Silvia Souza da Costa

Investigação por sequenciamento do exoma completo
da etiologia genética da deficiência intelectual familiar

Investigating the genetic etiology of familial intellectual
disability by whole exome sequencing

São Paulo
2018

Silvia Souza da Costa

Investigação por sequenciamento do exoma completo
da etiologia genética da deficiência intelectual familiar

Investigating the genetic etiology of familial intellectual
disability by whole exome sequencing

Tese apresentada ao Instituto de
Biociências da Universidade de
São Paulo, para a obtenção de
Título de Doutor em Ciências, na
Área de Biologia (Genética)

Orientadora: Profa. Dra. Carla Rosenberg

São Paulo

2018

Ficha Catalográfica

Costa, Silvia Souza da

Investigação por sequenciamento do exoma completo da etiologia genética da deficiência intelectual familiar.

136 páginas

Tese (Doutorado) - Instituto de Biociências da Universidade de São Paulo. Departamento de Genética e Biologia Evolutiva

1. Deficiência intelectual 2. Sequenciamento de nova geração 3. Sequenciamento do exoma completo

I. Universidade de São Paulo. Instituto de Biociências. Departamento de Genética e Biologia Evolutiva

Comissão Julgadora:

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Prof(a). Dr(a).

Profa. Dra. Carla Rosenberg
Orientadora

Este trabalho foi realizado com os auxílios financeiros da FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) e da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), concedidos à orientadora (projeto CEPID 2013/08028-1) e ao departamento (programa CAPES-PROEX), respectivamente.

Aos meus filhos, Manuela e João
À minha amiga, Ana C. V. Krepischi

"There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know."

Donald Rumsfeld

(Feb. 12, 2002, U.S. Department of Defense)

Agradecimentos

À Prof. Dra. Carla Rosenberg pela amizade e paciência. Agradeço pela orientação deste projeto e pela confiança depositada em mim todos esses anos.

À Profa. Dra. Ana Cristina Victorino Krepschi pela generosidade e co-orientação deste trabalho. Acredito que sem seu apoio eu não teria coragem para dar início a este projeto.

Aos pacientes e seus familiares, sem os quais esse trabalho não poderia ter sido realizado.

Ao Departamento de Genética e Biologia Evolutiva do IB-USP, pela possibilidade de realização deste trabalho.

Ao meu marido, Ju (Wilson) pelo amor, pela cumplicidade e pela família maravilhosa que temos.

Aos meus amores, Manuela e João, por me ensinarem todos os dias o valor da vida.

Aos Drs. Angela M. Vianna-Morgante, Paulo Otto, Débora Bertola, Rachel Sayuri Honjo Kawahira, Gisele Santos de Oliveira, Fernanda S. Jehee, Juliana Gurgel Giannetti, Paula Frassinetti Vasconcelos de Medeiros, Fernando Kok e Hélio van der Linden pela avaliação clínica e encaminhamento dos pacientes.

Aos colaboradores do Laboratório Nacional de Biociências – LNBio, Dr. Kleber Gomes Franchini, Dra. Juliana Ferreira de Oliveira e Dra. Mariana Maschietto (atualmente como pesquisadora no Centro Infantil Boldrini), pelo desenvolvimento da pesquisa com variante do gene *UBE2A*.

Aos colaboradores do Centro Internacional de Pesquisa (CIPE) do Hospital AC Camargo: Drs. Israel Tojal da Silva e Renan Valieris pela análise bioinformática e Dras. Dirce Maria Carraro, Giovana Tardin Torrezan e Ana Carolina Kerekes Miguez pelo auxílio técnico na confirmação de mosaicismo pela técnica de ddPCR.

À equipe do Centro de Pesquisas sobre o Genoma Humano e Células Tronco (CEGH-CEL), pelo suporte técnico. Especialmente à Naila e Meire pela amizade.

À Profa. Dra. Regina C. Mingroni Netto pelo uso de seu laboratório.

Às queridas amigas Maraisa e Ligia pelo apoio e amizade de tantos anos.

À Lilian por todo apoio, pela amizade e pela formatação desta tese.

Aos colegas de laboratório e alunos: Mara, Paulo, Fátima, Darine, Francine, Juliana Carnavalli, Juliana Sobral, Laura, Talita, Anne, Maria, Camila, Alex, André, Vinícius, Willian e Luiz Gustavo, pela ajuda e apoio.

Aos colegas dos demais laboratórios do Instituto de Biociências pelo apoio e conhecimento técnico.

Finalmente, agradeço a todos aqueles que de alguma forma ajudaram para que eu chegasse onde estou e contribuíssem na realização deste trabalho.

Índice geral

Resumo	1
Abstract	2
I. Introdução	3
I.1. Deficiência Intelectual	3
I.2. DI com padrões de herança mendelianos	7
I.2.1. DI ligada ao cromossomo X	8
I.2.2. DI autossômica dominante	10
I.2.3. DI autossômica recessiva	11
II. Justificativa	14
III. Objetivos	15
IV. Famílias	16
V. Métodos	19
V.1. Obtenção das amostras de DNA genômico	19
V.2. Sequenciamento do exoma completo (WES)	19
V.3. Análise de bioinformática dos dados do WES	21
V.4. Anotação, filtragem e seleção de variantes	22
V.5. Sequenciamento Sanger	24
V.6. Análise do padrão de inativação do cromossomo X	25
V.7. Droplet digital PCR (ddPCR)	25
VI. Resultados e discussão dos casos	27
VI.1. Resumo dos resultados	27
VI.2. Identificação de variantes patogênicas não relacionada a DI	29
VI.3. Apresentação das famílias portadora de variantes potencialmente patogênicas para a DI	31
VI.3.1. Família 1 – O espectro fenotípico de mutações em <i>UBE2A</i>	31
VI.3.2. Família 2 – Pacientes com fenótipo intermediário entre as síndromes de Opitz-Kaveggia (FG) e de Lujan-Fryns	43

VI.3.3. Família 3 – Novas variantes patogênicas no gene <i>CDK5RAP2</i> associadas à microcefalia primária autossômica recessiva 3 (MCPH3)	57
VI.3.4. Família 6 – Nova variante provavelmente patogênica no gene <i>SLC6A8</i> que causa DI recessiva ligada ao cromossomo X	65
VI.3.5. Família 8 – Mosaicismo em DI com padrão de herança dominante	75
VI.3.6. Família 9 – <i>ARHGAP4</i> : um novo gene de deficiência intelectual não síndrômica?	83
VI.3.7. Família 11 – Região 3' não-traduzida (3' UTR) de <i>PHF6</i> : uma variante a ser considerada para o fenótipo?	89
VI.3.8. Família 13 – Identificação de uma nova variante no gene <i>MED13L</i>	96
VI.3.9. Família 25 – Um caso familiar de DI segregando com padrão dominante associada a uma variante no gene <i>MBD5</i>	104
VII. Discussão geral	111
VIII. Conclusão	118
IX. Referências bibliográficas	119

Índice de figuras

- Figura 1.** Identificação de genes associados com a DI ao longo dos anos. O total de genes identificados em 2015 (linha cinza) ultrapassava 700 genes. Esses genes estão distribuídos nos diferentes padrões de herança mendeliana: linha verde são os genes de herança ligada ao X, linha azul genes de herança autossômica dominante e linha roxa genes de herança autossômica recessiva (modificado de Vissers et al. 2016). 7
- Figura 2.** Distribuição de genes de DI no genoma humano. Esquema mostrando ideogramas dos cromossomos humanos com padrão de bandas G; as setas verdes indicam o mapeamento dos genes já associados à DI síndrômica e não síndrômica. A figura evidencia que genes de DI estão espalhados por todos os cromossomos, exceto no Y (Chiurazzi and Pirozzi 2016). 8
- Figura 3.** Heredograma da Família 24, na qual uma variante patogênica no gene *HNF1A* segrega com diabetes mellitus num padrão de herança autossômica dominante. 30
- Figura 4.** Sequenciamento Sanger em amostras de DNA genômico mostrando a deleção em heterozigose da variante g.121434373delT (c.1137delT), encontrada no exon 6 do gene *HNF1A* nos irmãos afetados (II.1 e II.2). Apenas a presença da sequência normal foi observada no controle. 30
- Figura 5.** Representação esquemática do processo de adição covalente de ubiquitina (Ub) a proteínas alvo. E1, enzima ativadora de ubiquitina; E2, enzima conjugadora de ubiquitina e E3, ligase de ubiquitina [baseado em (Suresh et al. 2016)]. 32
- Figura 6.** Heredograma da Família 1, na qual segrega uma variante no gene *UBE2A* com um padrão de herança recessiva ligada ao cromossomo X. 35
- Figura 7.** (A) Diagrama do gene *UBE2A*: o gene contém seis exons e a seta indica a localização da variante identificada nos pacientes II.1 e II.3, no exon 5 (c.277C>G:p.Gln93Glu). (B) Diagrama da proteína UBE2A que contém um domínio catalítico de conjugação de ubiquitina (UBCc) - a seta indica a localização da variante Q93E (<http://smart.embl-heidelberg.de/>). 37
- Figura 8.** Imagem da superfície da proteína UBE2A. O resíduo catalítico C88 está indicado (esferas em azul), assim como o resíduo Q93, no qual a variante missense foi identificada (esferas vermelhas) na família 1 (PyMOL by Schrödinger - <https://pymol.org>). 37
- Figura 9.** Sequenciamento Sanger em amostras de DNA genômico mostrando a variante missense c.277C>G encontrada no gene *UBE2A*. (A) presença da variante em hemizigose nos irmãos afetados (II.1 e II.3) e em heterozigose na mãe (I.2) e (B) presença da variante nas irmãs heterozigotas (II.2, II.4 e II.5). 38
- Figura 10.** Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (I.2), evidenciando desvio total de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de apenas um dos alelos (1), 39

mostrando inativação preferencial de um dos cromossomos X em sangue periférico (razão de inativação 95:5).

Figura 11. Sequenciamento Sanger em amostras de cDNA derivado de mRNA de sangue periférico mostrando a variante *missense* cDNA.451C>G no exon 5 do gene *UBE2A*. Acima, presença da variante nos irmãos afetados (II.1 e II.3); abaixo, ausência da variante no controle e na mãe dos afetados (I.2). 39

Figura 12. Dismorfismos faciais da síndrome de DI ligada ao cromossomo X do tipo Nascimento. A) Família com três homens afetados descrita por Nascimento e col., na qual a mutação c.382C>T (Q128X), que cria um códon de parada prematura, foi identificada como a causa da síndrome. (B) Quatro afetados de uma mesma família, portadores da variante *missense* c.67G>A (G23R) e (C) dois afetados de uma mesma genealogia portadores da variante *missense* c.236C>G (P79R). 42

Figura 13. Heredograma da Família 2, na qual uma variante provavelmente patogênica no gene *MED12* segrega com DI num padrão de herança ligada ao cromossomo X. 48

Figura 14. (A), (B) e (C) Características faciais do paciente 2a (II.4), que incluem hipertelorismo e estrabismo divergente, orelhas pequenas, hipoplasia malar, filtro curto e mandíbula proeminente. (D) Imagem de corpo inteiro do paciente. (E) e (F) mostram a clinodactilia de 50 dedo. 49

Figura 15. (A) e (B) Características faciais do paciente 2b (II.5), que incluem braquicefalia à direita, hipertelorismo e estrabismo convergente à esquerda, orelhas pequenas, filtro nasal curto, má-oclusão dentária e mandíbula proeminente. (C) Imagem de corpo inteiro do paciente. (D) e (E) Clinodactilia de 5º dedo e polegares largos. 50

Figura 16. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante *missense* c.2179G>A encontrada no exon 15 do gene *MED12*. (A) presença da variante em hemizigose nos irmãos afetados (II.4 e II.5) e em heterozigose na mãe portadora (I.2) e (B) presença da variante nas irmãs portadoras (II.1 e II.2) e detecção apenas do alelo normal na irmã II.3. 52

Figura 17. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados, evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de apenas um dos alelos (1), mostrando inativação preferencial de um cromossomo X em aproximadamente 70% das células de sangue periférico (razão de inativação 70:30). 52

Figura 18. Heredograma da Família 3, na qual duas variantes diferentes do tipo frameshift foram detectadas no gene *CDK5RAP2* em heterozigose composta, segregando com DI num padrão de herança autossômica recessiva. 61

