamos os tipos de análises realizados aqui em análise padrão e análise otimizada. A análise padrão de WES e CMA consiste em utilizar um programa para filtrar e priorizar as variantes encontradas e depois selecionar as variantes patogênicas e provavelmente patogênicas que tenham relação

com o fenótipo. Já a análise otimizada, consistiu em utilizar outros tipos de análises nos dados de WES, principalmente nos pacientes em que não foi possível encontrar uma causa genética, para aumentar a taxa de DG em nossos pacientes. Para isso realizamos a análise sem filtros com o programa *Varstation*, realizamos a chamada de CNVs a partir dos dados de WES e realizamos análise de segregação em variantes de significado incerto cuja classificação seria alterada para provavelmente patogênica se fossem não herdadas, ou seja, *de novo*.

Tabela 5: Pacientes que apresentaram grandes regiões de homozigose na análise de CMA. ID - número de identificação do paciente; LOH: Perda de heterozigose.

| ID | Sexo | Diagnóstico Genético através de WES | Regiões de LOH | % de LOH Autossomos | Fenótipo |
|-----|------|-------------------------------------|---|---------------------|----------------------------|
| 16 | M | NSUN2:c.2035G>A | Consanguinidade | 4,36 | Síndrome de Lennox-Gastaut |
| 17 | F | SCN1A:c.602+1G>A | arr[hg19] 10q21.1q22.1(54715100_70608911)x2 hmz arr[hg19] 10q22.1q23.1(73435426_85459984)x2 hmz | 0,92 | Síndrome de Dravet |
| 31 | F | SCN1A:c.5701C>T | arr[hg19] Xp11.4q23(39605561_113497351)x2 hmz | 0 | Síndrome de Dravet |
| 33 | F | SCN1A:c.5456_5458delAGT | arr[hg19] 7q22.3q33(106012679_135619612)x2 hmz | 0,98 | Síndrome de Dravet |
| 58 | F | - | arr[hg19] 7q31.32q36.1(122485983_150496416)x2 hmz arr[hg19] 9q33.1q34.11(121363428_132926691)x2 hmz | 1,31 | EED indeterminada |
| 76 | F | - | arr[hg19] 3q13.11q22.1(104031239_132647420)x2 hmz arr[hg19] 8q12.3q13.3(62826793_73577554)x2 hmz arr[hg19] 11p14.3p13(24804459_35032596)x2 hmz arr[hg19] 15q25.1q26.1(79950522_92933859)x2 hmz | 2,07 | EED indeterminada |
| 97 | M | - | arr[hg19] 1p21.2p13.2(101203113_111854120)x2 hmz arr[hg19] 9p22.2p13.1(18497828_39213995)x2 hmz arr[hg19] 12q15q22(70937094_93476695)x2 hmz arr[Hg19] 18p11.21q21.1(11283157_47016787)x2 hmz | 2,97 | Síndrome de Dravet |
| 99 | M | NPRL2:c.883C>T | arr[hg19] 14q24.2q31.3(70650867_85973847)x2 hmz | 0,50 | Síndrome de Lennox-Gastaut |
| 128 | M | - | Consanguinidade | 4,20 | EED indeterminada |
| 149 | M | KCNT1:c.2944-2A>C | arr[hg19] 14q23.1q31.3(58473958_85984782)x2 hmz | 0,91 | EED indeterminada |
| 248 | M | - | Consanguinidade | 13,50 | EED indeterminada |
| 251 | F | KCNQ2:c.790T>C | Consanguinidade | 26,06 | EED indeterminada |

Utilizando a análise padrão, foi possível realizar o DG em 73 pacientes utilizando WES e em 10 pacientes utilizando CMA. Considerando as duas técnicas combinadas foi possível encontrar uma causa genética em 83 ou 34,7% dos pacientes em nossa casuística. Quando adicionamos os resultados da análise otimizada de WES aumentamos em 10 o número de pacientes com DG. Sendo assim, foi possível

o DG em 89 pacientes utilizando somente WES e em 10 pacientes utilizando CMA. Somando as duas técnicas combinadas e todas as análises que foram realizadas, foi possível encontrar uma causa genética em 93 ou 38,9% dos pacientes. Esse número representa uma alta taxa de DG para essas doenças considerando a nossa coorte e foi atingido devido a análise otimizada realizada nesse trabalho. As taxas de DG de ambas as técnicas separadas e combinadas utilizando análise padrão + análise otimizada podem ser observadas na Tabela 6.

Tabela 6: Taxa de diagnóstico genético dos 239 pacientes com EEDs de acordo com a técnica utilizando a análise padrão + análise otimizada.

| Nº Pacientes por Técnica | | % Pacientes com DG | | |
|--------------------------|-----|--------------------|-----|-------------|
| | | SNP/Indel | CNV | Total |
| WES | 234 | 33,8 | 4,7 | 38 |
| CMA | 236 | – | 4,6 | 4,6 |
| Total | 239 | 33 | 6,3 | 38,9 |

Na Tabela 7 podemos observar as taxas de DG em outros estudos que variam de acordo com o número, fenótipo e critérios de inclusão dos pacientes, o tipo de técnica usada e método do teste, o número de genes analisados no painel, etc⁶³. Separamos esses estudos em quatro grupos: primeiro podemos observar as taxas de DG em estudos que avaliaram painéis gênicos; em segundo temos estudos que avaliaram somente WES ou sequenciamento de genoma completo (WGS); o terceiro grupo reúne quatro estudos que utilizaram técnicas de detecção de variantes de sequência e estruturais combinadas; e o último grupo reúne estudos que analisaram somente CNVs utilizando WES e/ou CMA.

As taxas de DG na literatura utilizando painéis, variam de 10 a 48,5%, porém, cada estudo deve ser analisado dentro de seu contexto. Na coorte com 48,5% de diagnóstico, por exemplo, foi analisado painel de 265 genes em 33 pacientes sendo que dez deles possuíam história familiar de epilepsia, o que faz com que seja mais provável encontrar uma variante pois ela segrega junto com a doença nos indivíduos afetados da família, que também foram analisados⁹⁹. Em outro caso analisando também painel de 67 genes, o mesmo foi observado: foi possível obter 47% de diagnóstico em 19 pacientes, sendo oito deles com histórico familiar¹⁰⁰. Outros artigos

utilizando painéis avaliaram de 28 a 500 pacientes com epilepsias e doenças do neurodesenvolvimento e obtiveram de 10 a 34,7% de sucesso no diagnóstico, isso devido ao baixo número de genes (de 38 a 71 genes) utilizados nos painéis^{84,86–88,101}.

Utilizando WES, as taxas de DG variam de 11 a 72%. Os estudos com altas taxas de DG analisaram um baixo número de pacientes (de 10 a 71 pacientes^{43–45}) e muitos indivíduos possuíam malformações cerebrais, microcefalia e dismorfismos. Os estudos que combinaram técnicas de detecção tanto de variantes de sequência quanto de variantes estruturais atingiram uma taxa de DG de 15,4 a 31,7%, sendo mais próximas da atingida em nosso trabalho, porém, ainda mais baixas^{84,90,102}.

Quanto ao poder de diagnóstico das CNVs, a taxa de DG na literatura varia de 4,8 a 28%^{46–48,103,104}. Aqui também precisamos considerar cada coorte individualmente para entender a seleção de pacientes e os vieses envolvidos que influenciam significativamente as taxas observadas. No estudo que atingiu 28% de DG, eles mesmos chamam a atenção para um viés de verificação que pode estar influenciando os resultados¹⁰⁴.

O WES apresentou o maior rendimento diagnóstico, porém, o uso de CMA conseguiu adicionar quase cinco pontos percentuais a esse rendimento e, portanto, deve ser também considerado na investigação da etiologia genética das EEDs. Por outro lado, a análise de CMA encontrou um maior número de CNVs causais do que a análise por WES, pois, na análise com WES foram usados filtros mais restritos, onde apenas CNVs que continham genes já relacionados às EEDs foram analisadas. A análise de CNV usando WES foi usada como uma análise complementar nesse trabalho e se mostrou eficiente para aumentar o poder de diagnóstico do WES. Porém, muita informação é perdida quando se utiliza um painel de genes como filtro na análise de CNVs.

Em relação ao sexo dos pacientes, foi possível diagnosticar geneticamente 43,1% das pacientes do sexo feminino e 35,4% dos pacientes do sexo masculino. Segundo teste de Fisher, não houve diferença estatisticamente significativa entre os sexos ($p= 0,2333$). Em relação aos subtipos de EEDs, não observamos nenhuma relação genótipo-fenótipo com exceção do gene *SCN1A* para síndrome de Dravet. Lembrando que nossa análise aqui foi apenas observacional.

Tabela 7: Taxas de diagnóstico genético (DG) em diferentes estudos com pacientes com epilepsias de início precoce ou da infância. WES – sequenciamento completo do exoma; WGS – sequenciamento completo do genoma; CMA – análise de microarranjos cromossômicos; MLPA – amplificação multiplex de sondas dependente de ligação; CNV – variante de número de cópia.

| Estudo | População | Nº Pacientes | Fenótipos | % de DG (Referência) |
|--------------------------------|------------------|--------------|---|------------------------|
| Painel 65 genes | EUA | 500 | Encefalopatias Epilépticas | 10 ⁸⁸ |
| Painel 38-327 genes | Canadá | 93 | Encefalopatias Epilépticas | 12,7 ⁸⁷ |
| Painel 38-53 genes | EUA | 28 | Epilepsias | 21,4 ¹⁰¹ |
| Painel 35 genes | Japão | 53 | EE de Início Precoce | 23 ¹⁰⁵ |
| Painel 71 genes | Austrália | 105 | Encefalopatias Epilépticas | 28,5 ⁸⁶ |
| Painel 172 genes | Coréia do Sul | 150 | EEDs | 34,7 ⁸⁹ |
| Painel 67 genes | Itália | 19 | Epilepsias de Início Precoce | 47 ¹⁰⁰ |
| Painel 265 genes | Alemanha e Suíça | 33 | Epilepsias | 48,5 ⁹⁹ |
| WES | EUA | 264 | Espasmos Infantis e LGS | 11 ⁴² |
| WES | Itália | 71 | Adultos com EEDs | 25,3 ⁴⁵ |
| WGS | Austrália | 30 | EEDs | 50 ¹⁰⁶ |
| WES | EUA | 10 | Encefalopatias Epilépticas | 70 ⁴³ |
| WES | Canadá | 11 | Epilepsias da Infância | 72 ⁴⁴ |
| Painel 70 genes CMA | EUA | 8565 | Epilepsia e Doenças do Neurodesenvolvimento | 15,4 ⁸⁴ |
| WES CMA | Canadá | 18 44 | Espasmos Infantis | 28 7 ¹⁰² |
| Painel >4600 genes MLPA | Índia | 82 | EEDs | 31,7 ⁹⁰ |
| Painel 104 genes CMA WGS | Escócia | 390 | Epilepsias da Infância | 31 ⁸ |
| CNV a partir de WES | EUA | 349 | Espasmos Infantis e LGS | 4,8 ¹⁰³ |
| CMA | EUA | 805 | Epilepsias | 5 ⁴⁸ |
| CMA | EUA | 315 | Encefalopatias Epilépticas | 7,9 ⁴⁶ |
| CMA | Noruega | 21 | LGS | 19 ⁴⁷ |
| CNV a partir de WES e CMA | China | 100 | Deficiência Intelectual e Epilepsia | 28 ¹⁰⁴ |

Observamos também que os pacientes com grandes CNVs que envolvem muitos genes apresentam quadros mais complexos do que pacientes com variantes patogênicas de sequência (SNVs ou indels) ou CNVs pequenas que envolvem poucos genes. Esses pacientes apresentam quadros mais sindrômicos devido à somatória de efeitos de cada gene deletado e/ou duplicado. Pacientes com rearranjos cromossômicos complexos, como os casos de cromossomo isodicêntrico e translocação aqui apresentados, apresentam quadros ainda mais graves do que os pacientes com variantes de sequência, CNVs pequenas que envolvem poucos genes e CNVs grandes, mas que envolvem somente uma região de um cromossomo. Esses achados podem ser muito úteis para priorizar e otimizar a indicação de testes genéticos

na prática clínica, visando o melhor custo-benefício para os pacientes e para os serviços e planos de saúde. Quando há suspeita de causa genética e as outras causas já foram descartadas após exames específicos, para pacientes com quadros sindrômicos envolvendo má formações, dismorfismos, ADNPM e/ou problemas neuropsiquiátricos, por exemplo, a análise de CNVs por CMA tem grande chance de ser eficiente para p DG¹⁰⁷. Já pacientes sem essas características e com suspeita de doença monogênica de causa desconhecida são elegíveis para teste de WES ou painéis multigênicos. Porém, alguns autores sugerem que, após triagem diagnóstica inicial (envolvendo EEG, exame de ressonância magnética, triagem de doenças metabólicas) os pacientes que não possuem diagnóstico devem ser indicados diretamente para exame de painel ou WES, incluindo a chamada de CNVs. Se após WES ou painel não for possível o DG, então o CMA seria indicado. Se mesmo após CMA não houver DG, outras estratégias podem ser usadas, como WGS, reanálises e reinterpretções periódicas dos exames já realizados, avaliar mutações dos tipos repetição e expansão, sequenciamento de genoma mitocondrial, variantes epigenéticas, etc⁸⁵. Pacientes com suspeita de causa monogênica conhecida ainda podem se beneficiar do sequenciamento de gene único, como o caso do gene *SCN1A* na síndrome de Dravet, por exemplo. Com o avanço das técnicas e programas de processamento e chamada de variantes, tem sido cada vez mais viável a detecção de CNVs a partir de WES e, preferencialmente de WGS. Por enquanto CMA continua sendo o padrão ouro para detecção de CNV, mas a ideia de se fazer apenas um teste abrangendo o maior número de possibilidades em termos de identificação de variantes causais está cada vez mais próxima.

Quando analisamos as taxas de DG em relação aos países (Tabela 8), Peru teve o maior rendimento com 50% dos pacientes diagnosticados, sendo que a maior parte do diagnóstico foi possível no grupo de EED indeterminadas. Em segundo lugar com maior DG temos o Chile, seguido por Brasil, Honduras e Argentina. Honduras apresentou uma alta taxa de DG, porém, deve-se considerar o viés causado pelo baixo número de pacientes analisados (N=3). Nos pacientes do Uruguai não foi possível obter DG, porém, apenas quatro pacientes foram analisados e dois deles possuem variantes em potencial que ainda necessitam serem validadas. No grupo da Argentina estão os oito pacientes com EMES, mas o DG não foi possível em nenhum deles. O padrão de POCS observado no EEG desses pacientes é apresentado também por

crianças com epilepsias estruturais ou metabólicas associadas a diferentes tipos de lesões cerebrais que podem estar relacionadas a causas genéticas, mas também a etiologias desconhecidas. Além disso, tanto as crises quanto o padrão de POCS podem desaparecer com o tempo na maioria dos pacientes e, enquanto que o POCS desaparece por volta da puberdade, o desaparecimento da epilepsia e os danos cognitivos dependem da etiologia¹⁰⁸. Na seleção de pacientes do HC-UNICAMP, pacientes com EMES não foram incluídos, por isso, esses oito pacientes podem estar baixando a nossa taxa de DG.