Figura 19. Face do paciente 3a (I.1). 62

Figura 20. Face do paciente 3b (I.2). 62

Figura 21. Sequenciamento Sanger em amostras de DNA genômico mostrando as variantes identificadas em heterozigose composta no gene <i>CDK5RAP2</i> . (A) variante <i>frameshift</i> em heterozigose c.564_565dupGA no exon 7 do gene <i>CDK5RAP2</i> , presente nos irmãos (II.1 e II.2) e na mãe (I.2). (B) variante <i>frameshift</i> em heterozigose c.1246delG no exon 12 do gene <i>CDK5RAP2</i> , presente nos irmãos (II.1 e II.2) e no pai (I.1).	63
Figura 22. Heredograma da Família 6, na qual foi identificada uma variante <i>missense</i> no gene <i>SLC6A8</i> classificada como provavelmente patogênica, segregando com DI num padrão de herança ligada ao cromossomo X.	67
Figura 23. Face dos irmãos afetados. (A) paciente 6a (II.1), (B) paciente 6b (II.2) e (C) paciente 6c (II.3).	68
Figura 24. Imagem do arquivo BAM mostrando nas <i>reads</i> a posição da variante de nucleotídeo único na posição g.152954234G>C (c.205G>C) no exon 1 do gene <i>SLC6A8</i> visualizada no <i>software</i> VarSeq (Golden Helix). (A) I.2 (mãe) e (B) II.3 (afetada): observa-se a presença do nucleotídeo mutado (em amarelo) em heterozigose; a estrela laranja na sequência proteica indica a posição do aminoácido trocado (Ala69).	70
Figura 25. Imagem do arquivo BAM mostrando nas <i>reads</i> a posição da variante de nucleotídeo único na posição g.152954234G>C (c.205G>C) no exon 1 no gene <i>SLC6A8</i> visualizada no <i>software</i> VarSeq (Golden Helix). (C) II.1 (afetado) e (D) II.3 (afetado), observa-se a presença do nucleotídeo mutado (em amarelo) em hemizigose e a estrela laranja indica o aminoácido onde a troca ocorreu na proteína (Ala69).	71
Figura 26. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (A e B) e na filha afetada (C e D), evidenciando desvio parcial de inativação. (A e C) amplificação de dois alelos do loco <i>AR</i> , 2 e 3 na mãe e 1 e 2 na filha, mostrando heterozigose. (B e D) após o tratamento com a enzima sensível à metilação <i>HpaII</i> . Na mãe, inativação do alelo 2 do cromossomo X em aproximadamente 57% das células de sangue periférico (razão de inativação 57:43) e na filha, inativação do alelo 1 do cromossomo X em aproximadamente 43% das células de sangue periférico (razão de inativação 43:57).	72
Figura 27. Heredograma da Família 8, na qual foi detectada uma variante patogênica no gene <i>SETBP1</i> do tipo <i>frameshift</i> que segrega com DI num padrão de herança autossômica dominante.	78
Figura 28. Face da paciente 8a (II.1).	78
Figura 29. Face da paciente 8b (II.3).	79
Figura 30. Sequenciamento Sanger em amostras de DNA genômico mostrando a detecção da deleção em heterozigose da variante g.42529933delC (c.628delC), encontrada no exon 4 do gene <i>SETBP1</i> nas irmãs afetadas (II.1 e II.3) e apenas a presença da sequência normal nos pais (I.1 e I.2) e controle.	80

- Figura 31.** Detecção de mosaicismos para a variante c.628delC de *SETBP1* em células de sangue periférico da mãe das afetadas. Os gráficos representam os dados referentes às reações-controle e às reações para detecção das sequências-alvo (alelo selvagem e alelo com a variante). Os pontos pretos representam as reações com DNA sem amplificação (gotículas vazias), os pontos verdes, as reações de amplificação do alelo selvagem, os azuis, amplificação do alelo mutado, e os pontos laranjas são os duplo-positivos, que na mesma gotícula tinham os dois alelos (selvagem e mutado). (A) Paciente II.1, detecção de 49,2% do alelo mutado (pontos em azul) em amostra de DNA extraída de sangue periférico. (B) Mãe (I.2), detecção de 0,13% do alelo mutado (em azul) em amostra de DNA extraída de sangue periférico. (C) Pai (I.1), detecção apenas de alelos selvagens quanto à variante testada, em amostra de DNA extraída de sangue periférico. 81
- Figura 32.** (A) Características faciais de pacientes relatados na literatura com variantes de perda de função no gene *SETBP1* (Coe e col. 2014). (B) Características faciais das pacientes da família 8, portadoras da variante frameshift aqui identificada. 83
- Figura 33.** Heredograma da Família 9, na qual segrega uma variante no gene *ARHGAP4* num padrão compatível com herança recessiva ligada ao cromossomo X. 85
- Figura 34.** Face dos irmãos afetados. (A) paciente 9a (II.2) e (B) paciente 9b (II.3). 86
- Figura 35.** Sequenciamento Sanger em amostras de DNA genômico mostrando a variante missense c.1471C>T identificada no exon 11 do gene *ARHGAP4*. Presença da variante em hemizigose nos irmãos afetados (II.2 e II.3) e em heterozigose na mãe (I.2); ausência da variante alterada no controle masculino. 87
- Figura 36.** Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados, heterozigota quanto à variante no gene *ARHGAP4*, evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de um dos alelos (1), mostrando inativação preferencial de um dos cromossomos X em aproximadamente 78% das células de sangue periférico (razão de inativação 78:22). 87
- Figura 37.** Heredograma da Família 11; na qual segrega uma variante de significado indeterminado no gene *PHF6* em mãe clinicamente normal e em um casal de filhos, com padrão compatível com herança recessiva ligada ao X. 91
- Figura 38.** Sequenciamento Sanger em amostras de DNA genômico mostrando a variante de nucleotídeo único na posição g.133562204C>T encontrada na 3'UTR do gene *PHF6*. Presença da variante em hemizigose no irmão afetado (II.1) e em heterozigose na mãe (I.2). O pai (I.1) é portador do alelo normal. 93
- Figura 39.** Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (I.2), evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de um dos alelos (1), mostrando inativação preferencial de um cromossomo X em 68% das células de sangue periférico (razão de inativação 68:32). 94

Figura 40. Heredograma da Família 13, na qual uma variante provavelmente patogênica no gene <i>MED13L</i> segrega com DI num padrão de herança autossômica dominante.	100
Figura 41. Face do paciente 13a (II.2).	101
Figura 42. Face da paciente 13b (II.9).	101
Figura 43. Imagem do arquivo BAM mostrando nas <i>reads</i> a posição da variante g.116422099G>A no exon 20 do gene <i>MED13L</i> visualizada no <i>software</i> VarSeq (Golden Helix). (A) I.3 (afetado) e (B) II.9 (afetada): observa-se a presença da base mutada (em vermelho) em heterozigose; a estrela vermelha na sequência proteica indica a posição do aminoácido trocado (Gln1473).	102
Figura 44. (A) Exons codificadores do gene <i>MBD5</i> são mostrados em vermelho nas duas isoformas proteicas. (B) Organização da proteína MBD5. Domínio MBD (ligação a metil-CpG), segmento rico em prolina (P rich); domínio PWWP (prolina-triptofano-triptofano-prolina) e sinal de localização nuclear (NLS).	105
Figura 45. Heredograma da Família 25, na qual uma variante provavelmente patogênica no gene <i>MBD5</i> aparentemente segrega com DI e agressividade num padrão de herança autossômica dominante.	107
Figura 46. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante missense c.80G>A encontrada no gene <i>MBD5</i> . Presença da variante em heterozigose na mãe (II.2) e nos irmãos afetados (II.1, II.3 e II.4); variante ausente no controle masculino.	108

Índice de tabelas

Tabela 1. Caracterização das famílias com indivíduos afetados por DI incluídas neste estudo.	17
Tabela 2. Descrição das plataformas de captura e construção de bibliotecas genômicas para sequenciamento completo de exoma utilizadas.	21
Tabela 3. Variantes potencialmente causativas dos fenótipos em nove famílias com pacientes afetados por DI.	28
Tabela 4. Resumo do quadro clínico dos pacientes deste estudo e das famílias encontradas na literatura com síndrome de DI ligada ao X tipo Nascimento causada por mutação missense no gene <i>UBE2A</i> .	41
Tabela 5. Comparação das características clínicas das síndromes de Lujan e Opitz-Kaveggia.	46
Tabela 6. Comparação das características clínicas descritas em afetados pelas síndromes associadas a mutações no gene <i>MED12</i> (Lujan-Fryns, Opitz-Kaveggia e Ohdo) e nos irmãos afetados da Família 2.	55
Tabela 7. Genes conhecidos de microcefalia primária autossômica recessiva (MCPH) de acordo com a revisão de Zaqout e col. (2017).	59
Tabela 8. Frequência das características clínicas em pacientes com mutações ocasionando haploinsuficiência de <i>MED13L</i> .	99
Tabela 9. Número de famílias com variantes que potencialmente explicam a deficiência intelectual e respectivos padrões de herança.	113

Lista de abreviaturas e siglas

AAIDD	Associação Americana de Deficiência Intelectual e de Desenvolvimento (<i>American Association on Intellectual and Developmental Disabilities</i>)
AAP	Academia Americana de Pediatria (<i>American Academy of Pediatrics</i>)
ABraOM	Arquivo Brasileiro Online de Mutações
ACMG	Colégio Americano de Genética Médica e Genômica (<i>American College of Medical Genetics and Genomics</i>)
AR	Gene receptor de andrógenos (<i>Androgen Receptor</i>)
ASD	<i>Autism spectrum disorders</i>
BAM	<i>Binary Alignment Map</i>
BLAST	<i>Basic Local Alignment Search Tool</i>
BWA	<i>Burrows-Wheeler Aligner</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CCDS	<i>The consensus coding sequence project</i>
CEGH-CEL	Centro de Pesquisa do sobre o Genoma Humano e Células-Tronco
CNV	Variação no número de cópias (<i>Copy number variation</i>)
dbSNP	<i>Single Nucleotide Polymorphism Database</i>
DI	Deficiência intelectual
Die	Distância intercantal externa
Dii	Distância intercantal interna
DNA	Ácido desoxirribonucleico
DNPM	Desenvolvimento neuropsicomotor
dNTP	Desoxirribonucleotídeo fosfatado
ESP6300SI-V2	<i>Current Exome Variant Serve data</i>
EUA	Estados Unidos da América
ExAC	<i>Exome Aggregation Consortium</i>
FAPESP	Fundação de Amparo à pesquisa do Estado de São Paulo

FISH	Hibridização fluorescente in situ (<i>Fluorescence in situ hybridization</i>)
FMR1	Gene retardo mental X frágil 1 (<i>Fragile X Mental Retardation 1</i>)
FMRP	Proteína retardo mental X frágil (<i>Fragile X Mental Retardation Protein</i>)
IB-USP	Instituto de Biociências da Universidade de São Paulo
LoF	Perda de função (<i>Loss of Function</i>)
NCBI	<i>National Center for Biotechnology Information</i>
NGS	Sequenciamento de nova geração (<i>Next Generation Sequencing</i>)
NHLBI	<i>National Heart, Lung, and Blood Institute</i>
OMIM	<i>Online Mendelian Inheritance in Man</i>
PC	Perímetro cefálico
PCR	Reação em cadeia da polimerase
pH	Potencial hidrogeniônico
PhD-SNP	<i>Predictor of human Deleterious Single Nucleotide Polymorphisms</i>
PolyPhen	<i>Polymorphism Phenotyping</i>
POP7	Polímero de alta performance
QI	Quociente de inteligência
RefSeq	<i>The NCBI Reference Sequence</i>
RefSNP	SNP referência
RNA	Ácido ribonucleico
RNA _m	RNA mensageiro
RNM	Ressonância Magnética Nuclear
SAM	<i>Sequence Alignment Map</i>
SIFT	<i>Sorting Intolerant From Tolerant</i>
SNP	Polimorfismo de base única
SNV	Variante de nucleotídeo único ou de ponto (<i>Single Nucleotide Variant</i>)
SXF	Síndrome do cromossomo X frágil

UCSC	<i>University of California, Santa Cruz</i>
UTR	Região não traduzida (<i>untranslated region</i>)
USP	Universidade de São Paulo
VCF	<i>Variant Call Format</i>
VEP	<i>Variant Effect Predictor</i>
WES	Sequenciamento completo do exoma (<i>Whole exome sequencing</i>)
WGS	Sequenciamento completo do genoma (<i>Whole genome sequencing</i>)

Resumo

Deficiência intelectual (DI) é um distúrbio frequente do neurodesenvolvimento que afeta ~1% da população mundial e a causa genética é desconhecida em grande parte dos casos. O sequenciamento de exoma é uma ferramenta valiosa na elucidação da etiologia da DI e foi utilizada neste trabalho para investigar 25 famílias com dois ou mais indivíduos afetados. Nosso objetivo foi identificar novos genes ou variantes em genes já conhecidos que pudessem explicar a DI nessas famílias. Identificamos variantes que podem explicar a DI em nove das 25 famílias (36%); três variantes em genes já associados a DI de herança autossômica dominante (*SETBP1*, *MED13L* e *MBD5*), duas em um gene já associado a DI de herança autossômica recessiva (*CDK5RAP2* – heterozigose composta) e cinco em genes de herança ligada ao X (*UBE2A*, *MED12*, *SCL6A8*, *ARHGAP4* e *PHF6*). Nenhuma dessas variantes havia sido previamente descrita. O gene *ARHGAP4* aparece como um novo candidato e novos estudos serão necessários para estabelecer seu papel na DI. No gene *PHF6*, a variante identificada mapeia na região 3'UTR e potencialmente interfere com as interações de miRNA reguladores. A variante no gene *SETBP1*, que apresenta padrão de herança dominante, foi detectada em duas irmãs, indicando que um genitor seria portador obrigatório; essa variante foi detectada em 0,13% das células de sangue periférico da mãe com a utilização da técnica de PCR digital. Essa variante em *SETBP1*, juntamente com as variantes em *MED13L* e no gene *MBD5*, salientam a contribuição de mosaicismos gonadais assim como genitor também com quadro clínico na DI de herança autossômica dominante. A repetição da DI nas irmandades foi essencial para determinar a patogenicidade das variantes identificadas do tipo *missense*. Dois irmãos portadores de variante *missense* ainda não descrita no gene *UBE2A* apresentavam quadro clínico atipicamente leve, trazendo dúvidas sobre a patogenicidade dessa variante; o estudo funcional do gene demonstrou que a mutação de fato prejudicava a função da proteína. Embora estudos funcionais sejam o padrão-ouro para determinar a patogenicidade de uma variante, é difícil sua implementação na rotina diagnóstica.

Abstract

Intellectual deficiency (ID) is a common neurodevelopmental disorder affecting ~1% of the world population, and in which the genetic underlying condition is not determined in many cases. Whole exome sequencing is a valuable tool to elucidate ID etiology and was used in the present work to investigate 25 families with two or more affected individuals. Our main goal was to identify either new genes or variants in already known genes that might explain ID in these families. We detected variants in nine out of the 25 families (36%), with three variants in autosomal dominant ID genes (*SETBP1*, *MED13L* e *MBD5*), two in an autosomal recessive ID gene (compound heterozygote – *CDK5RAP2*) and five in X-linked ID genes (*UBE2A*, *MED12*, *SCL6A8*, *ARHGAP4* e *PHF6*). None of the variants had been previously described. *ARHGAP4* emerged as a new candidate gene and further studies are needed to establish its role in ID. In *PHF6*, the identified variant mapped to the 3'UTR region and potentially interferes with miRNA interactions. The variant in *SETBP1*, which segregates as an autosomal dominant, is present in two sisters, indicating that one parent is an obligate carrier; in fact, the variant was only identified in 0.13% of the mother's blood cells after using digital PCR. This case, together with the variants in *MED13L* and *MBD5*, highlight the contribution of gonadal mosaicism and affected parent in the autosomal dominant ID. The ID recurrence in the sibships was essential to determine the putative pathogenicity of missense variants. Two affected brothers carrying a novel missense variant in *UBE2A* exhibited an atypically mild phenotype, and raised questions regarding the pathogenicity of this variant; associated functional studies showed that the variant impairs gene function. Although functional studies are the gold standard to determine a variant pathogenicity, its implementation as a diagnostic routine is still difficult.