Tabela 8: Taxas de diagnóstico genético (DG) dos pacientes em relação ao país de origem.

| País | Nº de Pacientes | Nº Pacientes com DG | % Pacientes com DG |
|--------------|-----------------|---------------------|--------------------|
| Brasil | 146 | 61 | 41,8 |
| Argentina | 57 | 17 | 29,8 |
| Chile | 15 | 7 | 46,7 |
| Peru | 14 | 7 | 50 |
| Uruguai | 4 | 0 | 0 |
| Honduras | 3 | 1 | 33,3 |
| Total | 239 | 93 | 38,9 |

De todos os pacientes com DG, 51,6% possui variantes em genes ou regiões com possíveis abordagens terapêuticas, o que é um ponto muito importante quando falamos em medicina de precisão. Em mais da metade dos pacientes diagnosticados é possível realizar uma intervenção clínica, mudando e personalizando o tratamento para melhora das crises e da vida dos pacientes. O que demonstra, mais uma vez, a utilidade e aplicabilidade dos testes genéticos e a importância da realização de um DG o mais cedo possível em pacientes com EEDs. Os genes/diagnósticos para os quais há evidências de terapia de precisão, até o presente momento, em nossa coorte foram: *CDKL5*, *GAMT*, *KCNQ2*, *NPRL2*, *PCDH19*, *PRRT2*, *SCN1A*, *SLC6A1*, *STXBP1* e InvDup15^{63,85,109–111}.

É importante ressaltar que a relação genótipo-fenótipo foi muito bem estudada em todos os pacientes através de discussões periódicas com nosso grupo de especialistas da FCM-UNICAMP. As variantes encontradas foram relacionadas com os históricos presentes nos prontuários dos pacientes para seleção das variantes causais ou exclusão de variantes que não eram condizentes com os fenótipos.

O cenário aqui apresentado ainda pode sofrer alterações conforme o avanço das validações. Consideramos aqui somente as variantes patogênicas e provavelmente patogênicas já validadas ou que apresentaram altas coberturas, profundidades e parâmetros de qualidade adequados. Ainda, a ausência de diagnóstico molecular no restante dos pacientes não quer dizer que não haja variantes genéticas relacionadas ao fenótipo nesses indivíduos. Isso só ressalta a importância da aplicação de novas abordagens de análises, como a análise de novos genes que possam estar relacionados as EEDs e/ou análises que levem em conta um tipo de herança complexa.

Análise de Regressão Logística

Para essa análise foram usados os pacientes com dados pareados, ou seja, pacientes que possuem dados de WES e CMA ao mesmo tempo. Assim, foram usados inicialmente 208 pacientes, 203 controles e 914116 variantes. Após os filtros, as análises foram realizadas com 178 pacientes, 179 controles e 125796 variantes, sendo 105 pacientes do grupo sem DG e 73 pacientes no grupo com DG.

As distribuições entre os grupos foram comparadas e apresentadas em Gráficos Quantil-Quantil (QQ) que podem ser observados na Figura 11. Os gráficos, bem como os valores de lambda, indicam uma boa qualidade dos dados, ou seja, não há evidência de estruturação dentro dos grupos. O desvio da distribuição esperada do p-valor é evidente apenas na área da cauda, com lambdas próximos a um, sugerindo que a estratificação da população foi adequadamente controlada.

Os SNPs associados em cada grupo podem ser observados na Tabela 9. Na análise de regressão logística com variantes comuns, três SNPs (rs10962384, rs10786907 e rs12327363) ultrapassaram o primeiro valor de referência ($-\log_{10}(1e-05) = 5$) para o grupo total de pacientes, como podemos ver na Figura 12a, estando possivelmente associados a doença. Quando analisado o grupo com DG (Figura 12b) o único SNP com possível associação é o rs10962384 que também aparece no grupo total. Esse parece ser um SNP com uma associação mais robusta que pode estar influenciando o grupo de pacientes totais.

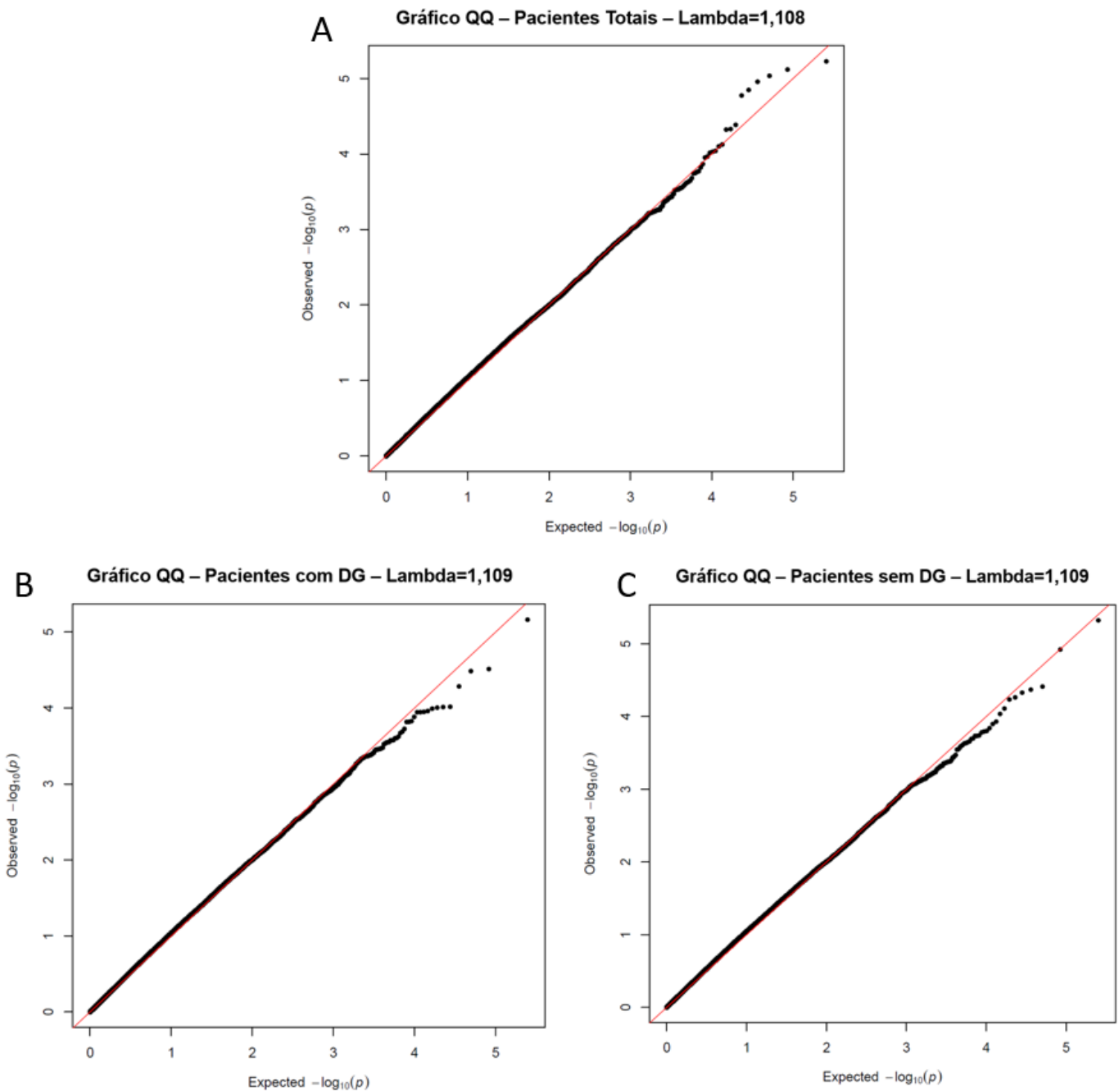


Figura 11: Gráficos QQ: O eixo Y mostra os valores de $-\log_{10} P$ observados, e o eixo X mostra os valores de $-\log_{10} P$ esperados. Cada SNP é representado por um ponto preto, e a linha vermelha indica a hipótese nula de não haver uma associação verdadeira. O desvio da distribuição esperada do p -valor é evidente apenas na área da cauda, com lambdas próximos a 1, sugerindo que a estratificação da população foi adequadamente controlada. (A) Gráfico QQ do grupo total de pacientes com lambda de 1,108. (B) Gráfico QQ de pacientes com diagnóstico genético (DG) e lambda de 1,109. (C) Gráfico QQ de pacientes sem diagnóstico genético e lambda de 1,109.

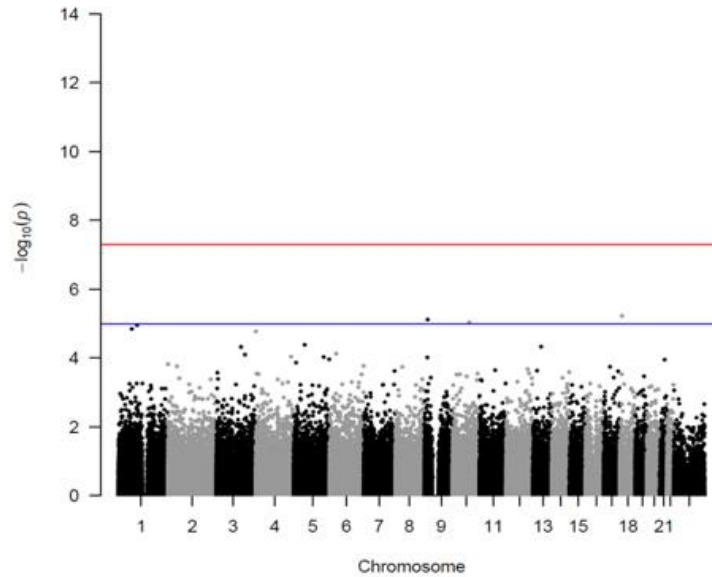
Tabela 9: SNPs associados a cada grupo mostrando suas posições, grupo em que estão associados e genes em que estão localizados. Chr: Cromossomo; Ref: Alelo de referência; Alt: Alelo alternativo.

| Chr | Posição | SNP | Ref | Alt | p-valor | Grupo | Gene | Tipo do Gene |
|-----|----------|------------|-----|-----|----------|--------------|---------------------|--------------|
| 6 | 98159736 | rs9374755 | T | C | 0.000005 | Sem DG | <i>LOC101927314</i> | Gene de RNA |
| 9 | 16336600 | rs10962384 | A | G | 0.000007 | Total/Com DG | <i>C9orf92</i> | Gene de RNA |
| 10 | 83908249 | rs10786907 | T | G | 0.000009 | Total | <i>NRG3</i> | Intrôn |
| 18 | 11410757 | rs12327363 | T | C | 0.000006 | Total | <i>LINC01928</i> | Gene de RNA |

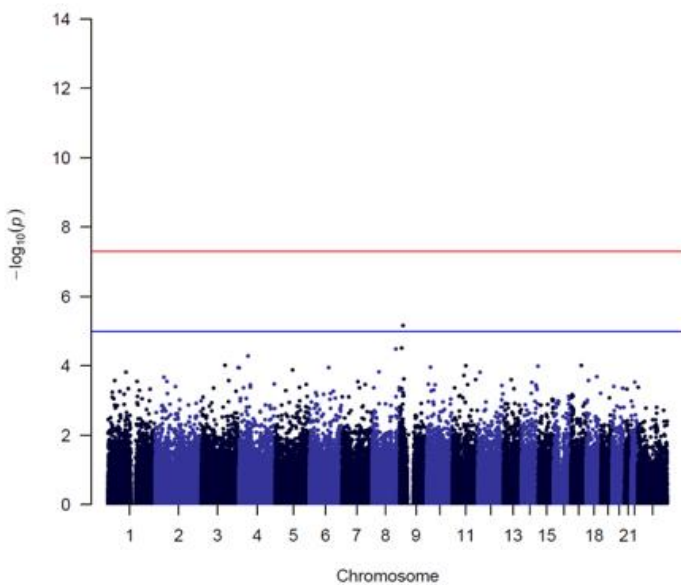
Quando analisamos somente o grupo de pacientes sem DG, apenas um SNP (rs9374755) ultrapassou o valor de referência, estando associado a esse grupo de pacientes (Figura 12c) e, possivelmente associado a doença, sendo um SNP candidato para a análise de herança complexa. Esse SNP está localizado em um gene de RNA (*LOC101927314*) categorizado como RNA longo não codificante (lncRNA). Tratam-se de RNAs com mais de 200pb que possuem papel essencial na regulação da transcrição e estão envolvidos em diversos processos biológicos^{112,113}. Não foram encontradas, ainda, publicações sobre esse gene específico, porém, a expressão anormal de lncRNAs ligados à epilepsia tem sido observada tanto em pacientes quanto em modelos animais. Devido à sua expressão no sistema nervoso e às suas funções nos processos celulares, uma desregulação ou expressão anormal de lncRNAs pode causar um comprometimento da função e do desenvolvimento cerebral, o que leva a vários distúrbios neurodegenerativos do desenvolvimento, como epilepsia, distúrbios do espectro do autismo, e doença de Alzheimer, porém, o papel funcional dessas moléculas na patogênese dessas doenças ainda está sendo explorado^{114,115}.

É importante ressaltar que nenhum SNP ultrapassou o segundo valor de referência ($-\log_{10}(5e-08) = 7.30103$), que seria o mais confiável para indicar associação. O que podemos sugerir aqui é que existe uma diferença entre os SNPs comuns encontrados entre os grupos com e sem DG quando comparados aos controles normais, ou seja, o grupo dos pacientes sem diagnóstico apresenta características diferentes dos grupos controle e com diagnóstico, o que seria uma evidência de que outro tipo de herança, que não a monogênica, pode estar envolvido nesse grupo de pacientes.

A Gráfico de Manhattan – Pacientes Totais



B Gráfico de Manhattan – Pacientes com DG



C Gráfico de Manhattan – Pacientes sem DG

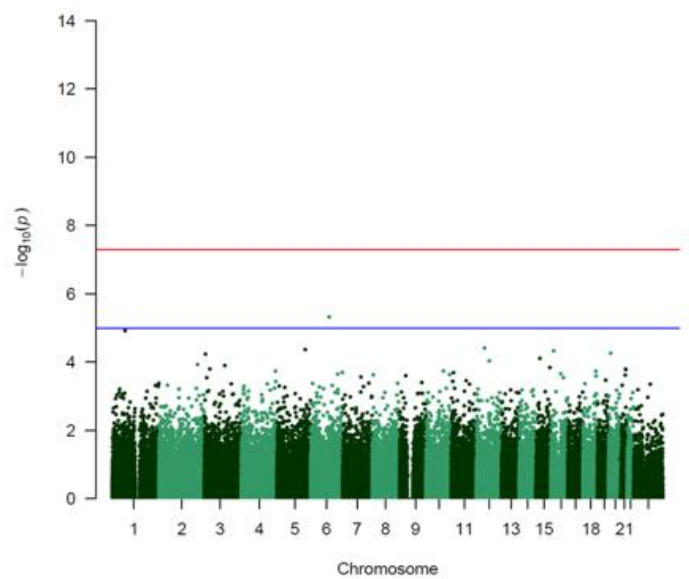


Figura 12: Gráficos de Manhattan obtidos a partir da análise de regressão logística. O eixo Y mostra $-\log_{10}$ dos p-valores, e o eixo X mostra as posições cromossômicas. A linha horizontal azul representa o valor de referência $-\log_{10}(1e-05) = 5$, e a linha horizontal vermelha representa o valor de referência $-\log_{10}(5e-08) = 7.30103$. (A) Gráfico de Manhattan para o grupo total de pacientes mostrando os três SNPs de interesse que ultrapassaram o primeiro valor de referência. (B) Gráfico de Manhattan para o grupo de pacientes com DG mostrando o único SNP de interesse que ultrapassou o primeiro valor de referência. (C) Gráfico de Manhattan para o grupo de pacientes sem DG mostrando o único SNP de interesse que ultrapassou o primeiro valor de referência.

Análise de SKAT-O

Nessa análise também foram usados os pacientes com dados pareados, porém, após os filtros, as análises foram realizadas com 117 pacientes, 130 controles e 299598 variantes, sendo 152694 variantes raras e 82248 variantes ultrarraras.

Quando analisamos conjuntos de variantes por gene, só encontramos evidências de associações significativas em todos os grupos quando foram analisadas as variantes ultrarraras (MAF = 0.0005), o que parece ser uma tendência desse tipo de análise, apesar de estarmos analisando doenças também raras.

No grupo total de pacientes, sete genes apresentaram p -valor ≤ 0.05 : *CACNA1H*, *AHDC1*, *TBCD*, *HERC1*, *WWOX*, *PTPN23* e *PCLO* (Tabela 10), sendo que seis deles também foram relacionados ao grupo de pacientes com DG (Tabela 11).

Tabela 10: Genes cujos grupos de variantes ultrarraras apresentaram p -valor significativo ($p \leq 0,05$) no grupo total de pacientes.

| Gene | Doenças Associadas aos Genes | p -valor | Nº de SNPs |
|----------------|---|-------------|------------|
| AHDC1 | EED, síndrome de Xia-Gibbs | 0.006988838 | 19 |
| CACNA1H | EED, epilepsia de ausências infantil | 0.004311538 | 34 |
| HERC1 | Microcefalia, dismorfismos faciais e atraso psicomotor | 0.014326329 | 21 |
| PCLO | Hipoplasia Pontocerebelar | 0.044966164 | 25 |
| PTPN23 | EED, epilepsia de início precoce refratária/ encefalopatia epilética e do desenvolvimento com hipomielinização e atrofia cerebral | 0.027643194 | 18 |
| TBCD | Encefalopatia precoce progressiva com atrofia cerebral e agenesia de corpo caloso | 0.008093057 | 26 |
| WWOX | EED, síndrome de West | 0.016897373 | 16 |

Identificamos grupos de variantes ultrarraras significativas nos genes *LAMC3*, *PTPN23*, *TBCD*, *FASN*, *RELN*, *CREBBP*, *ASPM*, *SPTAN1*, *RYR3* e *HERC1* para o grupo de pacientes sem DG (Tabela 12), sendo que sete desses genes são exclusivos para esse grupo. Esses sete genes candidatos podem conter variantes de predisposição às EEDs que estão, possivelmente, associadas a doença e devem ser consideradas sob a perspectiva de análises de herança complexa. Dos 10 genes associados a esse grupo, seis deles já foram relacionados às EEDs, sendo que apenas o gene *TBCD* aparece no diagnóstico genético em um paciente de nossa casuística.

Tabela 11: Genes cujos grupos de variantes ultrarraras apresentaram p-valor significativo ($p \leq 0,05$) no grupo de pacientes com DG.

| Gene | Doenças Associadas aos Genes | p-valor | Nº de SNPs |
|----------------|---|-------------|------------|
| AHDC1 | EED, síndrome de Xia-Gibbs | 0.001037311 | 19 |
| CACNA1H | EED, epilepsia de ausências infantil | 2.10004E-05 | 34 |
| HERC1 | Microcefalia, dismorfismos faciais e atraso psicomotor | 0.009037844 | 21 |
| PCLO | Hipoplasia Pontocerebelar | 0.044334276 | 25 |
| TBCD | Encefalopatia precoce progressiva com atrofia cerebral e agenesia de corpo caloso | 0.051756669 | 26 |
| WWOX | EED, síndrome de West | 0.019156085 | 16 |

Tabela 12: Genes cujos grupos de variantes ultrarraras apresentaram p-valor significativo ($p \leq 0,05$) no grupo de pacientes sem DG. * - genes exclusivos que só aparecem nesse grupo.

| Gene | Doenças Associadas aos Genes | p-valor | Nº de SNPs |
|----------------|---|-------------|------------|
| ASPM* | Microcefalia primária | 0.045517718 | 15 |
| CREBBP* | EED, síndrome de Menke-Hennekam | 0.032086627 | 16 |
| FASN* | EED | 0.009504475 | 29 |
| HERC1 | Microcefalia, dismorfismos faciais e atraso psicomotor | 0.053036286 | 21 |
| LAMC3* | Malformações occipitais do desenvolvimento cortical | 0.001417445 | 27 |
| PTPN23 | EED, epilepsia de início precoce refratária/ encefalopatia epilética e do desenvolvimento com hipomielinização e atrofia cerebral | 0.004014821 | 18 |
| RELN* | Lisencefalia e epilepsia de lobo temporal, familiar e Lisencefalia | 0.017928247 | 16 |
| RYR3* | EED | 0.047068168 | 27 |
| SPTAN1* | EED, síndrome de West | 0.045613193 | 10 |
| TBCD | Encefalopatia precoce progressiva com atrofia cerebral e agenesia de corpo caloso | 0.006100143 | 26 |

Após as análises de regressão logística e de SKAT-O, observamos que o grupo de pacientes com diagnóstico é um grupo muito bem caracterizado e acaba influenciando o resultado quando o grupo de pacientes totais é analisado. Já o grupo de pacientes sem diagnóstico, por ser um grupo mais heterogêneo, possui uma característica diferente dos outros grupos. Isso seria uma evidência de que pelo menos uma parte do efeito genético presente na etiologia das EEDs deve seguir um modelo complexo de herança.

Para epilepsias em geral, a análise de regressão logística tem sido usada para prever associações em termos de tratamentos, resposta à medicamentos, desfecho de cirurgias, etc. Já a análise de SKAT-O tem sido usada, principalmente em estudos

de farmacogenética. Por fim, é muito importante enfatizar que essa é a primeira vez que esses dois tipos de abordagens (regressão logística e SKAT-O) são utilizados para explorar possíveis associações a níveis de variantes e de genes causais nas EEDs e se mostram mais adequadas quando analisamos essas doenças pela perspectiva de herança complexa, como a herança poligênica, por exemplo.

Apesar de não ser ainda possível aplicar essas análises ao diagnóstico genético de rotina, nossos resultados indicam que uma etiologia genética deve ser considerada em pacientes com EED mesmo que resultados negativos sejam obtidos utilizando as técnicas genômicas de rotina.

À medida que as compreensões acerca das causas genéticas das epilepsias e acerca das doenças genéticas de herança complexa e multifatorial se expandem, o conceito e o uso da medicina de precisão (incluindo terapias gene específicas) também continua a se expandir, aumentando as descobertas genéticas com possibilidade de intervenção clínica e, acreditamos que as EEDs poderão se beneficiar desses novos conhecimentos e novos métodos de análise dos dados genéticos.

CONCLUSÕES

Foi possível obter pacientes com caracterizações clínicas adequadas e foi possível realizar uma caracterização genética bem abrangente no grupo de pacientes com EEDs da América Latina. Seguindo a ordem dos objetivos específicos propostos, pontuamos as seguintes conclusões:

- Foi possível reunir uma grande coorte com 275 pacientes com EEDs, sendo que a obtenção de amostras de DNA foi possível para 239 desses pacientes;
- Foi possível obter dados de WES para 234 pacientes e dados de CMA para 236 pacientes, sendo que a taxa de perda de pacientes durante os processos foi muito baixa;
- Obtivemos altas taxas de DG em nossos pacientes, sendo que 38% foram diagnosticados utilizando somente WES, 4,6% foram diagnosticados utilizando somente CMA, e 38,9% foram diagnosticados utilizando ambas as análises combinadas;
- Encontramos evidências de que pelo menos uma parte do efeito genético presente na etiologia das EEDs deve seguir um modelo complexo de herança e, portanto, uma causa genética deve ser considerada e explorada mesmo nos pacientes com resultados negativos de exames que utilizam as técnicas genômicas de rotina.

REFERÊNCIAS

1. McTague, A., Howell, K. B., Cross, J. H., Kurian, M. A. & Scheffer, I. E. The genetic landscape of the epileptic encephalopathies of infancy and childhood. *Lancet Neurol.* **15**, 304–316 (2016).
2. Hurst, D. L. Epidemiology of Severe Myoclonic Epilepsy of Infancy. **31**, 397–400 (1990).
3. Cowan, L. D. & Hudson, L. S. Epidemiology and Natural History of Infantile Spasms. 355–364 (1991).
4. Hino-fukuyo, N., Haginoya, K., Iinuma, K., Uematsu, M. & Tsuchiya, S. Neuroepidemiology of West syndrome and early infantile epileptic encephalopathy in Miyagi Prefecture, Japan. 299–301 (2009). doi:10.1016/j.eplepsyres.2009.09.012
5. Bayat, A., Hjalgrim, H. & Møller, R. S. The incidence of SCN1A -related Dravet syndrome in Denmark is 1 : 22 , 000 : A population-based study from 2004 to 2009. 36–39 (2015). doi:10.1111/epi.12927
6. Behr, C., Goltzene, M. A., Kosmalski, G., Hirsch, E. & Ryvlin, P. Epidemiology of epilepsy. *Rev. Neurol. (Paris)*. **172**, 27–36 (2015).
7. Ware, T. L. *et al.* Epidemiology and etiology of infantile developmental and epileptic encephalopathies in Tasmania. 504–510 (2019). doi:10.1002/epi4.12350
8. Symonds, J. D. *et al.* Early childhood epilepsies : epidemiology , classification , aetiology , and socio-economic determinants. *Brain* **144**, 2879–2891 (2021).
9. Keezer, M. R., Sisodiya, S. M. & Sander, J. W. Comorbidities of epilepsy : current concepts and future perspectives. *Lancet Neurol.* **15**, 106–115 (2016).
10. Jeffrey, J. S., Leathem, J., Ross, K., Sadleir, L. G. & Mefford, H. C. Developmental and epileptic encephalopathy : Personal utility of a genetic diagnosis for families. *Epilepsia Open* **6**, 149–159 (2021).
11. Berg, A. T. *et al.* Revised terminology and concepts for organization of seizures and epilepsies: Report of the ILAE Commission on Classification and Terminology, 2005-2009. *Epilepsia* **51**, 676–685 (2010).
12. Scheffer, I. E. *et al.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* **58**, 512–521 (2017).
13. Specchio, N. & Curatolo, P. Developmental and epileptic encephalopathies: What we do and do not know. *Brain* **144**, 32–43 (2021).
14. Tavyev, Y. J. & Scaglia, F. European Journal of Medical Genetics Molecular bases and clinical spectrum of early infantile epileptic encephalopathies. *Eur. J. Med. Genet.* **55**, 299–306 (2012).
15. Weber, Y. G. & Lerche, H. Genetic mechanisms in idiopathic epilepsies. *Dev. Med. Child Neurol.* **50**, 648–654 (2008).
16. Gonsales, M. C. *et al.* Recent developments in the genetics of childhood

- epileptic encephalopathies: Impact in clinical practice. *Arq. Neuropsiquiatr.* **73**, 946–958 (2015).
17. Lennox, W. THE HEREDITY OF EPILEPSY AS TOLD BY RELATIVES AND TWINS William. *JAMA Neurol.* **146**, 529–536 (1951).
 18. Metrakos, K. & Metrakos, J. D. Genetics of convulsive disorders II. Genetic and eletroencephalographic studies in centrencephalic epilepsy. *Neurology* (1961).
 19. Steinlein, O. K. *et al.* A missense mutation in the neuronal nicotinic acetylcholine receptor alpha 4 subunit is associated with autosomal dominant nocturnal frontal lobe epilepsy. *Nat Genet* **11**, 201–203 (1995).
 20. Charlier, C. *et al.* A pore mutation in a novel KQT-like potassium channel gene in an idiopathic epilepsy family. *Nat. Genet.* **18**, 53–55 (1998).
 21. Singh, N. *et al.* A novel potassium channel gene, KCNQ2, is mutated in an inherited epilepsy of newborns. *Nat. Genet.* **18**, 25–29 (1998).
 22. Wallace, R. H. *et al.* Febrile seizures and generalized epilepsy associated with a mutation in the Na⁺-channel β 1 subunit gene SCN1B. *Nat. Genet.* **19**, 366–370 (1998).
 23. Escayg, A. *et al.* Mutations of SCN1A, encoding a neuronal sodium channel, in two families with GEFS+2. *Nat. Genet.* **24**, 343–345 (2000).
 24. Claes, L. *et al.* De Novo Mutations in the Sodium-Channel Gene SCN1A Cause Severe Myoclonic Epilepsy of Infancy. *Am. J. Hum. Genet.* **68**, 1327–1332 (2001).
 25. Lopes-Cendes, I. & Oliveira Ribeiro, P. A. Aspectos genéticos das epilepsias: uma visão atual. *Rev. Médica Clínica Las Condes* **24**, 903–908 (2013).
 26. Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145 (2008).
 27. Chandler, M. R., Bilgili, E. P. & Merner, N. D. A Review of Whole-Exome Sequencing Efforts Toward Hereditary Breast Cancer Susceptibility Gene Discovery. *Hum. Mutat.* **37**, 835–846 (2016).
 28. Helbig, I. & Abou Tayoun, A. N. Understanding genotypes and phenotypes in epileptic encephalopathies. *Mol. Syndromol.* **7**, 172–181 (2016).
 29. Depienne, C. *et al.* Spectrum of SCN1A gene mutations associated with Dravet syndrome: Analysis of 333 patients. *J. Med. Genet.* **46**, 183–191 (2009).
 30. Brunklaus, A. *et al.* The clinical utility of an SCN1A genetic diagnosis in infantile-onset epilepsy. *Dev. Med. Child Neurol.* **55**, 154–161 (2013).
 31. Hildebrand, M. S. *et al.* Recent advances in the molecular genetics of epilepsy. *J. Med. Genet.* **50**, 271–279 (2013).
 32. Scheffer, I. E. *et al.* ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology. *Epilepsia* **58**, 512–521 (2017).
 33. Gonsales, M. C. *et al.* Multimodal Analysis of SCN1A Missense Variants

- Improves Interpretation of Clinically Relevant Variants in Dravet Syndrome. *Front. Neurol.* **10**, 1–10 (2019).
34. Warman Chardon, J., Beaulieu, C., Hartley, T., Boycott, K. M. & Dymont, D. A. Axons to Exons: the Molecular Diagnosis of Rare Neurological Diseases by Next-Generation Sequencing. *Curr. Neurol. Neurosci. Rep.* **15**, 1–8 (2015).
 35. Rehm, H. L. Disease-targeted sequencing: A cornerstone in the clinic. *Nature Reviews Genetics* (2013). doi:10.1038/nrg3463
 36. Myers, K. A., Johnstone, D. L. & Dymont, D. A. Epilepsy genetics: Current knowledge, applications, and future directions. *Clin. Genet.* **95**, 95–111 (2019).
 37. Bamshad, M. J., Ng, S. B., Bigham, A. W. & Tabor, H. K. Mendelian disease gene discovery. *Nat. Publ. Gr.* **12**, 745–755 (2011).
 38. Kiezun, A. *et al.* Exome sequencing and the genetic basis of complex traits. *NIH Public Access* **44**, 623–630 (2013).
 39. Foo, J. N., Liu, J. J. & Tan, E. K. Whole-genome and whole-exome sequencing in neurological diseases. *Nat. Rev. Neurol.* **8**, 508–517 (2012).
 40. Boycott, K. M., Vanstone, M. R., Bulman, D. E. & MacKenzie, A. E. Rare-disease genetics in the era of next-generation sequencing: Discovery to translation. *Nat. Rev. Genet.* **14**, 681–691 (2013).
 41. Olson, H. E. *et al.* A Recurrent De Novo PACS2 Heterozygous Missense Variant Causes Neonatal-Onset Developmental Epileptic Encephalopathy, Facial Dysmorphism, and Cerebellar Dysgenesis. *Am. J. Hum. Genet.* **102**, 995–1007 (2018).
 42. Epi4K, C. *et al.* De novo mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
 43. Veeramah, K. R. *et al.* Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia* **54**, 1270–1281 (2013).
 44. Dymont, D. A. *et al.* Whole-exome sequencing broadens the phenotypic spectrum of rare pediatric epilepsy: A retrospective study. *Clin. Genet.* **88**, 34–40 (2015).
 45. Minardi, R., Licchetta, L., Baroni, M. C., Pippucci, T. & Stipa, C. Whole-exome sequencing in adult patients with developmental and epileptic encephalopathy : it is never too late. *Clin. Genet.* 1–12 (2020). doi:10.1111/cge.13823
 46. Mefford, H. C. *et al.* Rare copy number variants are an important cause of epileptic encephalopathies. *Ann. Neurol.* **70**, 974–985 (2011).
 47. Lund, C., Brodtkorb, E., Røsby, O., Rødningen, O. K. & Selmer, K. K. Copy number variants in adult patients with Lennox-Gastaut syndrome features. *Epilepsy Res.* **105**, 110–117 (2013).
 48. Olson, H. *et al.* Copy number variation plays an important role in clinical epilepsy. *Ann. Neurol.* **75**, 943–958 (2014).
 49. Magalhães, P. H. M., Moraes, H. T., Athie, M. C. P., Secolin, R. & Lopes-Cendes, I. New avenues in molecular genetics for the diagnosis and