I. Introdução

I. Introdução

I.1. Deficiência Intelectual

Segundo a Associação Americana de Deficiência Intelectual e de Desenvolvimento (*American Association on Intellectual and Developmental Disabilities* - AAIDD), a deficiência intelectual (DI) é caracterizada por limitações significativas do indivíduo originadas antes de 18 anos de idade, tanto no funcionamento intelectual (raciocínio, aprendizagem, resolução de problemas) como no comportamento adaptativo, que abrange uma variedade de habilidades sociais e práticas diárias. É uma condição clínica complexa e altamente heterogênea, cuja prevalência foi estimada em aproximadamente 1% nos países desenvolvidos (Chiurazzi 2011; Maulik et al. 2011). No entanto, um diagnóstico formal de DI é feito apenas quando o teste de coeficiente de inteligência (QI) resulta em uma pontuação inferior a 70, valor de corte estabelecido pela Associação Americana de Psiquiatria em 1994 (Leonard and Wen 2002). A gravidade da DI é altamente variável e, convencionalmente, pode ser classificada de acordo com o QI dos indivíduos em: leve (QI entre 70 e 50), moderada (QI entre 50 e 35), grave (QI entre 35 e 20) e profunda (QI menor que 20). A maioria dos indivíduos com DI é identificada precocemente na infância devido a atrasos no desenvolvimento. Segundo a Academia Americana de Pediatria (*American Academy of Pediatrics* - AAP), o termo atraso do desenvolvimento (DD) é geralmente utilizado para crianças menores de cinco anos de idade e DI é aplicado a crianças mais velhas quando o teste de QI é válido. No Brasil, o termo atraso do desenvolvimento neuropsicomotor (DNPM) é utilizado para se referir a crianças com déficit cognitivo e atraso motor. A DI pode ocorrer isoladamente ou em combinação com malformações congênitas e outras características neurológicas, tais como epilepsia, deficiência sensorial e transtorno do espectro autista (*autism spectrum disorders* - ASD). Dessa forma, a DI pode ser classificada como sindrômica, quando o indivíduo apresenta outros sinais clínicos além da DI, e não sindrômica, quando a única característica evidente é a própria DI. A distinção de formas de DI sindrômicas e não sindrômicas se baseia no

envolvimento de outros órgãos, na presença de malformações e de uma face típica com dismorfismos. No entanto, não é incomum observar que mutações em um mesmo gene às vezes causam uma síndrome, enquanto outras levam a formas de DI não sindrômicas ou "isoladas" (Chiurazzi and Pirozzi 2016).

A DI leve é mais frequentemente observada (80-85%) e vários estudos relatam que sua prevalência é maior nos países em desenvolvimento. Este aumento de prevalência de DI leve foi atribuído a níveis socioeconômicos baixos que levam à desnutrição, privação cultural e falta de cuidado com a saúde. Apesar do conhecimento acumulado, pouco se sabe sobre os fatores genéticos que predis põem à DI leve, mas, claramente, nem todas as formas de DI leve são multifatoriais - algumas segregam em famílias com padrão de herança mendeliano. Já a DI moderada-profunda (15-20% dos casos) ocorre em todos os grupos sociais e em famílias de todos os níveis educacionais, tornando evidente o componente genético preponderante. As formas mais graves de DI geralmente são sindrômicas, ocorrendo em associação a outras manifestações clínicas. Assim, quanto mais grave a DI, maior a probabilidade de que esteja associada a manifestações físicas e neurológicas, e também de que um diagnóstico seja estabelecido (Chiurazzi and Oostra 2000; Chiurazzi and Pirozzi 2016; Maulik et al. 2011; Ropers 2008; Ropers 2010).

Os mecanismos genéticos que originam a DI incluem mutações gênicas e anomalias cromossômicas. As técnicas de bandas cromossômicas de maior resolução (800-2.000 bandas) possibilitaram a identificação de algumas anomalias cromossômicas estruturais associadas a fenótipos característicos e resultantes de desequilíbrio de dosagem de segmentos genômicos específicos. Usando bandeamento G, a citogenética clássica permite visualizar alterações iguais ou maiores que 4-10 Mb. Assim, síndromes causadas por deleções ou duplicações de regiões genômicas foram inicialmente identificadas em pacientes com alterações cromossômicas visíveis (Shaffer et al. 2007). Anomalias cromossômicas citogeneticamente visíveis explicam aproximadamente 15% de todos os casos de DI, sendo que 6-10% destes pacientes apresentam trissomia do cromossomo 21 (Ellison, Rosenfeld and Shaffer 2013; Rauch et al. 2006; Vissers, Gilissen and Veltman 2016).

Os limites de resolução da citogenética clássica foram superados pela introdução da técnica de hibridização *in situ* fluorescente (FISH) (Pinkel et al. 1986; Pinkel, Straume and Gray 1986), cuja utilização com sondas específicas para as extremidades dos cromossomos humanos possibilitou a detecção de microdeleções subteloméricas em pacientes com DI. Este tipo de abordagem revelou que os desequilíbrios subteloméricos crípticos ou submicroscópicos seriam responsáveis por mais de 2% da DI idiopática (Flint et al. 1995).

A introdução de técnicas baseadas em microarranjo genômico permitiu a detecção de alterações submicroscópicas em todo o genoma. O uso abrangente desta técnica de alta resolução mostrou que as deleções e duplicações submicroscópicas representam uma causa importante de DI. Cerca de 15-20% dos casos de DI são causados por desequilíbrios genômicos submicroscópicos conhecidos na literatura como *copy number variations* (CNVs) (Cooper et al. 2011; Hochstenbach et al. 2009; Miller et al. 2010; Rosenberg et al. 2006) ou alterações de número de cópias genômicas. Vários genes responsáveis por DI foram identificados por mapearem em regiões críticas de CNVs consideradas patogênicas. A identificação desses genes nas CNVs levou, em alguns casos, à detecção de mutações em indivíduos afetados por quadro clínico similar, porém não portadores da CNV, aumentando assim a lista de genes únicos identificados como causa de DI mendeliana (Ballif et al. 2012; Coe et al. 2014; Rosenfeld et al. 2015; Talkowski et al. 2011).

Em 1975 foi descrita a técnica de sequenciamento Sanger, determinante para a descoberta de genes causadores de doenças. O princípio de funcionamento deste método é a terminação aleatória do processo de replicação de uma sequência alvo de DNA por nucleotídeos terminadores (dideoxinucleotídeos). Os dideoxinucleotídeos não possuem o grupo hidroxila livre na extremidade 3' e, por isso, interrompem a polimerização da cadeia, gerando diferentes fragmentos, com tamanhos distintos, que são separados por eletroforese de acordo com o tamanho (Sanger and Coulson 1975). O sequenciamento Sanger foi aprimorado entre as décadas de 1980 a 2000, substituindo-se o método de detecção de radioatividade para o de

fluorescência. O sequenciamento Sanger foi também a técnica predominantemente utilizada para a realização do Projeto Genoma Humano.

A partir de 2003, com a publicação da primeira versão do genoma humano, iniciou-se um grande avanço na tecnologia de sequenciamento (Bentley 2000). Entre 2005 e 2010 foram desenvolvidas diferentes plataformas de sequenciamento de nova geração (*Next Generation Sequencing*, NGS). Em relação à técnica de sequenciamento Sanger, a técnica de NGS aumentou em várias ordens de magnitude a eficiência e a capacidade de produzir dados, enquanto o custo de 3 bilhões de dólares do Projeto Genoma Humano para sequenciar um genoma foi reduzido para próximo de U\$1000. O mecanismo que permitiu tal redução de custo foi a estratégia de sequenciamento massivo em paralelo, ou seja, métodos de sequenciamento automático de milhões de sequências de DNA simultaneamente. O NGS pode ser dividido em três etapas: preparo das bibliotecas, sequenciamento e análise dos dados gerados no sequenciador. Para auxiliar a análise desse imenso volume de dados foi necessário o desenvolvimento de ferramentas computacionais (Heather and Chain 2016). Podemos destacar os seguintes tipos de NGS: *target sequencing* ou captura de painel de genes (grupo de sequências de interesse), captura do exoma completo (*Whole Exome Sequencing*, WES - regiões codificadoras de todos os genes) e sem captura de alvos específicos, gerando o sequenciamento completo do genoma (*Whole Genome Sequencing* – WGS). Assim, o uso do NGS facilitou a pesquisa genética em DI, tendo grande poder para identificação de genes responsáveis por doenças raras (Chiurazzi and Pirozzi 2016; Vissers, Gilissen and Veltman 2016). Na Figura 1 podemos observar como a introdução de tecnologias moleculares aumentaram a eficiência de identificação de genes associados a DI (Vissers et al. 2016).

Figura 1. Identificação de genes associados com a DI ao longo dos anos. O total de genes identificados em 2015 (linha cinza) ultrapassava 700 genes. Esses genes estão distribuídos nos diferentes padrões de herança mendeliana: linha verde são os genes de herança ligada ao X, linha azul genes de herança autossômica dominante e linha roxa genes de herança autossômica recessiva (modificado de Vissers et al. 2016).

I.2. DI com padrões de herança mendeliana

A DI segrega em famílias apresentando todos os tipos de padrão de herança mendeliana: autossômicos e ligados ao cromossomo X, dominantes e recessivos. A Figura 2, retirada da revisão de Chiurazzi e Pirozzi (2016), mostra o mapeamento de mais de 800 genes associados a DI síndrômica e não síndrômica identificados em estudos de DI (Chiurazzi and Pirozzi 2016). Vários estudos na literatura mostram que a taxa de diagnóstico da DI utilizando WES pode variar de 24-42% (de Ligt et al. 2012; Deciphering Developmental Disorders 2015, 2017; Rauch et al. 2012; Vissers et al. 2016), a depender dos critérios de seleção dos pacientes.

Figura 2. Distribuição de genes de DI no genoma humano. Esquema mostrando ideogramas dos cromossomos humanos com padrão de bandas G; as setas verdes indicam o mapeamento dos genes já associados à DI síndrômica e não síndrômica. A figura evidencia que genes de DI estão espalhados por todos os cromossomos, exceto no Y (Chiurazzi and Pirozzi 2016).

I.2.1. DI ligada ao cromossomo X

Mutações em genes do cromossomo X constituem causa comum de DI moderada a grave em homens. Diferentes estudos realizados com coortes de afetados por DI permitiram estimar que há um excesso de 30 a 40% de homens afetados em relação às mulheres (Mandel and Chelly 2004; Ropers and Hamel 2005; Stevenson 2000). Também foi constatada a importância dos genes do cromossomo X como causa de DI devido à ocorrência de numerosas famílias nas quais a DI segrega de forma compatível com o padrão de herança ligada ao cromossomo X.

A síndrome do cromossomo X frágil (SXF) é a mais frequente dentre as síndromes de DI de herança ligada ao X e resulta de mutação no gene *FMR1*. A SXF afeta 1 em 4000-6000 homens e 1 em 8000-9000 mulheres; estima-se que 2-3% dos casos de DI em homens e 1% em mulheres sejam devidos a essa síndrome (Chelly and Mandel 2001; Crawford et al. 2001). A SXF é a causa de DI em cerca de 25% das famílias com DI ligada ao X, o que a faz a principal causa de DI herdada (Ropers and Hamel 2005). O gene *FMR1* apresenta na sua região 5' não traduzida uma repetição de trinucleotídeos (CGG)_n composta por 6-55 trincas nos alelos da população em geral. Alelos contendo repetições entre 55-200 trincas são classificados como pré-mutados; a proteína FMRP é produzida, porém esses alelos são instáveis, uma vez que a repetição pode se expandir durante a gametogênese feminina para 200 trincas ou mais, caracterizando então o que é chamado de mutação completa. Esses alelos com a mutação completa não são traduzidos e a ausência da proteína FMRP é a causa da síndrome do cromossomo X frágil.

Mutações que causam doenças monogênicas de DI ligada ao X já foram identificadas em 124 genes confirmados, além de 17 outros candidatos (Neri et al. 2018). Exceto por *FMR1*, nenhum desses genes individualmente explica mais do que 0,1% da DI (Lubs et al. 2012; Vissers et al. 2016). É interessante notar que os genes de DI no cromossomo X não explicam a totalidade do excesso de homens afetados (Mandel and Chelly 2004; Ropers and Hamel 2005) e alterações cromossômicas submicroscópicas no cromossomo X também devem contribuir. Adicionalmente, o trabalho de Jacquemont e col. (2014) sugere que homens apresentariam um limiar inferior às mulheres para manifestação de distúrbios de neurodesenvolvimento (Jacquemont et al. 2014); segundo essa hipótese, mutações em genes autossômicos poderiam resultar em maior manifestação de DI em homens do que em mulheres. Atualmente, embora mais de 120 genes de DI ligada ao X tenham sido identificados, tornou-se claro que formas de DI ligadas ao X são responsáveis por apenas cerca de 10% dos casos de DI no geral, e a grande maioria dos defeitos genéticos causativos de DI deve estar em autossomos (Hehir-Kwa et al. 2011; Kuss et al. 2011; Raymond et al. 2009; Ropers and Hamel 2005; Tarpey et al. 2009; Vissers et al. 2010).