- application of therapeutics to the epilepsies. *Epilepsy Behav.* **121**, 1–11 (2021).
50. Ohmori, I. *et al.* A CACNB4 mutation shows that altered Cav2.1 function may be a genetic modifier of severe myoclonic epilepsy in infancy. *Neurobiol. Dis.* **32**, 349–354 (2008).
 51. Singh, N. A. *et al.* A role of SCN9A in human epilepsies, as a cause of febrile seizures and as a potential modifier of Dravet syndrome. *PLoS Genet.* **5**, 1–12 (2009).
 52. Hawkins, N. A., Martin, M. S., Frankel, W. N., Kearney, J. A. & Escayg, A. Neuronal voltage-gated ion channels are genetic modifiers of generalized epilepsy with febrile seizures plus. *Neurobiol. Dis.* **41**, 655–660 (2011).
 53. Hawkins, N. A. & Kearney, J. A. Hlf is a genetic modifier of epilepsy caused by voltage-gated sodium channel mutations. *Epilepsy Res.* **119**, 20–23 (2016).
 54. Calhoun, J. D., Hawkins, N. A., Zachwieja, N. J. & Kearney, J. A. Cacna1g is a genetic modifier of epilepsy caused by mutation of voltage-gated sodium channel Scn2a. *Epilepsia* **57**, e103–e107 (2016).
 55. Calhoun, J. D., Hawkins, N. A., Zachwieja, N. J. & Kearney, J. A. Cacna1g is a genetic modifier of epilepsy in a mouse model of Dravet syndrome. *Epilepsia* **58**, e111–e115 (2017).
 56. Sheidley, B. R. & Poduri, A. Genetics in clinical epilepsy: Issues in genetic testing and counseling. *J. Pediatr. Epilepsy* **1**, 135–142 (2012).
 57. Wirrell, E. C. & Nabbout, R. Recent Advances in the Drug Treatment of Dravet Syndrome. *CNS Drugs* **33**, 867–881 (2019).
 58. Gardella, E. & Moller, R. S. Phenotypic and genetic spectrum of SCN8A - related disorders , treatment options , and outcomes. *Epilepsia* **60**, 77–85 (2019).
 59. Wang, D. *et al.* Glut-1 Deficiency Syndrome : Clinical , Genetic , and Therapeutic Aspects. *Ann. Neurol.* **57**, 111–118 (2005).
 60. Sofou, K. *et al.* Ketogenic diet in pyruvate dehydrogenase complex deficiency : short- and long-term outcomes. *J. Inherit. Metab. Dis.* **40**, 237–245 (2017).
 61. Zuberi, S. M. *et al.* ILAE Classification & Definition of Epilepsy Syndromes in the Neonate and Infant: Position Statement by the ILAE Task Force on Nosology and Definitions. *Int. Leag. against Epilepsy* **63**, 1349–1397 (2022).
 62. Ohtahara, S. & Yamatogi, Y. Epileptic encephalopathies in early infancy with suppression-burst. *J. Clin. Neurophysiol.* **20**, 398–407 (2003).
 63. Symonds, J. D. *et al.* Incidence and phenotypes of childhood-onset genetic epilepsies : a prospective population-based national cohort. *Brain* **142**, 2303–2318 (2019).
 64. Sharma, S. & Prasad, A. N. Genetic testing of epileptic encephalopathies of infancy: an approach. *Can. J. Neurol. Sci.* **40**, 10–6 (2013).
 65. Dravet, C. The core Dravet syndrome phenotype. *Epilepsia* **52**, 3–9 (2011).

66. Scheffer, I. E. & Nabbout, R. SCN1A-related phenotypes : Epilepsy and beyond. *Epilepsia* **60**, 17–24 (2019).
67. Steel, D., Symonds, J. D., Zuberi, S. M. & Brunklaus, A. Dravet syndrome and its mimics : Beyond SCN1A. *Epilepsia* **58**, 1807–1816 (2017).
68. Kelley, S. A. & Kossoff, E. H. Doose syndrome (myoclonic – astatic epilepsy) : 40 years of progress. *Dev. Med. Child Neurol.* **52**, 988–993 (2010).
69. Tang, S. & Pal, D. K. Dissecting the genetic basis of myoclonic-astatic epilepsy. *Epilepsia* **53**, 1303–1313 (2012).
70. Specchio, N. *et al.* International League Against Epilepsy classification and definition of epilepsy syndromes with onset in childhood: Position paper by the ILAE Task Force on Nosology and Definitions. *Epilepsia* **63**, 1398–1442 (2022).
71. Berg, A. T., Levy, S. R. & Testa, F. M. Evolution and course of early life developmental encephalopathic epilepsies : Focus on Lennox - Gastaut syndrome. *Epilepsia* **59**, 2096–2105 (2018).
72. Lee, Y. J. *et al.* Resective pediatric epilepsy surgery in Lennox-Gastaut syndrome. *Pediatrics* **125**, 58–66 (2010).
73. Asadi-pooya, A. A. Lennox-Gastaut syndrome : a comprehensive review. (2017).
74. Lesca, G. *et al.* Update on the genetics of the epilepsy-aphasia spectrum and role of GRIN2A mutations. *Epileptic Disord.* **21**, 41–47 (2019).
75. Tsai, M. *et al.* Clinical genetic study of the epilepsy-aphasia spectrum. *Epilepsia* **54**, 280–287 (2013).
76. Stephani, U. The natural history of myoclonic astatic epilepsy (Doose syndrome) and Lennox-Gastaut syndrome. *Epilepsia* **47**, 53–55 (2006).
77. Faria, A. *et al.* Varstation: a complete and efficient tool to support NGS data analysis. *bioRxiv* 833582 (2019). doi:10.1101/833582
78. Szpiech, Z. A. *et al.* Long Runs of Homozygosity Are Enriched for Deleterious Variation. *Am. J. Hum. Genet.* **93**, 90–102 (2013).
79. Rocha, C. S., Secolin, R., Rodrigues, M. R., Carvalho, B. S. & Lopes- Cendes, I. The Brazilian Initiative on Precision Medicine (BIPMed) : fostering genomic data-sharing of underrepresented populations. *npj Genomic Med.* **42**, 1–7 (2020).
80. Li, H. A statistical framework for SNP calling , mutation discovery , association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
81. Purcell, S. *et al.* PLINK : A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
82. Wu, M. C. *et al.* Rare-Variant Association Testing for Sequencing Data with the Sequence Kernel Association Test. *Am. J. Hum. Genet.* **89**, 82–93 (2011).

83. Secolin, R. *et al.* Distribution of local ancestry and evidence of adaptation in admixed populations. *Sci. Rep.* **9**, 1–12 (2019).
84. Lindy, A. S. *et al.* Diagnostic outcomes for genetic testing of 70 genes in 8565 patients with epilepsy and neurodevelopmental disorders. *Epilepsia* **59**, 1062–1071 (2018).
85. Beltrán-corbellini, Á. *et al.* Epilepsy Genetics and Precision Medicine in Adults : A New Landscape for Developmental and Epileptic Encephalopathies. *Front. Neurol.* **13**, 1–12 (2022).
86. Kothur, K. *et al.* Diagnostic yield of targeted massively parallel sequencing in children with epileptic encephalopathy. *Seizure* **59**, 132–140 (2018).
87. Mercimek-mahmutoglu, S. *et al.* Diagnostic yield of genetic testing in epileptic encephalopathy in childhood. *Epilepsia* **56**, 707–716 (2015).
88. Carvill, G. L. *et al.* Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat. Publ. Gr.* **45**, 825–830 (2013).
89. Na, J. H. *et al.* Targeted gene panel sequencing in early infantile onset developmental and epileptic encephalopathy. *Brain Dev.* **42**, 438–448 (2020).
90. Mitta, N. *et al.* Genotype-phenotype correlates of infantile-onset developmental & epileptic encephalopathy syndromes in South India : A single centre experience. *Epilepsy Res.* **166**, 1–10 (2020).
91. Nabbout, R. & Dulac, O. Epileptic Encephalopathies : A Brief Overview. *J. Clin. Neurophysiol.* **20**, 393–397 (2003).
92. Mei, D., Cetica, V., Marini, C. & Guerrini, R. Dravet syndrome as part of the clinical and genetic spectrum of sodium channel epilepsies and encephalopathies. *Epilepsia* **60**, 2–7 (2019).
93. Makoff, A. J. & Flomen, R. H. Detailed analysis of 15q11-q14 sequence corrects errors and gaps in the public access sequence to fully reveal large segmental duplications at breakpoints for Prader-Willi , Angelman , and inv dup (15) syndromes. *Genome Biol.* **8**, 14.1-14.16 (2007).
94. Van Dyke, D. L., Weiss, L., Logan, M. & Pai, G. S. The Origin and Behavior of Two Isodicentric Bisateilited Chromosomes. *Am J Hum Genet* **29**, 294–300 (1977).
95. Bataglia, D. *et al.* Epileptic disorders within the first year of life: neurological and cognitive outcome. *Eur. J. Pediatr. Neurol.* **3**, 95–103 (1999).
96. Fontes, M. I. B. *et al.* Genotype-phenotype correlation of 16p13.3 terminal duplication and 22q13.33 deletion: Natural history of a patient and review of the literature. *Am. J. Med. Genet. Part A* **170**, 766–772 (2016).
97. Sund, K. L. *et al.* Regions of homozygosity identified by SNP microarray analysis aid in the diagnosis of autosomal recessive disease and incidentally detect parental blood relationships. *Genet. Med.* **15**, 70–78 (2013).
98. Kearney, H. M., Kearney, J. B. & Conlin, L. K. Diagnostic implications of excessive homozygosity detected by SNP-Based microarrays: Consanguinity,

- uniparental disomy, and recessive single-gene mutations. *Clin. Lab. Med.* **31**, 595–613 (2011).
99. Lemke, J. R. *et al.* Targeted next generation sequencing as a diagnostic tool in epileptic disorders. *Epilepsia* **53**, 1387–1398 (2012).
 100. Della Mina, E. *et al.* Improving molecular diagnosis in epilepsy by a dedicated high-throughput sequencing platform. *Eur. J. Hum. Genet.* **23**, 354–362 (2015).
 101. Wang, J., Gotway, G., Pascual, J. M. & Park, J. Y. Diagnostic Yield of Clinical Next-Generation Sequencing Panels for Epilepsy. *JAMA Neurol.* **71**, 2013–2014 (2014).
 102. Michaud, J. L. *et al.* The genetic landscape of infantile spasms. *Hum. Mol. Genet.* **23**, 4846–4858 (2014).
 103. Epi4K, C., Project, E. P. & Mefford, H. Copy number variant analysis from exome data in 349 patients with epileptic encephalopathy. *Ann. Neurol.* **78**, 323–328 (2015).
 104. Kessi, M. *et al.* Rare Copy Number Variations and Predictors in Children With Intellectual Disability and Epilepsy. *Front. Neurol.* **9**, 1–12 (2018).
 105. Kodera, H. *et al.* Targeted capture and sequencing for detection of mutations causing early onset epileptic encephalopathy. *Epilepsia* **54**, 1262–1269 (2013).
 106. Palmer, E. E. *et al.* Diagnostic Yield of Whole Genome Sequencing After Nondiagnostic Exome Sequencing or Gene Panel in Developmental and Epileptic Encephalopathies. *Neurology* **96**, e1770–e1782 (2021).
 107. Striano, P. *et al.* Clinical significance of rare copy number variations in epilepsy: A case-control survey using microarray-based comparative genomic hybridization. *Arch. Neurol.* **69**, 322–330 (2012).
 108. Caraballo, R., Pavlidis, E., Nikanorova, M. & Loddenkemper, T. Encephalopathy with continuous spike-waves during slow-wave sleep: evolution and prognosis. *Epileptic Disord.* **21**, S15–S21 (2019).
 109. Nabbout, R. & Kuchenbuch, M. Impact of predictive, preventive and precision medicine strategies in epilepsy. *Nat. Rev. Neurol.* **16**, 674–688 (2020).
 110. Striano, P. & Minassian, B. A. From Genetic Testing to Precision Medicine in Epilepsy. *Neurotherapeutics* **17**, 609–615 (2020).
 111. Bayat, A., Bayat, M., Rubboli, G. & Møller, R. S. Epilepsy syndromes in the first year of life and usefulness of genetic testing for precision therapy. *Genes (Basel)*. **12**, 1–19 (2021).
 112. Ma, L. *et al.* LncBook : a curated knowledgebase of human long non-coding RNAs. *Nucleic Acids Res.* **47**, 128–134 (2019).
 113. Chen, J., Wang, Y., Wang, C., Hu, J. & Li, W. LncRNA Functions as a New Emerging Epigenetic Factor in Determining the Fate of Stem Cells. *Front. Genet.* **11**, 1–12 (2020).
 114. Wan, P., Su, W. & Zhuo, Y. The Role of Long Noncoding RNAs in Neurodegenerative Diseases. *Mol. Neurobiol.* **54**, 2012–2021 (2017).

115. Villa, C., Lavitrano, M. & Combi, R. Long Non-Coding RNAs and Related Molecular Pathways in the Pathogenesis of Epilepsy. *Int. J. Mol. Sci.* **20**, 1–14 (2019).

ANEXOS

I. Parecer do CEP

FACULDADE DE CIÊNCIAS
MÉDICAS - UNICAMP
(CAMPUS CAMPINAS)



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: BIORREPOSITÓRIO
ESTUDOS DE GENÉTICA MOLECULAR EM DOENÇAS NEUROPSIQUIÁTRICAS
FASE I

Pesquisador: Iscia Teresinha Lopes Cendes

Área Temática: Área 1. Genética Humana.

(Trata-se de pesquisa envolvendo genética humana não contemplada acima.);

Versão: 2

CAAE: 12112913.3.0000.5404

Instituição Proponente: Hospital de Clínicas da UNICAMP

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 257.020

Data da Relatoria: 12/04/2013

Apresentação do Projeto:

Trata-se de um projeto para implantação de biorepositório de doenças neuro-psiquiátricas e casos-controle. O estudo prevê recrutamento e coleta de 700 pacientes e 300 indivíduos controle.

O presente projeto pretende estudar os aspectos moleculares das seguintes doenças: epilepsias, malformações do desenvolvimento cortical, coreias, ataxias, paraparesias espásticas, distonias, transtorno afetivo bipolar, esquizofrenia, doenças musculares, doenças mitocondriais, doença de Parkinson, acidente vascular cerebral e demências. O projeto está dividido em sub-projetos, com a descrição detalhada das estratégias que serão utilizadas para cada uma dessas doenças.