I.2.2. DI autossômica dominante

A DI autossômica dominante quase sempre se deve a mutações *de novo* que podem explicar de 13-35% dos casos de pacientes com DI grave (Vissers et al. 2016). A quantidade de genes associados à DI de padrão de manifestação dominante é menor do que de genes recessivos, uma vez que indivíduos gravemente afetados raramente irão se reproduzir. De fato, a principal dificuldade para se identificar com segurança novos genes de DI autossômica dominante decorre da falta de casos familiares. Entretanto, vários estudos têm mostrado um grande aumento na detecção de genes responsáveis por DI autossômica dominante, principalmente devido à introdução do NGS.

Estudos com microarranjo genômico e sequenciamento de exoma demonstraram a importância de CNVs e de variações de nucleotídeo único (SNVs) como causa de DI com herança autossômica dominante (Gilissen et al. 2014). A técnica de NGS rapidamente permitiu identificar genes cujas mutações causam síndromes autossômicas dominantes associadas a DI, como a síndrome de Schinzel-Giedion, a síndrome de Kabuki e a síndrome de Bohring-Opitz (Hoischen et al. 2010; Hoischen et al. 2011; Ng et al. 2010).

Em 2010, Vissers e col. realizaram o sequenciamento de exoma de dez pacientes com DI grave e de seus genitores não afetados. Esta abordagem baseada em sequenciamento de exoma de trios (probando e genitores) foi feita pela primeira vez neste estudo e permitiu a identificação de mutações dominantes *de novo* como possível causa de DI (Vissers et al. 2010). A detecção de mutações *de novo* permitiu também a identificação de novos genes de DI autossômica dominante, mostrando que mutações em genes autossômicos expressos no cérebro são causa comum de DI idiopática (Ropers 2010).

Em 2012, o trabalho de de Ligt e col. confirmou a importância de mutações *de novo* como causa de DI grave; foram identificadas por sequenciamento de exoma completo mutações patogênicas em genes conhecidos em 16% dos pacientes estudados (de Ligt et al. 2012). Em outro estudo realizado com indivíduos por DI moderada ou grave, 41 probandos e seus pais, a causa molecular para a DI foi detectada em 29% dos pacientes, baseada em mutações

de novo em genes previamente relacionados com DI (Hamdan et al. 2014). Em 2015, foi realizado um grande estudo em coorte de 1.133 crianças com distúrbios graves do desenvolvimento, usando análise de exomas baseada em trios, resultando na detecção de mutações *de novo* em genes já associados com DI em 18% dos casos (Deciphering Developmental Disorders 2015). As variações na taxa de identificação da causa genética nesses estudos baseados em NGS podem ser explicadas por diferenças na seleção de pacientes, na interpretação de variantes e na qualidade do sequenciamento; entretanto, fica evidente a relevância do sequenciamento de exoma como teste para diagnóstico molecular.

Gilissen e col. (2014) realizaram o primeiro estudo utilizando o sequenciamento de genoma completo com pacientes afetados por DI grave e estimaram que 60% dos casos de DI poderiam ser explicados por mutações *de novo* em regiões codificadoras de genes já sabidamente associados a DI, sendo cerca de 39% dos casos causados por mutações do tipo SNVs, e 21%, CNVs. Embora o sequenciamento do genoma completo tenha sido usado neste estudo, a identificação e interpretação de variantes somáticas *de novo* fora das regiões codificadoras ainda é altamente complexa (Gilissen et al. 2014).

I.2.3. DI autossômica recessiva

A DI de herança autossômica recessiva é relativamente pouco explorada, uma vez que na Europa e nos Estados Unidos, que concentram grande parte da pesquisa mundial, a maioria dos casos de DI de herança recessiva são isolados, tanto por causa do tamanho pequeno das famílias quanto pelo risco de repetição de apenas 25%, situação que dificulta o mapeamento e identificação de novos genes (Gilissen et al. 2014; Musante and Ropers 2014; Najmabadi et al. 2011).

Musante e Ropers (2014) sugerem que a DI de herança recessiva representaria de 13-24% dos casos de afetados por DI na Europa. Os autores alertam para o fato de que essa taxa poderia ser uma superestimativa porque está baseada na suposição de que pais com uma única criança afetada terão a mesma probabilidade de buscar aconselhamento genético do que pais com duas ou mais crianças afetadas, o que provavelmente não é verdade (Musante and

Ropers 2014). Na Jordânia, por exemplo, onde existe uma frequência alta de casamentos consanguíneos, aproximadamente 32% de todos os casos de DI de herança recessiva são diagnosticados em filhos de primos em primeiro grau (Hamamy et al. 2007).

Casamentos consanguíneos representam um fator de risco importante para DI e anomalias congênitas, fato amplamente documentado pelo aumento significativo de DI na prole de casais consanguíneos (Bittles 2001; Durkin et al. 1998; Fernell 1998; Hamamy et al. 2007) e pela correlação entre a prevalência de anomalias congênitas e o coeficiente de consanguinidade (Bittles and Neel 1994).

Esforços em larga escala para mapear genes autossômicos recessivos para DI vêm, sobretudo, do grupo do Professor Ropers (Max Plank Institute, Alemanha), que combinou sequenciamento de exoma com mapeamento de regiões em homozigose em 136 famílias iranianas ou famílias de outras regiões com alto grau de consanguinidade, identificando inúmeros novos genes e *loci* (Kuss et al. 2011; Musante and Ropers 2014; Najmabadi et al. 2011).

Atualmente, mais de 300 genes foram identificados como causadores de DI autossômica recessiva. Devido à heterogeneidade da DI recessiva, é possível que a grande maioria dos genes ainda seja desconhecida (Musante and Ropers 2014; Ropers 2010; Vissers et al. 2016).

Duas abordagens principais têm sido utilizadas para identificar genes de DI de herança autossômica recessiva:

- 1) Estudo de irmandades com mais de um afetado: como exemplo, um estudo com 19 irmandades de afetados detectou a provável causa em nove delas, sendo três das alterações em genes mapeados no cromossomo X e as outras seis em autossomos (Schuurs-Hoeijmakers et al. 2013). Outros artigos identificando novos genes de DI a partir do estudo de irmandades foram publicados (Borck et al. 2015; Hildebrand et al. 2015). É importante notar que irmandades de afetados com genitores normais podem também ocorrer em associação a herança dominante, se

houver penetrância incompleta ou na presença de mosaicismismo gonadal (Chong et al. 2015), porém esses casos seriam raros e poderiam ainda ser considerados exceções.

- 2) Sequenciamento de regiões em homozigose (autozigoze) em famílias consanguíneas: embora a maioria dos estudos de identificação de genes recessivos de DI seja europeu, as famílias estudadas são, na sua maioria, Iranianas, Paquistanesas ou Turcas (Abbasi-Moheb et al. 2012; Abdullah et al. 2017; Bouyacoub et al. 2013; Larti et al. 2015; Mir et al. 2014; Oladnabi et al. 2015; Pajusalu et al. 2015; Rafiq et al. 2015; Rafiullah et al. 2016; Rajab et al. 2015; Roos et al. 2014), países nos quais o casamento consanguíneo é prática cultural.

II. Justificativa

II. Justificativa

Este trabalho de doutorado teve como objetivo central a identificação da causa genética da deficiência intelectual em irmandades com pelo menos dois afetados, com prévia exclusão da síndrome do cromossomo X frágil e de alterações cromossômicas detectáveis por microarranjos genômicos. Essa seleção pretendia privilegiar a DI de herança recessiva e, nos casos de irmandades que incluíam meninas, DI com padrão de herança provavelmente autossômica. Acredita-se que a maior parte dos genes a serem associados à DI sejam de padrão de herança autossômica. Essa abordagem permite verificar a eficiência da estratégia de NGS para triar a etiologia da DI em famílias brasileiras, assim como revelar novas variantes ou genes candidatos para DI.

Especificamente, o projeto contempla sequenciamento do exoma completo (*whole exome sequencing* - WES) de no mínimo dois irmãos em cada família e estudo de segregação de variantes raras selecionadas nos genitores, assim como em outros familiares quando relevante. Na análise dos exomas testamos as hipóteses de herança autossômica dominante (mosaicismo ou penetrância incompleta/expressividade variável), herança autossômica recessiva (em homozigose ou heterozigose composta) e herança ligada ao cromossomo X (mãe portadora e filhos afetados).

Investigar variantes genéticas que causam DI pode contribuir não somente para identificar variantes conhecidas e novas em genes já relatados, que permitirão estimar o impacto das mesmas na etiologia da DI, mas também para compreender o papel dessas variantes no desenvolvimento da inteligência e das habilidades cognitivas e sociais. Adicionalmente, a identificação da causa genética permite oferecer aconselhamento genético apropriado a essas famílias. Outro aspecto relevante é que novos genes candidatos a causar DI podem emergir deste tipo de análise, abrindo perspectivas para estudos futuros.

III. Objetivos

III. Objetivos

O objetivo central deste estudo foi a identificação da causa genética da DI idiopática em irmandades com pelo menos dois indivíduos afetados, com a detecção de novas variantes em genes conhecidos de DI ou novos genes associados a DI. Também almejávamos verificar, na ausência de grandes famílias consanguíneas, se o estudo de irmandades favoreceria a identificação de DI de padrão de herança recessiva.

Objetivos específicos:

1. Exclusão da síndrome do cromossomo X frágil e de alterações cromossômicas patogênicas detectáveis por microarranjo genômico.
2. Sequenciamento do exoma completo (*whole exome sequencing*, WES) de no mínimo dois irmãos de 25 famílias não aparentadas.
3. Análise dos dados genômicos para prospecção de variantes baseada em modelos de herança recessiva autossômica e ligada ao X e dominante autossômica e ligada ao X.
4. Estudo de segregação de variantes raras selecionadas nos genitores e, se necessário, em outros familiares, utilizando sequenciamento Sanger.
5. Aconselhamento genético das famílias com causa genética da DI identificada.

IV. Famílias

IV. Famílias

Foram selecionadas para o presente estudo famílias com irmandades com mais de um indivíduo portador de DI.

Portanto, vinte e cinco famílias com pacientes com DI foram incluídas no estudo: **onze** foram triadas no Serviço de Aconselhamento Genético do Departamento de Genética e Biologia Evolutiva, do Instituto de Biociências – USP, **duas** no Centro de Assistência Odontológica aos Portadores de Necessidades Especiais (CAOE) - Unidade Auxiliar de Estrutura Complexa da Faculdade de Odontologia do Campus de Araçatuba (FOA) da Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), e **doze** foram encaminhadas por médicos geneticistas colaboradores.

De modo geral, os pacientes apresentavam DI leve a moderada, avaliada com base em suas limitações cognitivas e comportamentais. Foram selecionados pacientes cujos quadros clínicos não se enquadravam em síndromes genéticas conhecidas e que tiveram resultados normais no teste molecular para a síndrome do cromossomo X frágil e na análise de microarranjo genômico (*array*-CGH).

O termo de consentimento livre e esclarecido foi assinado por todos os responsáveis pelos participantes dessa pesquisa. Este estudo faz parte de projeto aprovado pela Comissão de Ética em pesquisa – Seres Humanos – IBUSP (075/2007 e 2018/2.589.398).

A Tabela 1 sumariza as características básicas das 25 famílias, bem como os sinais clínicos dos 61 indivíduos afetados.

Tabela 1. Caracterização das famílias com indivíduos afetados por DI incluídas neste estudo.

Família	Número de afetados	Sexo dos afetados	Sumário do quadro clínico
Fam1	2	2 (M)	DI leve a moderado e dificuldade de fala
Fam2	2	2 (M)	DI leve, crises epiléticas e disgenesia de corpo caloso
Fam3	2	2 (M)	DI leve, microcefalia, dismorfismos faciais e manchas café-com-leite
Fam4	2	1 (M) / 1 (F)	DI leve, baixa estatura e face síndrômica
Fam5	2	2 (M)	DI moderada, atraso de fala e hiperatividade
Fam6	3	2 (M) / 1 (F)	DI moderada, convulsões, atraso de fala e agressividade
Fam7	2	2 (F)	DI profunda, convulsões, hipotrofia muscular com hipotonia e hipertrofia de mãos e pés
Fam8	2	2 (F)	DI moderada, ausência de fala e dismorfismos faciais
Fam9	2	2 (M)	Atraso de desenvolvimento neuropsicomotor e manifestações do espectro autista
Fam10	2	2 (F)	DI leve, hipotireoidismo, convulsões e baixa estatura
Fam11	2	1 (M) / 1 (F)	DI moderada a grave, ausência de fala e convulsões, menino com genitália ambígua e menina com quadro mais leve de DI
Fam12	2	1 (M) / 1 (F)	DI leve, dificuldade de fala e manifestações do espectro autista
Fam13	2	1 (M) / 1 (F)	DI moderada, dismorfismos faciais e ausência de fala

(cont.)

Tabela 1. Continuação.

Família	Número de afetados	Sexo dos afetados	Sumário do quadro clínico
Fam14	3	1 (M) / 2 (F)	DI leve, dismorfismos faciais e convulsões
Fam15	3	3 (F)	DI moderada a grave, convulsões e dificuldade de fala
Fam16	6	5 (M) / 1 (F)	DI leve e manifestações do espectro autista
Fam17	2	2 (M)	DI moderada a grave e dificuldade de fala
Fam18	3	3 (M)	DI leve, dificuldade de fala e manifestações do espectro autista
Fam19	2	1 (M) / 1 (F)	DI moderada a grave, convulsões, agressividade, baixa estatura e dificuldade de fala
Fam20	2	1 (M) / 1 (F)	DI leve e face síndrômica
Fam21	2	1 (M) / 1 (F)	DI leve, convulsões e atraso de fala
Fam22	2	2 (M)	DI moderada, dificuldade de fala, convulsão, dificuldade respiratória e hipoglicemia ao nascer
Fam23	2	1 (M) / 1 (F)	DI leve, dificuldade de fala e hipotireoidismo
Fam24	2	1 (M) / 1 (F)	DI profunda no menino e DI moderada na menina, ambos com diabetes
Fam25	5	5 (M)	DI moderada e agressividade
Total	61	39 (M) / 22 (F)	

M=masculino, F=feminino.

V. Métodos

V. Métodos

V.1. Obtenção das amostras de DNA genômico

O DNA genômico foi extraído a partir de leucócitos de sangue periférico utilizando-se a técnica de extração com fenol e clorofórmio (Sambrook et al. 1987), com modificações. A qualidade das amostras foi avaliada no equipamento *NanoDrop 2000 spectrophotometer* (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA), e as concentrações das amostras foram estimadas por fluorescência no equipamento *Qubit 2.0 Fluorometer* (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA).

V.2. Sequenciamento do exoma completo (WES)

Para investigação preliminar, foram encaminhadas amostras de DNA de duas das famílias estudadas para serem sequenciadas na empresa *BGI Americas Corporation Service Agreement*; as bibliotecas foram capturadas com o *kit SureSelect Human All Exon V5* (Agilent Technologies, Santa Clara, Califórnia, EUA) e o sequenciamento foi realizado no sequenciador IonTorrent (Thermo Fischer Scientific, Carlsbad, Califórnia, EUA).

Três famílias tiveram as bibliotecas preparadas e capturadas com o *kit Illumina Nextera Rapid Capture Expanded Exome Kit* (Illumina, Inc., San Diego, California, EUA), região alvo de 62Mb, incluindo todos os exons codificantes e UTRs de genes descritos no RefSeq, além de miRNAs (conforme descrito pelo fabricante Illumina, Inc., San Diego, California, EUA); o sequenciamento foi realizado no sequenciador HiSeq 2500 da Illumina (Illumina, Inc., San Diego, California, EUA) e todo o procedimento foi feito pela equipe do Centro de Estudos sobre o Genoma Humano e Células-tronco (CEGH-CEL).

A seguir, a técnica de sequenciamento de nova geração foi implementada e padronizada em nosso laboratório. Por ser um período de padronização, diferentes kits foram testados e comparados. As bibliotecas foram construídas a partir de amostras de DNA de pacientes e seus familiares usando o kit

SureSelectXT ou *SureSelect QXT Target Enrichment System for Illumina Paired-End Sequencing Library*, os procedimentos foram realizados conforme descrito pelo fabricante (Agilent Technologies, Santa Clara, Califórnia, EUA). A captura das regiões codificadoras para o sequenciamento foi feita da seguinte maneira:

- 15 famílias - kit *SureSelectXT Clinical Research Exome* (Agilent Technologies, Santa Clara, Califórnia, EUA), região alvo de 54Mb, incluindo todos os exons codificantes de genes descritos no RefSeq e enriquecimento em regiões já associadas a doenças;
- 5 famílias - *SureSelectXT Human All Exon V6*, região alvo de 60Mb, incluindo todos os exons codificantes de genes descritos no RefSeq (Agilent Technologies, Santa Clara, Califórnia, EUA).

O sequenciamento paralelo massivo dos fragmentos de DNA *paired-end* dessas 20 famílias foi realizado no sequenciador HiSeq 2500 da Illumina (Illumina, Inc., San Diego, California, EUA), no CEGH-CEL. A Tabela 2 identifica quais plataformas de sequenciamento foram usadas para cada família.

Tabela 2. Descrição das plataformas de captura e construção de bibliotecas genômicas para sequenciamento completo de exoma utilizadas.

Família	Kit de captura
Fam1	<i>SureSelect Human All Exon V5</i>
Fam2	<i>SureSelect Human All Exon V5</i>
Fam3	<i>SureSelectXT Clinical Research Exome</i>
Fam4	<i>SureSelectXT Clinical Research Exome</i>
Fam5	<i>SureSelectXT Clinical Research Exome</i>
Fam6	<i>SureSelectXT Clinical Research Exome</i>
Fam7	<i>SureSelectXT Clinical Research Exome</i>
Fam8	<i>SureSelectXT Clinical Research Exome</i>
Fam9	<i>SureSelectXT Clinical Research Exome</i>
Fam10	<i>SureSelectXT Clinical Research Exome</i>
Fam11	<i>Illumina Nextera Rapid Capture Expanded Exome</i>
Fam12	<i>Illumina Nextera Rapid Capture Expanded Exome</i>
Fam13	<i>Illumina Nextera Rapid Capture Expanded Exome</i>
Fam14	<i>SureSelectXT Clinical Research Exome</i>
Fam15	<i>SureSelectXT Clinical Research Exome</i>
Fam16	<i>SureSelectXT Clinical Research Exome</i>
Fam17	<i>SureSelectXT Clinical Research Exome</i>
Fam18	<i>SureSelectXT Clinical Research Exome</i>
Fam19	<i>SureSelectXT Clinical Research Exome</i>
Fam20	<i>SureSelectXT Clinical Research Exome</i>
Fam21	<i>SureSelectXT Human All Exon V6</i>
Fam22	<i>SureSelectXT Human All Exon V6</i>
Fam23	<i>SureSelectXT Human All Exon V6</i>
Fam24	<i>SureSelectXT Human All Exon V6</i>
Fam25	<i>SureSelectXT Human All Exon V6</i>

V.3. Análise de bioinformática dos dados do WES

Após o sequenciamento, os arquivos “.bcl” contendo as bases de DNA de cada fragmento sequenciado foram transformados em arquivos fastq pela equipe do CEGH-CEL. Os arquivos fastq foram considerados dados crus de

sequenciamento e armazenados como tal. A transformação desses dados crus em arquivos VCF (*Variant Call Format*) e BAM foi realizada em colaboração com o Dr. Israel Tojal da Silva, do Centro Internacional de Pesquisa e Ensino (CIPE) do AC Camargo Cancer Center, como descrito abaixo.

As *reads* (leituras de uma base) foram processadas para remoção de bases de baixa qualidade (*Phred quality score* <Q30) e de *reads* menores que 35 bp usando Sickle (Josh 2011), Scythe (Buffalo 2011) e FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Em seguida, as *reads* foram processadas e alinhadas ao genoma de referência estabelecido (versão UCSC hg19; <http://hgdownload.cse.ucsc.edu/downloads.html>), gerando arquivos “.sam” (*Sequence Alignment Map*). Em seguida, foi realizada a conversão dos arquivos SAM para BAM (*Binary Alignment Map*) com o *software* Burrows-Wheeler Aligner (BWA) (Li and Durbin 2010) e o *software* Samtools (Li et al. 2009). Foram marcadas as sequências provavelmente geradas por duplicatas de PCR (*mark duplicates*) e a sua remoção foi feita por meio da ferramenta Picard (<http://picard.sourceforge.net/index.shtml>).

A distribuição dos dados e a cobertura de sequências-alvo (*reads on-target*) foram avaliadas usando o pacote Picard CalculateHsMetrics. Para a recalibração e o realinhamento, identificação de variantes de nucleotídeo único (SNV) e inserções/deleções (indel), determinação de scores de qualidade básica (escala de Phred), utilizou-se o GATK (www.broadinstitute.org/gatk/). As variantes detectadas foram armazenadas em arquivos do tipo VCF.

V.4. Anotação, filtragem e seleção de variantes

A anotação e filtragem de variantes foram executadas utilizando o *software* VarSeq versão 1.5.0 (Golden Helix, Inc., Bozeman, Montana, EUA), usando arquivos VCF. Os arquivos BAM foram utilizados para inspeção visual da qualidade do sequenciamento de variantes selecionadas.

Foram utilizados diferentes bancos de dados públicos de frequência populacional de variantes: ExAC (<http://exac.broadinstitute.org/>), gnomAD (genome Aggregation Database - <http://gnomad.broadinstitute.org/>), ABraOM

(<http://abraom.ib.usp.br/>), 1K Genomes (<http://www.1000genomes.org/>) e dbSNP versão 150 (<http://www.ncbi.nlm.nih.gov/projects/SNP/>). Verificamos a presença das variantes nos bancos de dados clínicos: OMIM (<https://www.omim.org>), Clinvar (<https://www.ncbi.nlm.nih.gov/clinvar>) e HGMD (<http://www.hgmd.cf.ac.uk/ac/index.php>).

Foram selecionadas variantes com base em qualidade (Phred score ≥ 20 e Genotype Quality ≥ 20), frequência alélica na amostra ($\geq 10\%$), frequência populacional ($\leq 1\%$) e tipo – mutações não sinônimas (*missense*), mutações sem sentido (*nonsense*), mutações que afetam o processamento do RNAm (*splicing*) e mutações de mudança de quadro de leitura (*frameshift*), essas últimas causadas por inserções ou deleções (*indels*) que envolvem um número de pares de bases que não é múltiplo de três.

A patogenicidade de variantes *missense* foi avaliada *in silico* por seis diferentes algoritmos e tais dados são fornecidos pelo banco dbNSFP (<http://varianttools.sourceforge.net/Annotation/DbNSFP>, versão 2.4): SIFT (Scale-Invariant Feature Transform - <http://sift.bii.aster.edu.sg/>), PolyPhen-2 (Polymorphism Phenotyping v2 - <http://genetics.bwh.harvard.edu/pph2>), Mutation taster (<http://www.mutationtaster.org/>), Mutation assessor (<http://mutationassessor.org/>) e FATHMM (Functional Analysis through Hidden Markov Models (v2.3) - <http://fathmm.biocompute.org.uk/>). O potencial efeito prejudicial à função da proteína também foi avaliado usando o software VEP (Variant Effect Predictor) do Ensembl (<https://www.ensembl.org/>) para variantes selecionadas. As variantes foram priorizadas de acordo com o fenótipo associado usando a ferramenta VarElect (<https://varelect.genecards.org/>). Também utilizamos a ferramenta InterVar (<http://wintervar.wglab.org/>), que facilita a classificação clínica de variantes genéticas segundo os critérios do Colégio Americano de Genética Médica e Genômica (ACMG) 2015. O InterVar produz uma avaliação básica inicial a partir da anotação do ANNOVAR, classificando as variantes em benigna (*benign*), provavelmente benigna (*likely benign*), significado indeterminado (*variant of uncertain significance* - VUS), provavelmente patogênica (*likely pathogenic*) e patogênica (*pathogenic*), classificações que podem ser modificadas pelo usuário se forem adicionadas

evidências detalhadas adicionais, como estudo de segregação (Li and Wang 2017).

Foram selecionadas variantes priorizadas de acordo com os critérios descritos acima e que co-segregavam no par de irmãos estudado. Em cada caso, foi verificada a compatibilidade do gene identificado com o fenótipo dos pacientes, para genes conhecidos. Verificamos variantes em heterozigose nos dois irmãos para doenças com padrão de herança autossômica dominante; variantes em homozigose ou em heterozigose composta para doenças com padrão de herança autossômica recessivo e variantes em hemizigose (homens) ou heterozigose (mulheres) para doenças com padrão de herança ligada ao X recessiva ou dominante. Nas famílias cujas variantes selecionadas estavam em genes sabidamente sujeitos a *imprinting*, verificamos a herança parental da variante para afastar ou incluir a variante como relacionada ao quadro clínico dos afetados. A análise de segregação das variantes nas famílias foi realizada por sequenciamento Sanger.

V.5. Sequenciamento Sanger

Para validação e análise de segregação de variantes patogênicas e provavelmente patogênicas detectadas foi utilizado o sequenciamento bidirecional do segmento do gene em questão no sequenciador automático *ABI3730 Genetic Analyzer* (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA), no CEGH-CEL. Foram desenhados *primers* para a região de interesse utilizando o *software* Primer3 (<http://bioinfo.ut.ee/primer3-0.4.0/>). Seguiram-se amplificação por PCR convencional, purificação com ExoSAP-IT (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA), reação de sequenciamento com BigDye Terminator (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA), seguida de purificação final em coluna de sephadex (Amersham Biosciences, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA) e aplicação do produto final no sequenciador.

V.6. Análise do padrão de inativação do cromossomo X

A investigação do padrão de inativação do cromossomo X foi realizada nas genitoras de pacientes e em outras mulheres de famílias nas quais variantes detectadas no cromossomo X estavam segregando. Analisamos o padrão de metilação do gene *AR* (Xq12), como descrito por Allen e col., com modificações (Allen et al. 1992).

O exon 1 do gene *AR* possui um número variável de repetições CAG, traduzidas em uma cadeia de poliglutaminas de comprimento variável, no domínio aminoterminal de transativação da proteína. Além da repetição CAG ser altamente polimórfica (heterozigose de 0,9) existe, próximo a ela, um sítio de restrição da enzima *HpaII*, que se encontra metilado apenas no X inativo. Como a enzima *HpaII* é sensível a metilação, ela só é capaz de cortar os sítios que se encontram desmetilados, ou seja, aqueles localizados no cromossomo X ativo. Dessa maneira, pode-se determinar o padrão de inativação do cromossomo X. Os *primers* amplificam as repetições CAG do gene *AR* que flanqueiam o sítio de restrição da enzima, o que impede que os alelos digeridos pela *HpaII* sejam amplificados. Isso permite a determinação do padrão de inativação do cromossomo X a partir da análise dos alelos amplificados antes e após a digestão com a enzima *HpaII*. As amostras amplificadas após a reação de PCR foram sequenciadas no equipamento *ABI3730 Genetic Analyzer* (Applied Biosystems, Thermo Fischer Scientific, Carlsbad, Califórnia, EUA), do Centro de Pesquisas sobre o Genoma Humano e Células Tronco (CEGH-CEL), e analisadas no *software* GeneMapper. O desvio de inativação foi quantificado de acordo com Bittel e col. 2008 (Bittel et al. 2008).