Serão utilizadas diversas técnicas de biologia molecular para identificação de mutações conhecidas ou novas, como PCR, sequenciamento convencional e de terceira geração e análises de bioinformática. No Subprojeto 1 (Epilepsias e Malformações do Desenvolvimento Cortical), serão avaliadas mutações através da implantação da tecnologia de sequenciamento, baseada em equipamento de terceira geração e um novo sistema de detecção eletroônico, sistema Ion Torrent®. No subprojeto 2 será realizada a captura e o sequenciamento do exoma em amostras de DNA de

Endereço: Rua Tessália Vieira de Camargo, 126

Bairro: Barão Geraldo

CEP: 13.083-887

UF: SP

Município: CAMPINAS

Telefone: (19)3521-8936

Fax: (19)3521-7187

E-mail: cep@fom.unicamp.br

FACULDADE DE CIÊNCIAS
MÉDICAS - UNICAMP
(CAMPUS CAMPINAS)



indivíduos afetados por formas familiares de epilepsia mioclônica juvenil (EMJ), epilepsia do lobo temporal (ELT) e benigna rolandica epilepsia (BRE) e Malformações do desenvolvimento cortical (MCD) nas quais as mutações já descritas não foram identificadas. Nessa investigação serão utilizados kits de captura e enriquecimento Illumina TruSeq®, além do sequenciamento 11 usando a Illumina Hi-seq 2.000. Após o sequenciamento, será realizada análise de bioinformática nos dados obtidos a fim de identificar variantes potencialmente patogênicas.

Os pacientes serão recrutados nos ambulatórios do HC-UNICAMP (Neurogenética, Epilepsia, Epilepsia de difícil controle, Epilepsia infantil, Distúrbio do Movimento, Doenças Neuromusculares, Distonias, Toxina Botulínica, Demeências, Neurovascular ou Psiquiatria). Será coletado de 20-30 ml de sangue periférico para o estudo, além do exame clínico e anamnese. Será coletados dados do prontuário médico.

Objetivo da Pesquisa:

Identificar as mutações responsáveis pelos respectivos fenótipos, em um grupo de doenças neuropsiquiátricas, incluindo: epilepsias, malformações do desenvolvimento cortical, coreias, ataxias, paraparesias espásticas, distonias, transtorno afetivo bipolar, esquizofrenia, doenças musculares, doenças mitocondriais, acidente vascular cerebral, doença de Parkinson e demências.

Avaliação dos Riscos e Benefícios:

Os procedimentos a serem realizados são de pequeno risco para o indivíduo, sendo o principal deles a coleta de sangue por punção venosa.

Como benefícios, em alguns casos será possível a identificação das mutações associadas às doenças dos pacientes incluídos no estudo, o que implica na possibilidade de algoritmos mais eficientes de diagnóstico e tratamento. Os autores alertam que, muitas vezes, a mutação não resulta em uma mudança no tratamento, o que é compreensível quando se trata de estudos genéticos.

Informações geradas durante o projeto e que possam ter implicações na confirmação diagnóstica de indivíduos sintomáticos serão comunicadas aos profissionais responsáveis pelo acompanhamento destes pacientes, nas sessões de orientação e aconselhamento genético dos ambulatórios de Neurogenética, cuja pesquisadora principal é a responsável.

Comentários e Considerações sobre a Pesquisa:

O projeto está bem escrito e detalhado, assim como os procedimentos que serão aplicados aos voluntários saudáveis e portadores de doença. Apresenta relevância científica e social, a médio e longo prazos. O estudo será patrocinado pelo próprio pesquisador. Foram acrescentadas ao projeto principal as informações sobre o local onde serão recrutados e onde serão coletados os exames do grupo controle.

Considerações sobre os Termos de apresentação obrigatória:

O TCLE está bem redigido, é de fácil entendimento para o voluntário e prevê armazenamento de

Endereço: Rua Tessália Vieira de Camargo, 126
Bairro: Barão Geraldo CEP: 13.083-887
UF: SP Município: CAMPINAS
Telefone: (19)3521-8938 Fax: (19)3521-7187 E-mail: cep@fcm.unicamp.br

FACULDADE DE CIÊNCIAS
MÉDICAS - UNICAMP
(CAMPUS CAMPINAS)



material biológico. Estão presentes todos os termos de apresentação obrigatória previstos pela Resolução 196/96 e complementares, assim como o "Regulamento do Biorepositório de Doenças Neuropsiquiátricas!".

Recomendações:

Nada a declarar.

Conclusões ou Pendências e Lista de Inadequações:

Foram acrescentadas ao projeto principal as informações sobre o recrutamento dos voluntários do grupo controle, com priorização inicial para membros da família de pacientes, porém não portadores das doenças. As amostras serão coletadas nos ambulatórios de Genética e Neurologia HC-Unicamp e Hemocentro-UNICAMP.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Aprovado em reunião do colegiado, em 23 de Abril de 2013.

CAMPINAS, 26 de Abril de 2013

Assinador por:
Fátima Aparecida Bottcher Luiz
(Coordenador)

Endereço: Rua Tessália Vieira de Camargo, 126
Bairro: Barão Geraldo **CEP:** 13.083-887
UF: SP **Município:** CAMPINAS
Telefone: (19)3521-8936 **Fax:** (19)3521-7187 **E-mail:** cep@fcm.unicamp.br

II. Termo de Consentimento



Versão2- nov/2018

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO PARA PESQUISA (Menor de 18 anos ou legalmente incapaz)

Título do projeto: **Estudos de genética molecular em doenças neuropsiquiátricas – Fase I**
Pesquisadora Responsável: **Profa. Dra. Iscia Lopes Cendes (tel: 19 3521 8242)**

OBJETIVO DA PESQUISA:

O menor ou indivíduo legalmente incapaz, a qual você representa legalmente, está sendo convidado(a) a participar de um projeto de pesquisa em neurociência e neurotecnologia envolvendo pacientes e famílias de indivíduos com:

- | | | |
|---|--|---|
| <input type="checkbox"/> Acidente Vascular Cerebral | <input type="checkbox"/> Distonias | <input type="checkbox"/> Grupo Controle |
| <input type="checkbox"/> Ataxias | <input type="checkbox"/> Doenças mitocondriais | <input type="checkbox"/> Malformações corticais |
| <input type="checkbox"/> Crises Febris na Infância | <input type="checkbox"/> Doenças musculares | <input type="checkbox"/> Paraparesias |
| <input type="checkbox"/> Coréias | <input type="checkbox"/> Epilepsias | <input type="checkbox"/> Parkinson |
| <input type="checkbox"/> Demências | <input type="checkbox"/> Esquizofrenia | <input type="checkbox"/> Transtorno afetivo bipolar |

O objetivo desse estudo é identificar alteração genética que causa doenças e também estudar a população normal para comparação. Isso poderá melhorar o diagnóstico das doenças (se a alteração for encontrada) e poderá levar a um melhor tratamento no futuro. No entanto, muito provavelmente, o tratamento do(a) participante não será modificado com a participação nesse estudo.

ESCLARECIMENTOS:

- 1) A concessão do material é livre e não causará quaisquer prejuízos pessoais ou ao andamento do tratamento do participante.
- 2) As informações pessoais do(a) participante serão mantidas em sigilo e privacidade. Seus dados serão codificados e criptografados por sistemas informatizados. Somente dados relevantes para as pesquisas serão utilizados, mantendo como prioridade o anonimato dos(as) participantes. Toda informação médica, mas não os resultados dos testes genéticos decorrentes desse projeto de pesquisa, farão parte do prontuário médico e serão submetidos aos regulamentos do HC- UNICAMP referentes ao sigilo da informação médica.
- 3) Você, como responsável legal do(a) participante, terá acesso gratuito, se assim desejar:
 - a) às informações sobre as finalidades do armazenamento, incluindo seu responsável, os riscos e benefícios potenciais, as garantias de qualidade de conservação e integridade do material biológico, e as medidas para garantir a privacidade e confidencialidade;
 - b) às informações e resultados obtidos a partir do material biológico utilizado, inclusive, aquelas que impliquem em riscos para doenças não preveníveis ou riscos familiares;
 - c) ao aconselhamento genético, quando aplicável.
- 4) Você, como responsável legal, poderá retirar seu consentimento de uso da amostra do participante a qualquer momento, sem prejuízo ou penalização alguma às partes envolvidas. Uma vez desautorizada, você, como responsável legal do(a) concedente, será solicitado(a) a formalizar e assinar sua manifestação por escrito, cabendo ao biorrepositório o descarte ou devolução da amostra, se assim for solicitado. A pesquisadora responsável pode interromper a participação nesse estudo a qualquer momento que julgar apropriado.
- 5) Você, como responsável legal, será comunicado(a) se houver a necessidade de descarte do material armazenado, o que poderá ocorrer se a amostra não atender a critérios mínimos de qualidade para pesquisa, se houver dificuldades institucionais para seu armazenamento (espaço) ou se o biorrepositório deixar de existir. Caso o material seja descartado, será feito de acordo com as normas de biossegurança legais.
- 6) O material concedido ficará sob a guarda da Faculdade de Ciências Médicas do Hospital de Clínicas da UNICAMP (FCM-UNICAMP), sob cuidados dos(as) responsáveis pelo biorrepositório, e será utilizado para o estudo genético da doença em questão. Os dados gerados do estudo poderão ser divulgados em artigos e/ou congressos, resguardando-se sempre o sigilo quanto à identidade dos participantes da pesquisa.

Rubrica do(a) pesquisador(a)

Rubrica do(a) responsável legal

- 7) Este TCLE deverá ser elaborado em duas vias, rubricado em todas as páginas e assinado pelo(a) responsável legal e pelo(a) pesquisador(a) responsável. Uma via deverá ficar com o responsável legal e outra com o(a) pesquisador(a).
- 8) Em caso de dúvidas sobre as amostras concedidas ao biorrepositório, o(a) responsável pelo(a) participante poderá entrar em contato com: Profª D^{ra} Iscia Lopes Cendes, telefones: (19) 3521-7754 ou (19) 3521- 8242.
- 9) Em caso de denúncias ou reclamações sobre sua participação e sobre questões éticas de estudo, o(a) responsável pelo(a) participante poderá entrar em contato com a secretaria do Comitê de Ética em Pesquisa (CÉP) da UNICAMP das 08:30 às 11:30 e das 13:00 às 17:00 na Rua Tessália Vieira de Camargo, 126, CEP: 13083-887, Campinas – SP; telefones: (19) 3521-8936 ou (19) 3521-7187; e-mail: cep@fcm.unicamp.br.
- 10) Os esclarecimentos para material biológicos acima também são aplicáveis quando houver concessão somente de resultados de exames genéticos (genoma, exoma, painel gênico, etc), Neste caso, desconsidere os itens “RISCO E DESCONFORTO DOS PROCEDIMENTOS DE COLETAS” e “ACOMPANHAMENTO E ASSISTÊNCIA”.

COLETA DE MATERIAL:

O(s) pesquisador(a) fará perguntas sobre os antecedentes médicos e familiares do participante, que poderá ser convidado(a) a participar de exame físico, neurológico e/ou psiquiátrico para estabelecimento de seu estado clínico. Poderão ser solicitados exames como eletroencefalograma (EEG), eletromiografia (EMG) tomografia computadorizada e ressonância magnética de crânio. Será ainda convidado(a) a realizar coleta de material biológico, como especificado abaixo.

O(a) responsável pelo(a) participante autoriza a coleta/concessão do seguinte material:

- | | |
|---|---|
| <input type="checkbox"/> Biópsia de pele (2 a 4 mm) | <input type="checkbox"/> Resultados de exames genéticos |
| <input type="checkbox"/> Fezes (5g ou um copo de coleta) | <input type="checkbox"/> Sangue (20 a 30 ml) |
| <input type="checkbox"/> Raspado da mucosa bucal / saliva | |

Caso seja necessário, poderão fazer parte desta pesquisa informações clínicas disponíveis em prontuários médicos, com isso você, responsável pelo(a) participante:

- Autoriza** o uso de dados clínicos presentes no prontuário ou laudo médico do(a) participante.
 Não autoriza o uso de dados clínicos presentes no prontuário ou laudo médico do(a) participante.

RISCO E DESCONFORTO DOS PROCEDIMENTOS DE COLETAS:

Biópsia de pele: é uma pequena cirurgia, geralmente superficial. O procedimento é realizado por um médico, com instrumentos e materiais estéreis e há risco de infecção e dor local. Eventualmente pode ser necessário o uso de pontos para fechamento do local. Nesse caso, será necessário o uso de anestesia local e prescrição de medicamentos para controlar a dor e inflamação. Se houver sinal de infecção, antibióticos serão necessários. Há a necessidade de comparecer novamente ao ambulatório para retirada dos pontos e haverá orientação para a realização dos curativos e procedimentos de cuidado e higiene locais.

Fezes: será feita em casa, seguindo as orientações que forem passadas pelo(a) pesquisador(a), ou será feita no ambulatório. Não há desconforto ou risco associados, somente será necessário cuidado ao transportar o material, caso a coleta seja feita em casa, para evitar vazamentos, mau cheiro e contaminações.

Sangue: podem ocorrer dor e manchas roxas (equimoses) no local da coleta. O desconforto será mínimo, pois se trata de uma coleta de sangue geralmente da veia do braço que será realizada por profissional treinado e habilitado para realizar esse procedimento.

Para raspado de mucosa bucal/saliva e resultados de exames genéticos não existem riscos ou desconfortos associados.

ACOMPANHAMENTO E ASSISTÊNCIA:

Não haverá necessidade de acompanhamento médico específico. No caso de alguma intercorrência médica o(a) responsável pelo(a) participante deve procurar a pesquisadora responsável para assistência.

| | |
|------------------------------|---------------------------------|
| Rubrica do(a) pesquisador(a) | Rubrica do(a) responsável legal |
|------------------------------|---------------------------------|

Em caso de danos ocorridos de sua participação, o(a) participante tem direito a indenização prevista a qualquer cidadão pelo código civil. Se for paciente, sua participação será feita em conjunto com os retornos agendados para consulta no HC-UNICAMP, porém, se for necessário uma visita fora dos retornos agendados para consulta, poderemos ressarcir os gastos com transporte e alimentação; essa prerrogativa também é válida para voluntários pertencentes ao grupo controle, bastando para isso avisar a pesquisadora responsável previamente.

O diagnóstico de familiares do(a) participante que não apresentam sintomas (pré-sintomáticos) não faz parte dessa pesquisa e não é oferecido em nosso serviço. Mas se o(a) responsável pelo(a) participante desejar obter orientação/aconselhamento genético, será oferecido no ambulatório de neurogenética do HC/Unicamp, tel. (19) 3521-7754.

DISPONIBILIZAÇÃO PÚBLICA DOS DADOS DESTA PESQUISA:

Com o avanço das pesquisas na área da genética, é de grande importância compartilhar as informações de alguns testes genéticos em bancos de dados públicos. No compartilhamento desses dados é sempre assegurado que não haverá divulgação dos identificadores (nome, filiação, endereço, número de registro hospitalar), mas poderão ser incluídas informações relevantes para a interpretação dos dados genéticos (sexo, idade, país e região de procedência, diagnóstico e outras informações clínicas pertinentes). Sendo assim, o(a) responsável pelo participante:

- Autoriza** a disponibilização dos dados genéticos em bancos públicos, que poderão ser consultados por pesquisadores(as) da área médica, desde que esses dados não sejam vinculados com identificadores da amostra (veja acima exemplos de identificadores).
- Não autoriza** a disponibilização dos dados genéticos realizados neste projeto de pesquisa.