V.7. Droplet digital PCR (ddPCR)

A técnica de ddPCR foi utilizada em uma das famílias para investigação de mosaicismo somático nos genitores. Este método foi utilizado para investigar variante em mosaicismo somático devido à sua sensibilidade. A PCR digital permite uma quantificação absoluta e altamente sensível de ácidos nucleicos sem

a necessidade de curva padrão. No sistema de PCR Digital QX200 (Bio-Rad Laboratories, Inc, Berkeley, California, EUA), uma amostra é dividida em até 20.000 gotas, fragmentando assim o ensaio e diminuindo um possível viés de amplificação. Após a amplificação, as gotas contendo a sequência alvo são detectadas por fluorescência e classificadas como positivas para DNA selvagem ou sem a variante (fluorescência HEX) ou positivas para DNA com a variante (fluorescência FAM); gotas sem DNA e fluorescência são negativas. A sonda para o ensaio foi desenhada e comprada da empresa Bio-Rad (Bio-Rad Laboratories, Inc, Berkeley, California, EUA), ProbeFAM/HEXdHsaMDS178830062. A investigação foi realizada em colaboração com as Dras. Giovana Tardin Torrezan e Ana Carolina Kerekes Miguez, pesquisadoras do grupo da Dra. Dirce Maria Carraro do Centro Internacional de Pesquisa e Ensino (CIPE) do AC Camargo Cancer Center.

As reações foram realizadas em triplicata para cada amostra. Cada reação teve um volume final de 20 µl, contendo 10 µl de 2x ddPCR SuperMix for Probes (noUDTP) (Bio-Rad Laboratories, Inc, Berkeley, California, EUA), 2 µl (20x) de mix contendo *primers* e sondas selvagens (HEX) e mutadas (FAM), 22ng de DNA e água livre de nucleases UltraPure Distilled Water (Invitrogen) para alcançar o volume total de 20µl. As reações foram transferidas para cartuchos DG8 (Bio-Rad Laboratories, Inc, Berkeley, California, EUA) junto com 7 µl de Droplet Generation Oil for Probes (Bio-Rad Laboratories, Inc, Berkeley, California, EUA) para cada amostra, para a geração de gotículas na plataforma QX200 Droplet Generator (Bio-Rad Laboratories, Inc, Berkeley, California, EUA). Para cada reação foram geradas cerca de 20 mil gotículas de emulsão água-óleo. As gotículas foram transferidas para uma placa de 96 poços (Eppendorf), que foi selada na plataforma PX1 PCR Plate Sealer (Bio-Rad Laboratories, Inc, Berkeley, California, EUA) para reação de amplificação no T100 Thermal Cycler (Bio-Rad Laboratories, Inc, Berkeley, California, EUA). Após amplificação, a placa foi carregada na plataforma QX200 Droplet Reader (Bio-Rad Laboratories, Inc, Berkeley, California, EUA) e os resultados foram verificados no software QuantaSoft 1.7.4 (Bio-Rad Laboratories, Inc, Berkeley, California, EUA).

VI. Resultados e discussão dos casos

VI. Resultados e discussão dos casos

VI.1. Resumo dos resultados

No sequenciamento completo do exoma das 25 famílias estudadas obteve-se uma média de cobertura em sequências-alvo (*on target*) de 36x por amostra, com mediana de 92% das regiões-alvo cobertas a um mínimo de 1x profundidade (*read depth*) e 78% com 10x.

Em uma das famílias, foi identificada uma variante patogênica associada ao diabetes mellitus que, mesmo não estando relacionada a DI, está descrita abaixo. Não foram encontradas variantes patogênicas ou provavelmente patogênicas nos genes contidos na lista de achados secundários do Colégio Americano de Genética Médica e Genômica (ACMG) (Kalia et al. 2017).

Em nove famílias (36%) foram detectadas variantes que potencialmente explicam o quadro clínico dos pacientes avaliados. A Tabela 3 apresenta um resumo desses achados e o resultado de cada uma dessas famílias será discutido em detalhes a seguir.

Não detectamos variantes patogênicas/provavelmente patogênicas associadas ao quadro clínico de DI em 16 das famílias investigadas.

Tabela 3. Variantes potencialmente causativas dos fenótipos em nove famílias com pacientes afetados por DI.

Família	Gene	Citobanda	Posição genômica (GRCh37/hg19)	exon	Herança	Transcrito	cDNA	Proteína	Efeito	Previsão de patogenicidade	Herança da variante
Fam1	<i>UBE2A</i>	Xq24	X:118716586	5	Ligada ao X	NM_001282161.1	c.277C>G	p.Gln93Glu	<i>missense</i>	Danoso/ danoso/possivelmente danoso/danoso	Mat
Fam2	<i>MED12</i>	Xq13.1	X:70344949	15	Ligada ao X	NM_005120	c.2179G>A	p.Asp727Asn	<i>missense</i>	Tolerado/danoso/beni gno/tolerado	Mat
Fam3	<i>CDK5RAP2</i>	9q33.2	9:123280770 9:123298747	12 7	AR	NM_018249.5 NM_018249.5	c.1246delG c.564_565dupGA	p.Glu416Argfs p.Lys189Argfs	LoF (<i>frameshift</i>) LoF (<i>frameshift</i>)		Pat Mat
Fam6	<i>SLC6A8</i>	Xq28	X:152954234	1	Ligada ao X	NM_005629.3	c.205G>C	p.Ala69Pro	<i>missense</i>	Danoso/ danoso/possivelmente danoso/danoso	Mat
Fam8	<i>SETBP1</i>	18q12.3	18:42529933	4	AD	NM_015559.2	c.628delC	p.His210Thrfs	LoF (<i>frameshift</i>)		Mat
Fam9	<i>ARHGAP4</i>	Xq28	X:153178732	11	Ligada ao X	NM_001164741.1	c.1471C>T	p.Arg491Trp	<i>missense</i>	Danoso/ danoso/possivelmente danoso/danoso	Mat
Fam11	<i>PHF6</i>	Xq26.2	X:133562204	3' UTR	Ligada ao X	NM_032458.2 NM_001015877.1	c.*2844C>T c.*2514C>T				Mat
Fam13	<i>MED13L</i>	12q24.21	12:116422099	20	AD	NM_015335.4	c.4417C>T	p.Q1473X	LoF (<i>stop_gained</i>)		?
Fam25	<i>MBD5</i>	2q23.1	2:149216407	6	AD	NM_018328.4	c.80G>A	p.Arg27His	<i>missense</i>	Danoso/ danoso/possivelmente danoso/danoso	Mat

Mat=materno, Pat= paterno, LoF= *loss-of-function*; AD= autossômica dominante; AR=autossômica recessivo; previsão de patogenicidade: SIFT/Mutation taster/Polyphen2/PROVERAN; UTR=região não traduzida

VI.2. Identificação de variante patogênica não relacionada a DI

A Família 24 foi encaminhada pelo Dr. Hélio van der Linden, neuropediatra de Goiânia. Neste estudo, foi investigada a causa genética do quadro clínico de dois irmãos, um do sexo masculino e outra do sexo feminino, afetados por DI e diabetes mellitus. O heredograma desta família é mostrado na Figura 3.

A análise do sequenciamento de exoma completo dos pacientes portadores de DI resultou na identificação em ambos de uma variante *frameshift* em heterozigose no exon 6 do gene *HNF1A* (OMIM*142410). Este gene está localizado em 12q24.31 e mutações em heterozigose estão associadas a doenças relacionadas ao diabetes mellitus (OMIM #612520 - DIABETES MELLITUS, INSULIN-DEPENDENT20, IDDM20; OMIM #600496 - MATURITY-ONSET DIABETES OF THE YOUNG, TYPE 3; MODY3 e OMIM #125853 - DIABETES MELLITUS, NONINSULIN-DEPENDENT; NIDDM).

A presença da variante *frameshift* [NC_000012.11:g.121434373delT:NM_001306179.1:c.1137delT:p.Val380Serfs] foi confirmada nos irmãos por sequenciamento Sanger (Figura 4). Esta variante está relatada no banco de dados ClinVar como patogênica, NM_000545.5(HNF1A). Esses dados indicam que esta variante é responsável pelo diabetes mellitus dos irmãos; a mãe dos afetados por DI e duas irmãs também tinham diabetes mellitus, evidenciando a segregação da mutação. A causa da DI permanece não esclarecida.

Figura 3. Heredograma da Família 24, na qual uma variante patogênica no gene *HNF1A* segrega com diabetes mellitus num padrão de herança autossômica dominante.

Figura 4. Sequenciamento Sanger em amostras de DNA genômico mostrando a deleção em heterozigose da variante g.121434373delT (c.1137delT), encontrada no exon 6 do gene *HNF1A* nos irmãos afetados (II.1 e II.2). Apenas a presença da sequência normal foi observada no controle.

VI.3. Apresentação das famílias portadoras de variantes potencialmente patogênicas para a DI

VI.3.1. FAMÍLIA 1 - O ESPECTRO FENOTÍPICO DE MUTAÇÕES EM *UBE2A*

Resumo

A irmandade aqui estudada é composta por dois irmãos do sexo masculino afetados por DI leve a moderada, atraso de fala e disartria, nos quais foi identificada uma variante do tipo mudança de nucleotídeo (*missense*) no gene *UBE2A* (OMIM *312180), que codifica uma enzima conjugadora de ubiquitina (E2). Este gene está localizado no braço longo do cromossomo X (Xq24), e é associado à síndrome de deficiência intelectual ligada ao cromossomo X tipo Nascimento (OMIM #300860), caracterizada por quadro clínico variável que pode incluir dismorfismos craniofaciais, anomalias da pele, micropênis, DI moderada a grave, atraso motor, fala prejudicada ou ausente e convulsões. A variante *missense* c.277C>G (p.Gln93Glu) detectada nessa família ocorre no códon de um aminoácido evolutivamente conservado e não está documentada em nenhum dos grandes bancos de variantes genéticas populacionais. As análises *in silico* classificaram a variante c.277C>G como "danosa" (SIFT e MutationTaster) e "possivelmente danosa" (Polyphen-2). Estudos funcionais da proteína *UBE2A* construída com esta variante foram realizados por colaboradores no Laboratório Nacional de Biotecnologia (LNBio), em comparação com a proteína selvagem. Os resultados demonstraram que o resíduo Q93, no qual a mutação ocorreu, está diretamente envolvido na montagem de cadeias de poliubiquitina do sítio ativo da enzima *UBE2A* e que a mutação modificou a atividade enzimática normal. Os resultados também mostraram que a atividade alterada da proteína mutante Q93E poderia ser restaurada em pH mais alto, fornecendo a primeira evidência de reversão do efeito de uma mutação em *UBE2A*.

O gene *UBE2A* e a deficiência intelectual

Em 2006, Nascimento e col. em nosso Departamento descreveram pela primeira vez uma forma de DI ligada ao cromossomo X causada por mutação de perda de função no gene *UBE2A*. O gene *UBE2A*, localizado no braço longo do cromossomo X (Xq24), codifica a enzima E2 conjugadora de ubiquitina, que participa do processo de ubiquitinação de proteínas e é altamente conservada em eucariontes.

A ubiquitinação é um mecanismo importante na regulação dos níveis de proteína em células de mamíferos, como por exemplo: direcionamento de proteínas para degradação, ativação e inativação de fatores de transcrição, internalização de receptores e modificações epigenéticas (Bochtler et al. 1999; Conaway et al. 2002; Dover et al. 2002; Suresh et al. 2016; Weissman 2001). O processo de ubiquitinação consiste na adição covalente de moléculas de ubiquitina a proteínas específicas em uma reação enzimática que requer a participação de três enzimas principais: (I) as enzimas E1 (ativadoras), responsáveis pela ativação da ubiquitina, (II) as enzimas E2 (conjugadoras), que se ligam à ubiquitina ativada, e (III) as enzimas E3 (ligase), que se ligam às proteínas-alvo e transferem a ubiquitina ativada da enzima E2 para o substrato (Figura 5).

A deficiência do sistema de ubiquitinação pode levar a muitos efeitos patológicos, incluindo câncer, distúrbios neurodegenerativos e diabetes (Haddad et al. 2013; Howard et al. 2007; Petroski 2008; Riley et al. 2013; Ross and Poirier 2004; Zhao et al. 2016).

Figura 5. Representação esquemática do processo de adição covalente de ubiquitina (Ub) a proteínas alvo. E1, enzima ativadora de ubiquitina; E2, enzima conjugadora de ubiquitina e E3, ligase de ubiquitina [baseado em (Suresh et al. 2016)].

No trabalho de Nascimento e col. (2006) foram descritas duas gerações de uma família com três indivíduos do sexo masculino afetados por DI e aparentados por suas mães fenotipicamente normais. As mães, portadoras obrigatórias de uma mutação no cromossomo X, apresentavam desvio total de inativação do cromossomo X em linfócitos de sangue periférico. O sequenciamento dos exons do gene *UBE2A* detectou uma mutação do tipo introdução de códon de parada prematuro (*nonsense*) que segregava com a doença, a substituição de uma citosina por uma timina na posição 382 da região codificadora do gene. A mutação c.382C>T gera um códon de parada prematuro UAG (Q128X). Os pacientes tinham quadro clínico semelhante, incluindo DI grave com comprometimento significativo de fala, convulsões e hirsutismo generalizado de tronco e membros (Nascimento et al. 2006). Em 2010, Budny B. e col. relataram duas diferentes mutações *missense* no gene *UBE2A* (G23R e R11Q) em pacientes de duas famílias que desenvolveram uma forma sindrômica de DI com fenótipo similar aos pacientes portadores da mutação *nonsense* Q128X (Budny et al. 2010). No mesmo ano, uma deleção de 400 kb incluindo o gene *UBE2A* inteiro foi detectada em um menino de 4 anos de idade com DI sindrômica; a deleção foi herdada de sua mãe, clinicamente normal (Honda et al. 2010). Posteriormente, estudos realizados com as mutações I87MfsX14, R7W e Q128X mostraram que a proteína UBE2A tem um papel crítico na manutenção da função neuronal (Haddad et al. 2013). Czeschik J.C. e col. (2013) compararam os quadros clínicos de pacientes com diferentes alterações no gene *UBE2A* (deleções intragênicas, mutações *missense*, *nonsense*, *frameshift* e grandes deleções que incluíam o gene inteiro), concluindo que ocorria um padrão heterogêneo de dismorfismos faciais e outras características físicas e de desenvolvimento, o que sugeria que pacientes com mutações neste gene poderiam ser subdiagnosticados (Czeschik et al. 2013).