Em caso de falecimento ou condição incapacitante, os direitos sobre o material armazenado deverão ser concedidos a: _____ (pedir para o(a) responsável legal indicar o nome de uma pessoa a ser contatada).

Diante destas informações, o(a) responsável pelo(a) participante declara consentimento livre para ceder o material de origem biológica para o Biorrepositório do BRAINN, assinando o presente termo:

Campinas, _____ de _____ de 20_____.

Nome do(a) participante

HC do(a) participante

Nome do(a) responsável pelo(a) participante

Assinatura do(a) responsável pelo(a) participante

Nome da testemunha

Assinatura da testemunha

RESPONSABILIDADE DO(A) PESQUISADOR(A):

O(a) pesquisador(a) explicou a (nome do(a) responsável pelo(a) participante) _____ o objetivo do estudo, os procedimentos requeridos e os possíveis riscos e vantagens que poderão advir do estudo, usando o melhor de seu conhecimento e se compromete a fornecer uma via desse termo de consentimento ao participante.

Nome do(a) pesquisador(a) ou associado(a)

Assinatura do(a) pesquisador(a) ou associado(a)

III. Listas de Genes Usadas nas Análises

Tabela 13: Lista EEDG contendo genes relacionados às EEDs

| | | | | | | | |
|----------|---------|--------|---------|--------|---------|----------|---------|
| AARS1 | CAD | FARS2 | HDAC4 | MAPK10 | PLPBP | SERPINI1 | SZT2 |
| ADGRL2 | CASK | FGF12 | HNRNPU | MBD5 | PNKP | SIK1 | TBC1D24 |
| ADORA2A | CDKL5 | FLNA | HOXD@ | MECP2 | PNPO | SLC19A3 | TCF4 |
| ALG13 | CHD2 | FOLR1 | IQSEC2 | MEF2C | POLG | SLC1A2 | TNK2 |
| AP3B2 | CHRNA7 | FOXG1 | IRF2BPL | MTMR1 | PRRT2 | SLC25A22 | TOR1A |
| ARFGEF2 | CLCN4 | GABRA1 | KCNA2 | MTOR | PTPN23 | SLC2A1 | TRMT9B |
| ARHGEF15 | CNPY3 | GABRB3 | KCNH5 | NECAP1 | QARS1 | SLC35A2 | TRPM1 |
| ARHGEF9 | CNTNAP2 | GABRG2 | KCNQ2 | NEDD4L | RARS2 | SLC35A3 | TWNK |
| ARX | CPA6 | GNAO1 | KCNQ3 | NRG2 | RB1 | SLC9A6 | UBA5 |
| ASAH1 | CSNK1E | GRIA3 | KCNQ5 | OR10H2 | RHOBTB2 | SNAP25 | VRK2 |
| ATP1A2 | CYFIP2 | GRIN1 | KCNT1 | PACS2 | SCN1A | SPTAN1 | WDR45 |
| ATP6V1A | DCX | GRIN2A | KCNT2 | PCDH19 | SCN1B | SRGAP2 | YWHAG |
| BRAT1 | DOCK7 | GRIN2B | KCTD3 | PCDHG@ | SCN2A | SRPX2 | ZEB2 |
| CACNA1A | EEF1A2 | GRIN2D | KCTD7 | PIGA | SCN3A | STXBP1 | ZMYND8 |
| CACNA1E | ENG | HADHB | KLF13 | PIGQ | SCN8A | SYNGAP1 | ZNF182 |
| CACNA2D2 | ERBB4 | HCN1 | MAGI2 | PLCB1 | SEPSECS | SYNJ1 | |

Tabela 14: Lista EG contendo genes relacionados às epilepsias em geral.

| | | | | | | | |
|----------|---------|--------|---------|----------|----------|----------|---------|
| AARS1 | CDK5 | EFHC1 | GUF1 | MBD5 | PIGP | SERPINI1 | TMTC3 |
| ACER3 | CDKL5 | EIF2S3 | HACE1 | MDH2 | PIGQ | SGCE | TNK2 |
| ADAM22 | CENPE | EMX2 | HADHB | MECP2 | PIGT | SHH | TOR1A |
| ADGRG1 | CENPJ | ENG | HCN1 | MED17 | PIK3R2 | SIK1 | TPP1 |
| ADGRL2 | CERS1 | EPM2A | HDAC4 | MEF2C | PLAA | SIX3 | TRIO |
| ADGRV1 | CHD2 | EPRS1 | HECW2 | MFSD2A | PLCB1 | SLC12A5 | TRMT10A |
| ADORA2A | CHRNA2 | ERBB4 | HERC1 | MFSD8 | PLEKHG2 | SLC12A6 | TRMT9B |
| ADRA2B | CHRNA4 | ERMARD | HEXA | MOCS1 | PLPBP | SLC13A5 | TRPM1 |
| ADSL | CHRNA7 | EXOSC3 | HEXB | MOCS2 | PNKP | SLC19A3 | TSC1 |
| AIMP2 | CHRN2 | EXT2 | HNRNPU | MPDZ | PNPO | SLC1A2 | TSC2 |
| ALDH7A1 | CILK1 | FARS2 | HOXD@ | MTMR1 | POLG | SLC20A2 | TSEN15 |
| ALG13 | CLCN2 | FGF12 | IER3IP1 | MTOR | POLG2 | SLC25A12 | TSEN2 |
| AMPD2 | CLCN4 | FIG4 | IQSEC2 | NACC1 | PPP1R15B | SLC25A22 | TSEN54 |
| AMT | CLDN5 | FLNA | IRF2BPL | NDE1 | PPP3CA | SLC2A1 | TUBA1A |
| ANKLE2 | CLN3 | FOLR1 | ITPA | NECAP1 | PPT1 | SLC35A2 | TUBA8 |
| AP3B2 | CLN5 | FOXG1 | KANSL1 | NEDD4L | PRDM8 | SLC35A3 | TUBB2A |
| ARFGEF2 | CLN6 | FRRS1L | KATNB1 | NHLRC1 | PRICKLE1 | SLC45A1 | TUBB2B |
| ARHGEF15 | CLN8 | GABBR2 | KCNA1 | NPC1 | PRICKLE2 | SLC6A1 | TUBB3 |
| ARHGEF9 | CLP1 | GABRA1 | KCNA2 | NPC2 | PRRT2 | SLC6A8 | TUBG1 |
| ARV1 | CLTC | GABRA2 | KCNB1 | NPRL2 | PTCH1 | SLC6A9 | TWNK |
| ARX | CNNM2 | GABRA3 | KCNC1 | NPRL3 | PTPN23 | SLC9A6 | UBA5 |
| ASAH1 | CNPY3 | GABRA5 | KCNH5 | NR4A2 | QARS1 | SMC1A | UBE3A |
| ASPM | CNTN2 | GABRB1 | KCNJ10 | NRG2 | RAB11A | SMS | UFC1 |
| ATN1 | CNTNAP2 | GABRB2 | KCNMA1 | NRXN1 | RARS2 | SNAP25 | UFM1 |
| ATP13A2 | COL4A2 | GABRB3 | KCNQ2 | NSDHL | RB1 | SNIP1 | VARA1 |
| ATP1A2 | CPA6 | GABRD | KCNQ3 | NTRK2 | RBFOX1 | SPATA5 | VPS53 |
| ATP6V1A | CPLX1 | GABRG2 | KCNQ5 | NUS1 | RELN | SPTAN1 | VRK2 |
| ATP7A | CSNK1E | GAL | KCNT1 | OCLN | RHOBTB2 | SRGAP2 | WASF1 |
| ATP8A2 | CSTB | GAMT | KCNT2 | OPHN1 | ROGDI | SRPX2 | WDR45 |
| BRAF | CTSD | GATM | KCTD17 | OR10H2 | RORB | ST3GAL3 | WDR45B |
| BRAT1 | CYFIP2 | GBA | KCTD3 | OTUD6B | RPH3A | ST3GAL5 | WDR62 |
| BSCL2 | DCX | GCSH | KCTD7 | PACS2 | RTN4IP1 | STAMPB | WDR73 |
| CACNA1A | DEAF1 | GLDC | KIF11 | PAFAH1B1 | RTTN | STRADA | WWOX |
| CACNA1D | DENND5A | GNAO1 | KIF2A | PCDH12 | SARS1 | STX1B | XPR1 |
| CACNA1E | DEPDC5 | GOSR2 | KIF5C | PCDH19 | SASS6 | STXBP1 | YWHAG |
| CACNA1H | DHDDS | GPAA1 | KLF13 | PCDHG@ | SCARB2 | SUOX | ZEB2 |
| CACNA2D2 | DIAPH1 | GRIA3 | LAMB1 | PCLO | SCN1A | SYN1 | ZMYND8 |
| CACNB4 | DIP2A | GRIA4 | LAMC3 | PDHA1 | SCN1B | SYNGAP1 | ZNF182 |
| CAD | DLAT | GRIN1 | LGI1 | PDHX | SCN2A | SYNJ1 | |
| CASK | DNAJC5 | GRIN2A | LIAS | PDP1 | SCN3A | SZT2 | |
| CASR | DNM1 | GRIN2B | LMNB2 | PIGA | SCN8A | TBC1D24 | |
| CCDC88A | DOCK7 | GRIN2D | MAGI2 | PIGC | SCN9A | TBCD | |
| CCDC88C | EEF1A2 | GRN | MAPK10 | PIGN | SEPSECS | TCF4 | |

Tabela 15: Lista PG contendo genes potenciais para epilepsias.

| | | | | | | | | | |
|----------|----------|---------|---------|----------|--------|---------|---------|---------|---------|
| AAAS | BRCA2 | COL6A2 | F2 | GRM3 | KCNC4 | MAN2A2 | NONO | PTEN | SOBP |
| AASS | BRD2 | COL6A3 | FADD | GRM4 | KCND1 | MANBA | NOTCH1 | PTGS1 | SON |
| ABAT | BRWD3 | COQ2 | FAM111A | GRM5 | KCND2 | MAOA | NPC1L1 | PTH | SOX2 |
| ABCC8 | BSN | COQ4 | FAN1 | GRM6 | KCND3 | MAPRE2 | NR2F1 | PUF60 | SOX5 |
| ABHD12 | BSND | COQ6 | FAR1 | GRM7 | KCNE1 | MAPT | NRAS | PURA | SPR |
| ACADS | C12orf57 | COQ8A | FARP2 | GRM8 | KCNE2 | MATN4 | NRG3 | PUS3 | SPTLC2 |
| ACADSB | C18orf25 | COQ9 | FASN | GTPBP3 | KCNE3 | MBTPS2 | NSD1 | QDPR | SQSTM1 |
| ACMSD | C3 | CORO1A | FASTKD2 | GUCY1A1 | KCNE4 | MCCC1 | NSD2 | RAB18 | SSR4 |
| ACO2 | CACNA1B | COX10 | FAT4 | GYS1 | KCNE5 | MCCC2 | NUBPL | RAB39B | ST7 |
| ACOT7 | CACNA1C | COX6B1 | FBXL4 | HADHA | KCNF1 | MCM8 | OCA2 | RAI1 | ST8SIA2 |
| ACOX1 | CACNA1F | COX8A | FBXO28 | HAX1 | KCNG1 | MCM9 | OFD1 | RANBP2 | STAT1 |
| ACSF3 | CACNA1G | CP | FGD1 | HCFC1 | KCNG2 | ME2 | OPA1 | RANGAP1 | STAT2 |
| ACTB | CACNA1I | CPS1 | FGF8 | HCK | KCNG3 | MED12 | OPLAH | RAPGEF6 | STK11 |
| ACTG1 | CACNA1S | CPT2 | FGFR2 | HCN2 | KCNG4 | MED13L | OPRM1 | RAPSN | STT3A |
| ACVR1 | CACNA2D1 | CRB2 | FGFR3 | HCN3 | KCNH1 | MED25 | ORAI1 | RBFOX3 | STT3B |
| ACY1 | CACNA2D3 | CREBBP | FKTN | HCN4 | KCNH2 | METTL23 | OTC | RBM10 | STYXL1 |
| ADD3 | CACNA2D4 | CRH | FLG | HEG1 | KCNH3 | MID2 | OTX2 | RBP4 | SUCLA2 |
| ADK | CACNB1 | CRHR1 | FLT4 | HEPACAM | KCNH4 | MINK1 | PAK3 | RBPJ | SUCO |
| ADNP | CACNB2 | CRLF1 | FMC1 | HERC2 | KCNH6 | MKS1 | PARN | RBSN | SV2A |
| AFG3L2 | CACNB3 | CRYAB | FMN2 | HESX1 | KCNH7 | MLC1 | PC | RD3 | SVIL |
| AGL | CACNG1 | CSF1R | FMR1 | HGSNAT | KCNH8 | MLLT3 | PCDH15 | RFT1 | SYN2 |
| AGO4 | CACNG2 | CSNK1G1 | FOXRED1 | HIBCH | KCNIP1 | MLYCD | PCDH7 | RFX3 | SYP |
| AGPS | CACNG3 | CSPP1 | FSTL5 | HIP1 | KCNJ1 | MMAA | PCDHB13 | RMND1 | SYT14 |
| AGTR2 | CACNG4 | CTC1 | FTL | HIVEP2 | KCNJ11 | MMADHC | PCDHB4 | RNASET2 | SYT2 |
| AHI1 | CACNG5 | CTH | FTO | HMGCS2 | KCNJ12 | MOGS | PDE10A | RNF213 | TAF1 |
| AIFM1 | CACNG6 | CTNND2 | FTSJ1 | HNF1B | KCNJ14 | MPC1 | PDSS2 | RPL10 | TANGO2 |
| AIMP1 | CACNG7 | CTSF | GABBR1 | HNRNPH1 | KCNJ15 | MPDU1 | PDYN | RRM2B | TAP1 |
| AKT2 | CACNG8 | CUL4B | GABRA4 | HS2ST1 | KCNJ16 | MPP7 | PET100 | RTN4R | TBCK |
| AKT3 | CALM2 | CUX1 | GABRA6 | HSD17B10 | KCNJ2 | MRAP | PEX13 | RUBCN | TBL1XR1 |
| ALDH18A1 | CALN1 | CXCR4 | GABRE | HSD17B4 | KCNJ3 | MRI1 | PEX14 | RYR1 | TBP |
| ALDH4A1 | CAMTA1 | CYB5R3 | GABRG1 | HSPD1 | KCNJ4 | MRPS22 | PEX19 | RYR2 | TDP1 |
| ALG1 | CARD9 | CYP26C1 | GABRG3 | HSPG2 | KCNJ5 | MSC | PEX2 | RYR3 | TDP2 |
| ALG11 | CARS2 | CYP27B1 | GABRP | HTR1A | KCNJ6 | MSX2 | PEX3 | SACS | TECPR2 |
| ALG12 | CAV3 | D2HGDH | GABRQ | HTR1B | KCNJ8 | MT-ATP6 | PEX5 | SATB2 | TELO2 |
| ALG2 | CBL | DARS2 | GABRR1 | HTR1D | KCNJ9 | MT-ATP8 | PEX7 | SCN10A | TENM2 |
| ALG3 | CC2D2A | DENND2B | GABRR2 | HTR1E | KCNK1 | MT-CO3 | PGAP1 | SCN11A | THAP1 |
| ALG6 | CD46 | DGKD | GAD1 | HTR1F | KCNK10 | MT-CYB | PGAP2 | SCN2B | THOC2 |
| ALG9 | CD59 | DHCR24 | GALC | HTR2A | KCNK12 | MTFMT | PGAP3 | SCN3B | THRB |
| AMACR | CELF4 | DHFR | GAS2L2 | HTR2C | KCNK13 | MTHFR | PGK1 | SCN4A | TICAM1 |
| AMER1 | CELSR3 | DHTKD1 | GATA6 | HTR3A | KCNK15 | MTMR10 | PGM3 | SCN4B | TK2 |
| ANK2 | CENPW | DIP2C | GATAD2B | HTR3B | KCNK16 | MTMR11 | PHF6 | SCN5A | TMEM139 |
| ANK3 | CEP164 | DKC1 | GBE1 | HTR3C | KCNK17 | MT-ND4 | PHF8 | SCN7A | TMEM70 |
| ANKH | CEP290 | DLD | GCH1 | HTR3D | KCNK2 | MT-ND5 | PHGDH | SCNM1 | TMLHE |