Recentemente, uma nova mutação *missense* (c.76G>A) foi identificada segregando em uma família japonesa, herdada de mãe com desvio de inativação do cromossomo X de razão 92:8 (Tsurusaki et al. 2017). Giugliano T. e col. (2018) também relataram uma nova mutação *missense* (c.330G>A) que faz com que o exon 5 não seja transcrito, além de promover transcrição fora do quadro de

leitura, provavelmente codificando uma forma truncada da enzima UBE2A; esta mutação não foi detectada em DNA de células sanguíneas maternas, sugerindo mosaïcismo germinativo materno uma vez que dois filhos herdaram a mutação (Giugliano et al. 2018).

Os diferentes tipos de alterações do gene *UBE2A* já identificadas - *missense*, *nonsense*, *frameshift*, microdeleções intragênicas ou deleções de segmentos maiores que englobam o gene *UBE2A* - podem explicar parcialmente o espectro fenotípico observado nos pacientes com a síndrome de DI ligada ao X do tipo Nascimento (Budny et al. 2010; Czeschik et al. 2013; de Leeuw et al. 2010; Haddad et al. 2013; Honda et al. 2010; Nascimento et al. 2006). Com relação às diferentes mutações *missense* já identificadas, o cenário é mais complexo: algumas resultam em quadro clínico indistinguível das mutações de perda de função, enquanto outras foram identificadas em pacientes com quadro clínico menos típico do originalmente descrito.

Neste trabalho, nós relatamos uma nova família com dois irmãos do sexo masculino, ambos afetados por quadro clínico de DI leve a moderada e comprometimento de fala, causado por uma nova variante *missense* no gene *UBE2A* que foi identificada por sequenciamento total de exoma.

Descrição da Família 1 (pacientes 1a e 1b)

A família aqui apresentada foi convidada a participar do projeto por ser composta por dois irmãos afetados por DI leve/moderada, nascidos de pais saudáveis e não consanguíneos, e com três irmãs não afetadas. O heredograma desta família é mostrado na Figura 6.

Figura 6. Heredograma da Família 1, na qual segrega uma variante no gene *UBE2A* com um padrão de herança recessiva ligada ao cromossomo X.

Paciente 1a (II.1)

Nascimento a termo, com 4.160g; cesárea foi indicada por causa de sofrimento fetal com eliminação de mecônio na bolsa amniótica; índices de Apgar (1 e 5 min) de 7 e 9 foram considerados satisfatórios. Eletroencefalograma realizado aos 6 anos de idade mostrou foco irritativo na região centro-temporal direita, sem quadro convulsivo associado ou outros sinais de epilepsia. O QI foi avaliado também aos 6 anos e mostrou escores de 65 (verbal) e 82 (teste/critério de Raven). Examinado aos 34 anos de idade, era portador de DI leve a moderada, sem história de atraso pronunciado na fonação, mas com gagueira (de média intensidade) e dificuldade importante na articulação correta das palavras. Tem altura 181cm, peso 86kg, hipertelorismo ocular (distância intercantal interna (dii) 31mm - < percentil 25 - e distância intercantal externa (die) 103mm - > percentil 97,5), perímetro cefálico (PC) 59cm, redemoinhos de cabelo, leve hipoplasia de face média, sobrancelhas espessas, inclinação normal das fissuras palpebrais, ponte nasal reta, boca com dimensões normais, micrognatia leve, pilificação aumentada na região posterior da nuca (cabelos de implantação normal) mas sem hirsutismo exagerado ou generalizado, genitália normal (sic; operado de

fimose), pés de tamanho normal para a altura (calça número 43) e semi-planos (valgos), dorso do pé de aspecto normal e textura da pele normal.

Paciente 1b (II.3)

Nascimento a termo de parto cesariano com 3.890g, em boas condições. Examinado aos 32a6m de idade, é portador de DI leve a moderada; sem história de atraso pronunciado na fonação, mas com dificuldade muito importante na articulação correta das palavras. EEG normal, sem história de convulsões. Apresenta biótipo diferente do irmão; altura 171cm, peso 55kg, sem hipertelorismo ocular (dii 33mm - percentil 25-50 - e die 92mm - percentil 75-97,5), PC 57cm, redemoinhos de cabelo e calvície frontal moderada, hipoplasia de face média, face estreita e triangular, sobrancelhas espessas, inclinação normal das fissuras palpebrais, impressão de olhos amendoados, ponte nasal reta, boca com dimensões normais, genitália normal (sic), pilificação aumentada na região anterior do tórax e nos antebraços mas sem hirsutismo exagerado ou generalizado, pés ligeiramente valgos, de tamanho normal (calça número 42), dorso do pé de aspecto normal e textura da pele normal.

Resultados

A análise do sequenciamento de exoma completo dos pacientes II.1 e II.3 resultou na identificação de uma variante *missense* em hemizigose no exon 5 do gene *UBE2A* [NM_003336:c.277C>G:p.(Gln93Glu)]. A proteína UBE2A contém um domínio catalítico de conjugação de ubiquitina (UBCc) e a substituição do aminoácido p.(Gln93Glu) ou Q93E está localizada neste domínio (Figura 7). Também observamos que a mutação Q93E está localizada ao lado do sítio ativo da proteína (C88) e, portanto, a interação da proteína com o peptídeo de ubiquitina poderia ser prejudicado (Figura 8).

Esta variante *missense* não foi previamente relatada em nenhum dos bancos de dados populacionais consultados (AbraOM, 1K Genomes, ESP6500, ExAC e gnomAD). A mutação ocorreu em um aminoácido que é evolutivamente conservado em diferentes espécies e sua patogenicidade foi avaliada usando

algoritmos, sendo classificada como “danosa” por SIFT, “possivelmente danosa” por Polyphen-2 e “danosa” por MutationTaster.

Figura 7. (A) Diagrama do gene *UBE2A*: o gene contém seis exons e a seta indica a localização da variante identificada nos pacientes II.1 e II.3, no exon 5 (c.277C>G:p.Gln93Glu). (B) Diagrama da proteína UBE2A que contém um domínio catalítico de conjugação de ubiquitina (UBCc) - a seta indica a localização da variante Q93E (<http://smart.embl-heidelberg.de/>).

Figura 8. Imagem da superfície da proteína UBE2A. O resíduo catalítico C88 está indicado (esferas em azul), assim como o resíduo Q93, no qual a variante *missense* foi identificada (esferas vermelhas) na família 1 (PyMOL by Schrödinge - <https://pymol.org>).

A presença da variante em hemizigose foi confirmada por sequenciamento Sanger nos irmãos e em heterozigose na mãe (Figura 9A). O sequenciamento Sanger foi também realizado em material das três irmãs fenotipicamente normais, mostrando que todas são portadoras da variante em heterozigose (Figura 9B).

Figura 9. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante *missense* c.277C>G encontrada no gene *UBE2A*. (A) presença da variante em hemizigose nos irmãos afetados (II.1 e II.3) e em heterozigose na mãe (I.2) e (B) presença da variante nas irmãs heterozigotas (II.2, II.4 e II.5).

O padrão de inativação do cromossomo X foi investigado pela análise de metilação do loco do gene *AR* (*Androgen Receptor* – receptor de andrógenos (Allen et al. 1992), evidenciando que as portadoras heterozigotas apresentavam desvio quase total de inativação do cromossomo X em linfócitos de sangue periférico. A Figura 10 mostra o padrão de inativação do cromossomo X das células de sangue periférico da mãe dos afetados (95:5); nas irmãs II.2, II.4 e II.5 o desvio observado foi de 90:10, 93:7 e 94:6; respectivamente.

Após extração de mRNA total de sangue periférico dos afetados e de sua mãe, a mutação no gene *UBE2A* foi detectada apenas nas amostras de cDNA dos pacientes, enquanto na amostra de cDNA da mãe houve a detecção apenas do alelo sem a mutação (Figura 11). A presença na mesma família de homens afetados hemizigotos para a variante e mulheres portadoras fenotipicamente normais com desvio quase total de inativação do cromossomo X com a mutação corrobora a hipótese de que a mutação detectada no gene *UBE2A* seja patogênica.

Figura 10. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (I.2), evidenciando desvio total de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de apenas um dos alelos (1), mostrando inativação preferencial de um dos cromossomos X em sangue periférico (razão de inativação 95:5).

Figura 11. Sequenciamento Sanger em amostras de cDNA derivado de mRNA de sangue periférico mostrando a variante *missense* cDNA.451C>G no exon 5 do gene *UBE2A*. Acima, presença da variante nos irmãos afetados (II.1 e II.3); abaixo, ausência da variante no controle e na mãe dos afetados (I.2).

Discussão

Nesta família foi identificada em dois afetados por DI síndrômica uma nova variante *missense* no gene *UBE2A*, cujas mutações estão associadas à síndrome de DI ligada ao X tipo Nascimento. Desde 2006, quando a primeira família foi descrita por Nascimento e col., outras 22 famílias foram relatadas com indivíduos do sexo masculino com mutações ou deleções que afetam o gene *UBE2A*. Foram descritas nove famílias com deleções contendo a totalidade do gene *UBE2A* (Czeschik et al. 2013; de Leeuw et al. 2010; Honda et al. 2010; Thunstrom et al. 2015; Utine et al. 2014), duas famílias com microdeleções intragênicas (Czeschik et al. 2013; Tucker et al. 2014) e 12 famílias com mutações *missense*, *nonsense* e *frameshift* (Budny et al. 2010; Czeschik et al. 2013; Giugliano et al. 2018; Haddad et al. 2013; Nascimento et al. 2006; Niranjan et al. 2015; Tsurusaki et al. 2017; Tzschach et al. 2015). Alguns autores sugerem que pacientes com esta síndrome poderiam ser clinicamente diagnosticados, embora quase todos os indivíduos com DI do tipo Nascimento tenham sido identificados por métodos moleculares. No total, seis das mutações já descritas no gene *UBE2A* são *missense*. Na Tabela 4, temos um resumo do quadro clínico dos pacientes descritos neste trabalho e de outros cinco portadores de mutações *missense* com dados fenotípicos disponíveis na literatura. A DI e o comprometimento de fala são características apresentadas por todos os pacientes, enquanto outros sinais clínicos podem ou não estar presentes; o hirsutismo e as convulsões aparentemente são características menos observadas nos pacientes com mutações *missense*. Os pacientes aqui descritos não apresentam o quadro clínico típico da síndrome, principalmente as características faciais que estão presentes na maioria dos pacientes já descritos. A Figura 12 mostra as faces dos pacientes relatados por Nascimento e col. (2016), portadores da mutação *nonsense* Q128X, pacientes com a mutação *missense* G23R (Budny et al. 2010) e pacientes com a mutação *missense* P79R (Czeschik et al. 2013), todos apresentando uma face considerada típica da síndrome de DI ligada ao cromossomo X tipo Nascimento.

Tabela 4. Resumo do quadro clínico dos pacientes deste estudo e das famílias encontradas na literatura com síndrome de DI ligada ao X tipo Nascimento causada por mutação *missense* no gene *UBE2A*.

	Budny e col. (família 1)	Budny e col. (família 2)	Czeschik e col. (família B)	Haddad e col. (família 2)	Tsurusaki e col. (família 2)	Presente Estudo	Total
Mutação	p.G23R	p.R11Q	p.P79R	p.R7W	p.G26R	p.Q93E	6 variantes
Nº de pacientes	4	1	2	3	1	2	13
Desvio de inativação em portadoras	sim	n.a.	sim	sim	sim	sim	Em todas as mulheres portadoras investigadas
Deficiência intelectual	4/4	1/1	2/2	3/3	1/1	2/2	13/13 (100%)
Atraso motor	4/4	1/1	2/2	3/3	1/1	0 /2	11/13 (85%)
Comprometimento de fala	4/4	1/1	2/2	3/3	1/1	2/2	13/13(100%)
Dismorfismos faciais	4/4	1/1	2/2	3/3	1/1	0/2	11/13 (85%)
Alterações na pele	3/4	1/1	2/2	2/3	1/1	0/2	9/13 (69%)
Hirsutismo	3/4	1/1	0/2	0/3	n.a.	0/2	4/12 (33%)
Anomalias urogenitais	3/4	1/1	2/2	2/3	0/1	0/2	8/13 (62%)
Convulsões	3/4	0/1	1/2	0/3	0/1	0/2	4/13 (31%)

n.a= não avaliado.

Embora muitos autores sugiram que a ausência da proteína UBE2A e até mesmo mutações *missense* causariam um quadro clínico reconhecível, o número pequeno de pacientes descritos até agora limita o conhecimento sobre o espectro das características clínicas dos afetados. Um quadro clínico mais leve, como dos pacientes aqui descritos, é uma possível explicação para a subidentificação de pacientes com a síndrome de deficiência intelectual ligada ao X tipo Nascimento.

Figura 12. Dismorfismos faciais da síndrome de DI ligada ao cromossomo X do tipo Nascimento. A) Família com três homens afetados descrita por Nascimento e col., na qual a mutação c.382C>T (Q128X), que cria um códon de parada prematura, foi identificada como a causa da síndrome. B) Quatro afetados de uma mesma família, portadores da variante *missense* c.67G>A (G23R) e C) dois afetados de uma mesma genealogia portadores da variante *missense* c.236C>G (P79R).

Nos casos em que uma variante *missense* nova é identificada em um gene já associado a quadro clínico causado por mutações de perda de função da proteína, fica a dúvida se tal variante teria impacto clínico. Para investigar o impacto da variante *missense* detectada nesse estudo, foi realizado um trabalho em colaboração com a Dra. Juliana Ferreira de Oliveira, pesquisadora do grupo do Dr. Kleber Gomes Franchini, do Laboratório Nacional de Biociências – LNBio.