| | | | | | | | | | |
|----------|----------|---------|--------|----------|--------|---------|---------|----------|----------|
| ANO10 | CERT1 | DLG2 | GCK | HTR3E | KCNK3 | MTO1 | PHOX2B | SCO2 | TNFAIP6 |
| ANOS1 | CHD1L | DLG3 | GDI1 | HTR4 | KCNK4 | MTR | PIGG | SDHD | TPK1 |
| AP1S2 | CHD3 | DMBX1 | GFAP | HTR5A | KCNK5 | MYH14 | PIGL | SEC24D | TRAF3 |
| AP3D1 | CHD4 | DMD | GFM1 | HTR6 | KCNK6 | MYH6 | PIGM | SEMA5B | TRAPPC11 |
| AP4B1 | CHD8 | DNAJC6 | GFRA1 | HTR7 | KCNK7 | MYO9B | PIGO | SETBP1 | TRAPPC9 |
| AP4E1 | CHKB | DNM1L | GIGYF2 | HTT | KCNK9 | MYOCD | PIGV | SETD2 | TRIM8 |
| AP4S1 | CHL1 | DNM3 | GIPC1 | HUWE1 | KCNMB2 | MYT1L | PIGW | SETD5 | TRMT44 |
| APOE | CHN1 | DNMT1 | GJC2 | IDH2 | KCNMB3 | NAA10 | PIGY | SEZ6 | TRPM6 |
| APP | CHRD1 | DNMT3A | GJD2 | IDS | KCNMB4 | NADK2 | PIK3AP1 | SHANK1 | TSNARE1 |
| ARG1 | CHRFAM7A | DOCK6 | GK | IFIH1 | KCNN1 | NALCN | PLA2G6 | SHANK3 | TSPAN7 |
| ARHGDI | CHRNA1 | DOCK8 | GLI2 | IFNAR2 | KCNN2 | NANS | PLP1 | SHROOM4 | TSPYL4 |
| ARID1A | CHRNA10 | DOLK | GLRA1 | IGSF8 | KCNN3 | NAPB | PLXNB2 | SKI | TTN |
| ARID1B | CHRNA3 | DPAGT1 | GLRA2 | IL10 | KCNN4 | NARS2 | PMP22 | SLC16A1 | TTR |
| ARNT2 | CHRNA5 | DPM1 | GLRA3 | IL1B | KCNQ1 | NAT8L | PNPLA8 | SLC17A5 | TUBA3E |
| ARSA | CHRNA6 | DPM2 | GLRB | IL1RAPL1 | KCNQ4 | NCKAP5 | PNPT1 | SLC1A1 | TUBB4A |
| ASNS | CHRNA9 | DPYD | GLUL | IL1RN | KCNRG | NDN | POGZ | SLC1A3 | TXN2 |
| ASXL1 | CHRN1 | DPYS | GLYCTK | IL27RA | KCNS1 | NDP | POLR3B | SLC1A4 | UBE2A |
| ATIC | CHRN3 | DRD1 | GMEB2 | IL6 | KCNS2 | NDUFA1 | POMC | SLC25A1 | UBR5 |
| ATP1A3 | CHRN4 | DRD2 | GMPPB | INO80 | KCNS3 | NDUFA2 | POMGNT1 | SLC25A15 | UNC80 |
| ATP5F1A | CHRNA | DRD3 | GNA11 | INPP4A | KCNV1 | NDUF3F3 | POMK | SLC25A2 | UPB1 |
| ATP6AP1 | CHRNA | DRD4 | GNAQ | INSR | KCNV2 | NDUF3F6 | POMT1 | SLC25A20 | UQCC2 |
| ATP6AP2 | CHRNA | DRD5 | GNB1 | IRF3 | KDM5C | NDUFB11 | POMT2 | SLC26A1 | USP9X |
| ATP6V0A2 | CLCN1 | DYNC1H1 | GNPAT | ISG15 | KDM6A | NDUFS4 | PPP1R3C | SLC30A3 | VARS2 |
| ATP6V0C | CLCN3 | DYRK1A | GPHN | ITGB1BP1 | KIF1A | NDUFS8 | PPP2R1A | SLC33A1 | VPS11 |
| ATPAF2 | CLCN5 | EARS2 | GPSM2 | ITPR1 | KIF3C | NDUFV1 | PPP2R5D | SLC39A8 | VPS35 |
| ATXN10 | CLCN6 | EBP | GPT2 | IVNS1ABP | KIF4A | NEB | PQBP1 | SLC46A1 | VPS35L |
| ATXN2 | CLCN7 | ECM1 | GPX4 | JAM3 | KMT2A | NEDD4 | PRAG1 | SLC4A10 | WDR19 |
| AUH | CLCNKA | EFTUD2 | GRIA1 | JRK | KPNA7 | NELL1 | PRKDC | SLC4A3 | XK |
| AUTS2 | CLCNKB | EGF | GRIA2 | KARS1 | KPTN | NEU1 | PRKN | SLC6A3 | XPNPEP3 |
| AVPR2 | CLIC2 | EHMT1 | GRID1 | KAT6A | KRAS | NEXMIF | PRKX | SLC7A11 | YAP1 |
| B3GNT2 | CLPB | EIF2B1 | GRID2 | KCNA10 | KRIT1 | NF1 | PRNP | SLC7A6OS | YWHAE |
| B4GAT1 | CLPP | EIF3E | GRIK1 | KCNA3 | L1CAM | NGLY1 | PRODH | SLC8A1 | ZBTB18 |
| BCAP31 | CLSTN1 | ELMO1 | GRIK2 | KCNA4 | LAMA2 | NHS | PROK2 | SLC9A1 | ZC4H2 |
| BCKDHA | COA8 | ELOVL4 | GRIK3 | KCNA5 | LARS2 | NID1 | PROS1 | SLC9A9 | ZDHHC15 |
| BCKDK | COG4 | EPG5 | GRIK4 | KCNA6 | LMAN2L | NIN | PRRC2B | SLCO1B7 | ZFP57 |
| BCL10 | COG6 | EPHA5 | GRIK5 | KCNA7 | LRFN5 | NIPA1 | PSAP | SMARCA2 | ZFYVE26 |
| BCS1L | COG7 | EPHB2 | GRIN2C | KCNAB1 | LRP2 | NIPA2 | PSAT1 | SMARCA4 | ZMYND11 |
| BDNF | COG8 | ERCC6 | GRIN3A | KCNAB2 | LRPPRC | NKAIN3 | PSEN1 | SMARCB1 | ZNF12 |
| BEX3 | COL18A1 | ERLIN2 | GRINA | KCNAB3 | LRRK2 | NLGN1 | PSEN2 | SMARCE1 | ZNF44 |
| BMP4 | COL2A1 | ETFDH | GRIP1 | KCNB2 | MAF | NOD2 | PSMB8 | SMG9 | ZSWIM6 |
| BMP5 | COL3A1 | ETHE1 | GRM1 | KCNC2 | MAGEL2 | NOL11 | PSMD1 | SNRPN | |
| BOLA3 | COL4A1 | EXOC6B | GRM2 | KCNC3 | MAN1B1 | NOL3 | PSPH | SNX14 | |

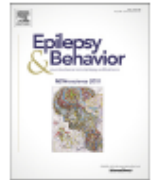
IV. Artigo de Revisão Publicado como Primeira Autora

Epilepsy & Behavior 121 (2021) 106428



Contents lists available at ScienceDirect

Epilepsy & Behavior

journal homepage: www.elsevier.com/locate/yebeh

Review

New avenues in molecular genetics for the diagnosis and application of therapeutics to the epilepsies



Pedro H.M. Magalhães¹, Helena T. Moraes¹, Maria C.P. Athie, Rodrigo Secolin, Iscia Lopes-Cendes^{*}

^a Departments of Medical Genetics and Genomic Medicine, School of Medical Sciences, University of Campinas (UNICAMP), Campinas, SP, Brazil

^b Brazilian Institute of Neuroscience and Neurotechnology (BRANN), Campinas, SP, Brazil

ARTICLE INFO

Article history:

Received 5 May 2019

Revised 14 June 2019

Accepted 6 July 2019

Available online 7 August 2019

Keywords:

Genome-wide association studies

Complex inheritance

Developmental epileptic encephalopathies

ABSTRACT

Genetic epidemiology studies have shown that most epilepsies involve some genetic cause. In addition, twin studies have helped strengthen the hypothesis that in most patients with epilepsy, a complex inheritance is involved. More recently, with the development of high-density single-nucleotide polymorphism (SNP) microarrays and next-generation sequencing (NGS) technologies, the discovery of genes related to the epilepsies has accelerated tremendously. Especially, the use of whole exome sequencing (WES) has had a considerable impact on the identification of rare genetic variants with large effect sizes, including inherited or *de novo* mutations in severe forms of childhood epilepsies. The identification of pathogenic variants in patients with these childhood epilepsies provides many benefits for patients and families, such as the confirmation of the genetic nature of the diseases. This process will allow for better genetic counseling, more accurate therapy decisions, and a significant positive emotional impact. However, to study the genetic component of the more common forms of epilepsy, the use of high-density SNP arrays in genome-wide association studies (GWAS) seems to be the strategy of choice. As such, researchers can identify loci containing genetic variants associated with the common forms of epilepsy. The knowledge generated over the past two decades about the effects of the mutations that cause the monogenic epilepsy is tremendous; however, the scientific community is just starting to apply this information in order to generate better target treatments.

© 2019 Elsevier Inc. All rights reserved.

1. Genetic predisposition to the epilepsies: from monogenic to complex inheritance

Remarkably successful efforts in identifying genes involved in the monogenic forms of epilepsy have led to the misconception that only in these rare forms of the disease does genetics have an important role [1,2]. Genetic epidemiology studies have shown that 70% to 80% of the epilepsies have the involvement of some genetic cause while only 20% to 30% are related to other etiologies [3–6]. In addition, twin studies have helped strengthen the hypothesis that, in most patients with epilepsy, a complex inheritance is involved [7–11].

The *CHRNA4* gene, which encodes a nicotinic acetylcholine receptor, was the first gene related to a monogenic form of epilepsy to be identified in 1995 [12,13]. The main approaches used before the advent of high throughput deoxyribonucleic acid (DNA) sequencing and high-density genome data were based on a traditional linkage analysis of

large families that segregated the disease following monogenic inheritance [14–18]. Based on these studies, many genes were found to be related to different forms of epilepsy in the early 2000s, most of them encoding ion-channel subunits [15,16]. However, despite the evidence of high heritability [19], these results were poorly replicated in the scenario in which common genetic variants contributed to the development of the more common epilepsies [19–22]. Nevertheless, the challenge to identify genetic variants associated with complex inheritance epilepsies resides in the large number of genes and their low effect size, which means that variants themselves just slightly raise the risk of developing epilepsy [16,23,24]. To address this issue better, it is also important to highlight the necessity to carry out multicenter studies to advance the field of complex epilepsies.

More recently, with the development of high-density single-nucleotide polymorphism (SNP) microarrays and next-generation sequencing (NGS) technologies, the discovery of genes related to the epilepsies has accelerated tremendously [20]. In particular, the use of whole exome sequencing (WES) has had a considerable impact on the identification of rare genetic variants with a rather large effect size, including inherited or *de novo* mutations in the severe forms of childhood epilepsy [15]. However, WES is still underpowered in the search for common variants with low effect size [16]. Therefore, to study the

^{*} Corresponding author at: Department of Medical Genetics and Genomic Medicine, School of Medical Sciences, University of Campinas – UNICAMP, Tessália Vieira de Camargo, 126, Cidade Universitária Zeferino Vaz, Campinas, SP 13083-887, Brazil.

E-mail address: icendes@unicamp.br (I.Lopes-Cendes).

¹ Both authors contributed equally to this publication.



V. Artigo Publicado como Coautora



Genetics and Molecular Biology, 43, 4, e20200047 (2020)
Copyright © Sociedade Brasileira de Genética.
DOI: <https://doi.org/10.1590/1678-4685-GMB-2020-0047>

Research Article
Genomics and Bioinformatics

The impact of post-alignment processing procedures on whole-exome sequencing data

Murilo Guimarães Borges^{1,2,3} , Helena Tadiello de Moraes^{1,2} ,
Cristiane de Souza Rocha^{1,2} and Iscia Lopes-Cendes^{1,2}

¹*Universidade Estadual de Campinas (UNICAMP), Faculdade de Ciências Médicas, Departamento de Genética Médica e Medicina Genômica, Campinas, SP, Brazil.*

²*Instituto Brasileiro de Neurociência e Neurotecnologia (BRAINN), Campinas, SP, Brazil.*

³*Universidade Estadual de Campinas (UNICAMP), Centro de Engenharia Biomédica. Campinas, SP, Brazil.*

Abstract

The use of post-alignment procedures has been suggested to prevent the identification of false-positives in massive DNA sequencing data. Insertions and deletions are most likely to be misinterpreted by variant calling algorithms. Using known genetic variants as references for post-processing pipelines can minimize mismatches. They allow reads to be correctly realigned and recalibrated, resulting in more parsimonious variant calling. In this work, we aim to investigate the impact of using different sets of common variants as references to facilitate variant calling from whole-exome sequencing data. We selected reference variants from common insertions and deletions available within the 1K Genomes project data and from databases from the Latin American Database of Genetic Variation (LatinGen). We used the Genome Analysis Toolkit to perform post-processing procedures like local realignment, quality recalibration procedures, and variant calling in whole exome samples. We identified an increased number of variants from the call set for all groups when no post-processing procedure was performed. We found that there was a higher concordance rate between variants called using 1K Genomes and LatinGen. Therefore, we believe that the increased number of rare variants identified in the analysis without realignment or quality recalibration indicated that they were likely false-positives.

Keywords: Sequence alignment, quality recalibration, variant discovery, BIPMed, LatinGen.

Received: February 26, 2020; Accepted: September 18, 2020.

Introduction

Advances in sequencing methods have made it possible to interrogate the genome in its most basic components at an affordable price and in a timely manner (Goodwin *et al.*, 2016). A single sequencing reaction generates reads that, after processing, make it possible to compare the resulting assembly against a given reference genome (Li and Durbin, 2009; DePristo *et al.*, 2011). Based on these achievements, the rapid adoption of such techniques are being directly applied to medicine, inaugurating the new era of genome medicine (Karczewski, 2013). Nonetheless, understanding the pathogenicity of a given variant is not a straightforward task and it demands prior knowledge or biological and *in-silico* validation (Thusberg *et al.*, 2011). In this perspective, databases containing variants previously related to disease have an unquestionable role while linking phenotypes to genotypes (Landrum *et al.*, 2018).

The common variants from a given population are also relevant to clinical diagnosis, while evidencing that normal alterations are not necessarily linked to disease (International HapMap Consortium, 2010; 1000 Genomes Project Consortium, 2015). Many such projects have been implemented by initiatives across the globe, providing a

much richer picture of human variation across regions and/or populations (Haga, 2017; Wijmenga and Zernakova, 2018; Stark *et al.*, 2019). In Latin America, one of the first initiatives was the Brazilian Initiative on Precision Medicine (BIPMed), which expanded into additional projects in the Latin American Database of Genomic Variation (LatinGen), aiming to assist with genomic data sharing in Latin America. Currently, despite being an initiative that encompasses all of Latin America, only two databases are contributing to genetic variation from the reference population (or healthy individuals), from the Brazilian population. Thus, LatinGen materializes what many studies reinforce, the need for a better understanding of the admixture in Latin American populations, as well as other underrepresented populations, in large scale sequencing studies (Ruiz-Linares *et al.*, 2014; Petrovski and Goldstein 2016; Popejoy and Fullerton, 2016; van Rooij *et al.*, 2017).

The use of post-alignment bioinformatics procedures has been suggested to reduce false-positive discovery rates in massive DNA sequencing data (McKenna *et al.*, 2010). Insertions and deletions are most likely to be misinterpreted by the alignment algorithms, which may produce several false single-nucleotide variants in the call-set. The use of high-quality, commonly-known variants tends to minimize such mismatching and allows reads to be correctly realigned and recalibrated. This procedure tends to increase the number of true genetic variants identified (Vo and Phan, 2018). In this work, we aim to investigate the impact of using different

Send correspondence to Iscia Lopes-Cendes. Universidade Estadual de Campinas, Faculdade de Ciências Médicas, Departamento de Genética Médica e Medicina Genômica, Campinas, SP, Brazil. E-mail: icendes@unicamp.br.

VI. Lista de Pacientes

Tabela 16: Lista de todos os pacientes analisados no estudo. ID - número de identificação do paciente; DG - Estado do diagnóstico genético do paciente.

| ID | Sexo | Idade coleta | Idade 1ª crise | Local de Coleta | Hipótese Diagnóstica | WES | CMA | DG |
|----|------|--------------|----------------|--------------------|--|-----|-----|--------------|
| 1 | M | 240 | 30 | Campinas - Brasil | Síndrome de Doose vs. Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 2 | M | - | 30 | Campinas - Brasil | Síndrome de Doose | OK | OK | Validar |
| 4 | M | - | 8 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 6 | M | - | 43 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 7 | F | - | 36 | Campinas - Brasil | Síndrome de Doose | OK | - | Inconclusivo |
| 9 | M | - | 36 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 12 | F | 84 | 8 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 13 | F | 120 | 72 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 15 | M | - | 3 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 16 | M | - | 15 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 17 | F | - | 4 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 18 | M | - | 36 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 19 | M | - | 8 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 20 | M | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 21 | F | 216 | 48 | Campinas - Brasil | Síndrome de Doose | OK | OK | OK |
| 22 | F | 168 | 3 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 24 | F | - | 4 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 25 | M | - | 30 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 26 | F | 312 | 35 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 27 | M | 168 | 36 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 28 | F | - | 7 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 29 | F | - | 3 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 30 | M | 168 | 30 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 31 | F | 192 | 7 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 32 | M | - | 18 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 33 | F | - | - | São Paulo - Brasil | Síndrome de Dravet | OK | OK | OK |
| 34 | F | 84 | 6 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 35 | F | - | 5 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 36 | F | 180 | 23 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 37 | F | - | 18 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 38 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 39 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 40 | F | - | 5 | São Paulo - Brasil | Síndrome de Dravet | OK | OK | OK |
| 41 | M | - | 4 | São Paulo - Brasil | Síndrome de Dravet | OK | OK | OK |
| 42 | M | - | 11 | Campinas - Brasil | Síndrome de Dravet | - | OK | Inconclusivo |
| 43 | M | - | 36 | São Paulo - Brasil | Síndrome de Doose | OK | OK | OK |
| 44 | M | - | 24 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 45 | F | - | 6 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |

| | | | | | | | | |
|----|---|-----|-----|-----------------------|----------------------------|----|----|--------------|
| 46 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 47 | M | - | 56 | São Paulo - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 48 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 49 | M | 204 | 5 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 50 | M | 84 | - | São Paulo - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 51 | M | 84 | - | São Paulo - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 52 | F | - | 6 | Porto Alegre - Brasil | Síndrome de Dravet | OK | OK | OK |
| 53 | F | - | 8 | Porto Alegre - Brasil | Síndrome de Dravet | OK | OK | OK |
| 54 | M | 204 | 2 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 55 | M | 204 | 8 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 56 | M | 36 | 12 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 57 | M | 444 | 46 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 58 | F | 156 | 24 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 59 | F | 60 | 11 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 60 | M | 204 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 61 | M | 84 | 48 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 62 | M | 108 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 63 | M | 48 | 24 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 64 | M | 192 | 108 | Campinas - Brasil | EED indeterminada | - | OK | Inconclusivo |
| 65 | F | 24 | 15 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 66 | M | 156 | 10 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 67 | M | 192 | 60 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 69 | F | 48 | 27 | Campinas - Brasil | EED indeterminada | - | OK | Inconclusivo |
| 70 | F | 36 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 71 | M | 180 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 72 | F | 72 | 35 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 73 | M | 12 | 2 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 74 | M | 108 | 14 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 75 | M | 12 | 8 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 76 | F | - | 9 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 77 | M | 96 | 0 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 78 | F | 2 | - | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 79 | F | 132 | 2 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 80 | M | 108 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 81 | F | 72 | 4 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 82 | F | 96 | 84 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 83 | M | 12 | 5 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 84 | M | 84 | 12 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 86 | M | 252 | 30 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 87 | M | 12 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 88 | M | 120 | 8 | Campinas - Brasil | Síndrome de West | OK | OK | Inconclusivo |
| 89 | M | 228 | 7 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 90 | M | - | 4 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 91 | M | 36 | 5 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 92 | M | 60 | 0 | Campinas - Brasil | EED indeterminada | OK | OK | OK |

| | | | | | | | | |
|-----|---|-----|-----|-------------------|---|----|----|--------------|
| 94 | M | 156 | - | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 95 | F | 216 | 24 | Salvador - Brasil | Síndrome de Dravet | OK | OK | OK |
| 96 | F | 132 | 3 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 97 | M | 48 | 18 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 98 | M | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 99 | M | - | - | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 100 | F | 120 | 41 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 102 | M | 156 | 24 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 103 | F | 168 | 7 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 104 | F | 180 | 60 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 105 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 106 | M | 204 | - | Curitiba - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 108 | M | 84 | 6 | Natal - Brasil | Síndrome de Dravet | OK | OK | OK |
| 109 | F | 36 | - | Curitiba - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 110 | M | - | - | Curitiba - Brasil | EED indeterminada | OK | OK | OK |
| 111 | M | 84 | 0 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 112 | M | - | - | Campinas - Brasil | EED indeterminada | OK | - | Validar |
| 113 | M | - | - | Campinas - Brasil | EED indeterminada | OK | - | OK |
| 114 | M | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 115 | F | 72 | 7 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 116 | M | - | - | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 125 | M | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 126 | M | 192 | 36 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 127 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 128 | M | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 129 | M | 84 | 15 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 130 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 131 | M | 48 | 46 | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 132 | M | 24 | 8 | Campinas - Brasil | Síndrome de West | OK | OK | OK |
| 133 | F | 120 | 10 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Validar |
| 134 | F | 72 | - | Curitiba - Brasil | EED indeterminada | OK | OK | OK |
| 136 | F | 204 | 132 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 137 | M | 36 | - | Curitiba - Brasil | EED indeterminada | OK | OK | OK |
| 138 | F | 24 | - | Curitiba - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 139 | M | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 140 | M | 192 | 61 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 141 | F | - | - | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 142 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 143 | M | 108 | 8 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 144 | M | 24 | 19 | Campinas - Brasil | Síndrome de West → Síndrome de Lennox-Gastaut | OK | OK | OK |
| 145 | F | 84 | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 149 | M | 180 | 132 | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 158 | F | 48 | 6 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 159 | M | 168 | 36 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |

| | | | | | | | | |
|-----|---|-----|-----|--------------------------|---|----|----|--------------|
| 160 | F | 12 | 1 | Campinas - Brasil | Síndrome de West | OK | OK | OK |
| 161 | F | 5 | - | Campinas - Brasil | Síndrome de West | OK | OK | OK |
| 167 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 168 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 169 | M | - | - | Campinas - Brasil | Síndrome de West | - | OK | Inconclusivo |
| 170 | M | - | - | Brasília - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 171 | M | 120 | 3 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 173 | M | 132 | 12 | Campinas - Brasil | Síndrome de Dravet | OK | OK | Inconclusivo |
| 174 | M | 192 | 17 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 175 | F | 168 | 24 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Validar |
| 176 | F | 228 | 168 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 177 | F | 168 | 60 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Validar |
| 178 | M | 96 | 36 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | OK |
| 179 | M | 180 | 24 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 180 | F | 144 | 36 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 181 | F | 60 | 14 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 182 | M | 204 | 24 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 183 | M | 120 | 36 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 184 | F | 96 | 12 | Buenos Aires - Argentina | Síndrome de Landau-Kleffner | OK | OK | Inconclusivo |
| 185 | F | 24 | - | Campinas - Brasil | EED indeterminada | OK | OK | OK |
| 186 | F | 72 | 5 | Natal - Brasil | EED indeterminada | OK | OK | Validar |
| 187 | M | 66 | 7 | Tegucigalpa - Honduras | EED indeterminada | OK | OK | OK |
| 188 | F | 60 | 10 | Tegucigalpa - Honduras | EED indeterminada | OK | OK | Inconclusivo |
| 189 | M | 108 | 10 | Tegucigalpa - Honduras | Síndrome de Dravet | - | OK | Inconclusivo |
| 190 | F | 156 | 120 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 191 | M | 192 | 36 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 192 | M | 84 | 7 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 193 | F | 156 | 6 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 194 | M | 120 | 60 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 195 | M | 192 | 84 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 196 | M | 144 | 72 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 197 | M | 108 | 36 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 198 | F | 108 | 60 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 199 | M | 120 | 10 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 200 | F | 96 | 7 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 201 | F | 108 | 2 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 202 | F | 108 | 72 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 203 | M | 60 | - | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 204 | F | 108 | 36 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 205 | F | 156 | 36 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | OK |
| 206 | F | 72 | 24 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 207 | M | 120 | 24 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Validar |
| 208 | M | 108 | 8 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 209 | M | 36 | 18 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |

| | | | | | | | | |
|-----|---|-----|-----|--------------------------|--|----|----|--------------|
| 210 | M | 156 | 60 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 211 | M | 96 | 48 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 212 | M | 48 | 2 | Buenos Aires - Argentina | Síndrome de Doose vs. Síndrome de Lennox-Gastaut | OK | OK | OK |
| 213 | F | - | - | Buenos Aires - Argentina | EED indeterminada | OK | OK | OK |
| 214 | F | 10 | 4 | Buenos Aires - Argentina | Síndrome de Doose vs. Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 215 | F | 10 | 2 | Buenos Aires - Argentina | Síndrome de Doose vs. Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 216 | M | 144 | 36 | Buenos Aires - Argentina | EED indeterminada | OK | OK | Inconclusivo |
| 217 | F | - | - | Lima - Peru | EED indeterminada | OK | OK | Inconclusivo |
| 218 | M | - | - | Lima - Peru | EED indeterminada | OK | OK | Inconclusivo |
| 219 | M | - | - | Lima - Peru | EED indeterminada | OK | OK | OK |
| 220 | F | 9 | 1 | Lima - Peru | Síndrome de West | OK | OK | Inconclusivo |
| 221 | F | 60 | 36 | Lima - Peru | EED indeterminada | OK | OK | OK |
| 222 | F | - | - | Lima - Peru | EED indeterminada | OK | OK | OK |
| 223 | F | - | - | Lima - Peru | EED indeterminada | OK | OK | OK |
| 224 | F | 4 | 2 | Lima - Peru | Síndrome de West | OK | OK | Inconclusivo |
| 225 | M | 192 | 3 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | OK |
| 226 | M | 84 | 18 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 227 | F | 96 | 22 | Buenos Aires - Argentina | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 228 | M | 48 | 4 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | OK |
| 229 | M | 24 | 11 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | OK |
| 230 | M | 60 | 36 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 231 | F | 228 | 15 | Buenos Aires - Argentina | Síndrome de Doose | OK | OK | Inconclusivo |
| 232 | F | 21 | 0 | Buenos Aires - Argentina | Síndrome de Dravet | OK | OK | Inconclusivo |
| 233 | F | 72 | 36 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 234 | F | 108 | 2 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 235 | M | 60 | 48 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 236 | F | 156 | 156 | Buenos Aires - Argentina | EED indeterminada | OK | OK | Inconclusivo |
| 237 | F | 132 | 120 | Buenos Aires - Argentina | Epilepsia com Estado de Mal Eléctrico do Sono | OK | OK | Inconclusivo |
| 238 | F | 24 | 0 | Buenos Aires - Argentina | Síndrome de Dravet | OK | OK | Inconclusivo |
| 239 | M | 24 | 7 | Buenos Aires - Argentina | Síndrome de West | OK | OK | OK |
| 240 | M | 48 | 2 | Buenos Aires - Argentina | EED indeterminada | OK | OK | OK |
| 241 | M | 24 | 11 | Buenos Aires - Argentina | Síndrome de Dravet | OK | OK | Inconclusivo |
| 242 | F | 36 | 5 | Buenos Aires - Argentina | Síndrome de Dravet | OK | OK | Inconclusivo |
| 243 | M | 27 | 9 | Buenos Aires - Argentina | EED indeterminada | OK | OK | OK |
| 244 | F | 7 | 0 | Campinas - Brasil | EED indeterminada | OK | OK | Validar |
| 245 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 246 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 247 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 248 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 249 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 250 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 251 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 252 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |

| | | | | | | | | |
|-----|---|-----|-----|----------------------|----------------------------|----|----|--------------|
| 253 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 254 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 255 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 256 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 257 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | OK |
| 258 | M | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Inconclusivo |
| 259 | F | - | - | Valparaíso - Chile | EED indeterminada | OK | OK | Validar |
| 260 | M | 102 | 32 | Lima - Peru | Síndrome de Doose | OK | OK | Inconclusivo |
| 261 | M | 7 | 0 | Lima - Peru | Síndrome de West | OK | OK | OK |
| 262 | M | 10 | 6 | Lima - Peru | Síndrome de West | OK | OK | OK |
| 263 | F | 25 | 8 | Lima - Peru | Síndrome de West | OK | OK | Inconclusivo |
| 264 | F | - | - | Lima - Peru | EED indeterminada | OK | OK | OK |
| 265 | M | 9 | 4 | Lima - Peru | Síndrome de West | OK | OK | Inconclusivo |
| 266 | F | 21 | 3 | Montevideo - Uruguay | Síndrome de West | OK | OK | Inconclusivo |
| 267 | M | 14 | 8 | Montevideo - Uruguay | Síndrome de West | OK | OK | Validar |
| 268 | F | - | - | Montevideo - Uruguay | EED indeterminada | OK | OK | Inconclusivo |
| 269 | F | 31 | 5 | Montevideo - Uruguay | Síndrome de West | OK | OK | Validar |
| 270 | M | 10 | 0 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 271 | M | 10 | 48 | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 272 | F | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 273 | F | - | 156 | Campinas - Brasil | Síndrome de Lennox-Gastaut | OK | OK | Inconclusivo |
| 307 | M | 84 | 5 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |
| 308 | M | - | - | Campinas - Brasil | Síndrome de Doose | OK | OK | Inconclusivo |
| 309 | M | - | - | Campinas - Brasil | EED indeterminada | OK | OK | Inconclusivo |
| 310 | M | 60 | 4 | Campinas - Brasil | Síndrome de Dravet | OK | OK | OK |