A variante Q93E está localizada próxima à cisteína catalítica de UBE2A e, portanto, os experimentos iniciais foram realizados para determinar se esta variante altera a atividade catalítica de UBE2A. Os resultados identificaram um prejuízo na aminólise pela substituição patogênica nessa posição por um ácido glutâmico. Os dados obtidos nos experimentos realizados com a proteína selvagem e a mutante Q93E demonstraram que a variante resulta em uma aminólise defeituosa, afetando o ambiente químico do sítio catalítico da UBE2A, sem perturbar a capacidade de UBE2A para conjugar com ubiquitina. Os resultados também mostraram que a atividade de aminólise do mutante Q93E pode ser restaurada em pH mais alto, fornecendo a primeira evidência de reversão do efeito de uma variante em *UBE2A*. Portanto, estes estudos demonstram que a variante aqui descrita é patogênica e que a reversão da atividade da proteína mutada, dependente do ambiente, poderia explicar o quadro clínico menos típico e mais leve encontrado nos pacientes dessa família (Oliveira e col. 2018, *Mechanistic insights revealed by a UBE2A mutation linked to intellectual disability, Nature Chemical Biology manuscript NCHEMB-A180407510A, in press*).

VI.3.2. FAMÍLIA 2 - PACIENTES COM FENÓTIPO INTERMEDIÁRIO ENTRE AS SÍNDROMES DE OPITZ-KAVEGGIA (FG) E DE LUJAN-FRYNS

Resumo

Neste estudo foi investigada a causa genética de quadro clínico de dois irmãos do sexo masculino afetados por DI leve (atraso do DNPM) e disgenesia de corpo caloso. A análise de sequenciamento de exoma revelou uma variante *missense* no gene *MED12* (OMIM*300188), que codifica a subunidade 12 do

mediador da transcrição de RNA polimerase II. Este gene está localizado em Xq13.1 e mutações e deleções em homens foram associadas a quadros clínicos variados que podem incluir, entre outros, DI, agenesia do corpo caloso, hipotonia e epilepsia (OMIM #309520 – LUJAN-FRYNS SYNDROME; OMIM #300895 – OHDO SYNDROME, X-LINKED, OHDOX e OMIM #305450 – OPITZ-KAVEGGIA SYNDROME, OKS).

A variante *missense* c.2179G>A (p.Asp727Asn; rs192331892) detectada nessa família afeta um aminoácido evolutivamente conservado e é muito rara, relatada nos bancos de dados ExAC e 1K Genomes com frequência de 0,00002 e 0,0003, respectivamente. As análises *in silico* utilizando os algoritmos FATHMM MKL e MutationTaster classificaram esta variante como danosa ou de impacto moderado, de acordo com o VEP (*Variant Effect Predictor*). Os dados em conjunto, isto é, variante muito rara e provavelmente danosa à função proteica, em gene já associado a DI ligada ao cromossomo X, com quadro clínico compatível com alteração em *MED12*, indicam fortemente a patogenicidade da variante em questão.

O gene *MED12*

O gene *MED12*, localizado em Xq13, é uma subunidade do complexo multiprotéico que tem papel crítico e central na transcrição da RNA polimerase II, podendo funcionar na ativação ou inibição da transcrição, dependendo dos fatores com os quais interage. Em humanos, mutações germinativas no gene *MED12* já foram associadas a doenças neuropsiquiátricas, incluindo esquizofrenia e psicose (Philibert and Madan 2007), assim como mutações somáticas foram descritas em leiomioma uterino e câncer de próstata (Barbieri et al. 2012; Makinen et al. 2011; Makinen et al. 2013). Adicionalmente, mutações germinativas neste gene causam três síndromes de DI ligada ao cromossomo X que compartilham vários sinais clínicos, abaixo descritas: síndrome de Opitz-Kaveggia (FG), síndrome de Lujan-Fryns e síndrome Ohdo ligada ao X. A proteína MED12 participa da regulação de sinais envolvidos no crescimento, diferenciação e desenvolvimento celulares e essa atuação poderia explicar seu papel em diferentes síndromes de DI (Ding et al. 2008; Graham and Schwartz 2013).

A síndrome de Opitz-Kaveggia (síndrome de FG, OMIM #305450), foi descrita pela primeira vez em 1974, em uma família de cinco homens apresentando DI, macrocefalia, ânus imperfurado e hipotonia (Opitz and Kaveggia 1974). Em 2007, Risheg e col. foram os primeiros a associar mutações no gene *MED12* à síndrome Opitz-Kaveggia, em seis famílias nas quais uma mutação no exon 21 [c.2881C>T (R961W)] segregava em homens afetados e mulheres portadoras. Inicialmente, a mesma mutação foi encontrada em duas de 24 famílias estudadas que apresentavam DI ligada ao X; os afetados portadores da mutação haviam sido diagnosticados com a síndrome de Opitz-Kaveggia; posteriormente, utilizando uma coorte de 41 indivíduos (29 probandos masculinos não aparentados e diagnosticados com a síndrome de Opitz-Kaveggia e 12 mulheres com filhos afetados com o mesmo diagnóstico) identificaram a mesma mutação em um homem afetado e em duas mulheres heterozigotas, portadoras obrigatórias. A mesma mutação foi também detectada na família originalmente descrita para a síndrome Opitz-Kaveggia (Opitz and Kaveggia 1974) (Risheg et al. 2007). Em 2011, foi descrita uma nova mutação associada à síndrome de Opitz-Kaveggia numa família com três homens afetados, hemizigotos para uma mutação *missense* c.2873G>A (p.G958E), também no exon 21 do gene *MED12* (Rump et al. 2011). Dessa forma, ficou evidenciado que mutações diferentes no exon 21 do gene *MED12* causam a síndrome Opitz-Kaveggia.

Schwartz e col. (2007) estudaram a família originalmente relatada por Lujan e col. (Lujan et al. 1984) e outra família, também com afetados pela síndrome de Lujan-Fryns (OMIM #309520), identificando em ambas uma mutação no exon 22 do gene *MED12*, c.3020A>G (p.N1007S). Esses resultados mostraram que as síndromes de Lujan-Fryns e de Opitz-Kaveggia, consideradas até então síndromes distintas de DI ligadas ao X, são resultado de heterogeneidade alélica, sendo causadas por diferentes mutações no gene *MED12* (Schwartz et al. 2007).

Na Tabela 5 temos um resumo das características clínicas observadas nas síndromes de Opitz-Kaveggia e Lujan-Fryn. O espectro fenotípico causado por mutações germinativas no gene *MED12* apresenta características clínicas como

agenesia de corpo caloso, macrocefalia, testa alta, hipotonia, ausência ou atraso de fala, DI e distúrbios comportamentais, que podem ou não estar presentes nos afetados de ambas as síndromes; embora tais sinais clínicos tenham servido de base para descrever duas síndromes distintas, mutações no gene podem estar associadas a fenótipos parcialmente superpostos às duas síndromes, como discutiremos mais à frente.

Tabela 5. Comparação das características clínicas das síndromes de Lujan e Opitz-Kaveggia.

Principais características	Lujan	Opitz-Kaveggia
Estatura elevada	+	-
Voz anasalada	+	-
Raiz nasal alta	+	-
Dedos longos e hiperextensíveis	+	-
Macrocefalia	+	+/-
Polegares largos	+/-	+
Implantação frontal alta do cabelo	-	+
Hipertelorismo/telecanto	-	+
Orelhas pequenas	-	+
Pregas palmares horizontais ligadas	-	+
Sindactilia	-	+
Anomalias anais	-	+
Constipação infantil	-	+
Problemas urogenitais	-	+
Espasticidade	-	+
Agenesia de corpo caloso	+	+
Testa alta proeminente	+	+
Inclinação ascendente das fissuras palpebrais	+	+
Estrabismo	+	+
Hipoplasia maxilar	+	+
Má-oclusão dentária	+	+
Palato estreito/alto	+	+
Dentes amontoados	+	+
Micrognatia/retrognatia	+	+
Hipotonia	+	+
Convulsões	+	+
Deficiência intelectual	+	+
Distúrbios comportamentais	+	+

Baseado em Schwartz et al. (2007).

Verloes e col. (2006) classificaram as síndromes de blefarofimose associadas com DI em cinco subgrupos distintos, uma delas foi denominada síndrome de Ohdo ligado ao X ou tipo Maat-Kievit-Brunner (OSMKB; OMIM #300895) (Verloes et al. 2006). Vulto-van Silfhout e col. (2013) realizaram o sequenciamento de exoma em duas famílias com afetados do sexo masculino diagnosticados com a síndrome de Ohdo (OSMKB); foram detectadas mutações *missense* no gene *MED12* [c.3443G>A (p.Arg1148His) no exon 24, e c.3493T>C (p.Ser1165Pro) no exon 25], que segregavam com o fenótipo. Após esta descoberta, uma coorte de nove homens com o diagnóstico clínico da síndrome tiveram os 45 exons do gene *MED12* sequenciados pela técnica de Sanger e uma terceira mutação *missense*, no exon 37, foi detectada [c.5185C>A (p.His1729Asn)]. Assim, a detecção de três mutações *missense* diferentes em três famílias não relacionadas mostrou que mutações em *MED12* podem também causar a síndrome de Ohdo (Vulto-van Silfhout et al. 2013).

Recentemente, outros fenótipos foram relacionados a mutações em *MED12*, em especial DI não-sindrômica. Lesca e col. (2013) relataram uma mutação de mudança de quadro de leitura (*frameshift*) c.5898dupC no exon 41 do gene *MED12* em uma família com vários afetados por DI profunda, incluindo uma mulher; outras nove mulheres da família portadoras da mutação apresentavam diferentes graus de comprometimento intelectual (Lesca et al. 2013). Em 2015, dois trabalhos relataram novas mutações *missense* no gene *MED12* (c.3067^a>G, exon 22, c.5922G>T, exon 41), detectadas em pacientes considerados não sindrômicos (Bouazzi et al. 2015; Yamamoto and Shimojima 2015). Em 2016, outra família com dois irmãos e uma sobrinha foi relatada, todos com DI e a mutação *missense* c.2312T>C (p.Ile77Thr, exon 16); nessa família, os pacientes do sexo masculino apresentavam um fenótipo semelhante ao descrito por Lesca e col. (2013), sugerindo a existência de um quarto distúrbio relacionado ao gene *MED12* (Prontera et al. 2016).

Em 2018, Charzewska e col. reportaram seis famílias com variantes *missense* no gene *MED12*. Três são variantes novas: c.2861T>G (p.Val954Gly, exon 21), c.3271G>A (p.Glu1091Lys, exon 23) e c.4111C>T (p.Pro1371Ser, exon 29); uma variante reportada em pacientes com a síndrome de Ohdo e relatada

pela primeira vez em mulheres afetadas c.3443G>C (p.Arg1148His, exon 24). Outras duas variantes c.2444G>A (p.Arg815Gln, exon 18) e c.3883C>T (p.Arg1295Cys, exon 28) foram descritas em outro estudo dos autores, mas sem detalhes clínicos das famílias. Estes autores expandiram o espectro fenotípico associado ao gene *MED12* e utilizaram a modelagem molecular como uma importante abordagem *in silico* para avaliar os efeitos funcionais das variantes detectadas (Charzewska et al. 2018).

Neste trabalho, nós relatamos uma família com dois irmãos do sexo masculino afetados por DI leve, dismorfismos faciais e disgenesia de corpo caloso. Nossos dados sugerem que a variante *missense* no exon 15 do gene *MED12*, identificada por sequenciamento completo de exoma, é responsável pelo quadro clínico.

Descrição da Família 2 (pacientes 2a e 2b)

A Família 2 foi encaminhada pela Dra. Juliana Gurgel Giannetti. Os pacientes são nascidos de pais saudáveis e não consanguíneos e têm três irmãs não afetadas. O heredograma desta família é mostrado na Figura 13.

Figura 13. Heredograma da Família 2, na qual uma variante provavelmente patogênica no gene *MED12* segrega com DI num padrão de herança ligada ao cromossomo X.

Paciente 2a (II.4)

Nascimento a termo de parto normal, com 3.300g. Aos 6 meses foram notadas hipotonia e hipoatividade. Apresentou atraso de DNPM, tendo sentado depois de 1 ano de idade e andado com 2 anos de idade. Aos 3 anos, em avaliação com neuropediatra devido a irritabilidade, foi diagnosticada epilepsia em exame de EEG. Exame de ressonância magnética de encéfalo revelou disgenesia de corpo caloso e alteração volumétrica em hipocampo e cerebelo. Examinado aos 25 anos de idade (Figura 14), apresenta DI leve, hipertelorismo ocular e estrabismo divergente, orelhas pequenas, filtro curto e narinas antevertidas, hipoplasia malar e ângulo da mandíbula proeminente. O paciente tem clinodactilia de 5º dedo à direita e nasceu com dedo extranumerário bilateral, retirado cirurgicamente. Tem altura de 177,5cm, peso de 66,8kg e PC de 55cm.

Figura 14. (A), (B) e (C) Características faciais do paciente 2a (II.4), que incluem hipertelorismo e estrabismo divergente, orelhas pequenas, hipoplasia malar, filtro curto e mandíbula proeminente. (D) Imagem de corpo inteiro do paciente. (E) e (F) mostram a clinodactilia de 5º dedo.

Paciente 2b (II.5)

Nascimento a termo de parto normal, com 3.000g. Atraso de DNPM foi notado aos 6 meses, quando ainda não apresentava sustentação cefálica; andou com mais de 2 anos de idade. Exame de ressonância magnética de encéfalo revelou disgenesia de corpo caloso associado a redução volumétrica cerebelar. Examinado aos 22 anos de idade (Figura 15), apresenta DI leve, braquicefalia direita, orelhas pequenas, estrabismo convergente à esquerda, filtro nasal curto, má-oclusão dentária, palato alto e estreito, proeminência do ângulo da mandíbula com aumento de partes moles adjacentes, retrognatia, constipação intestinal crônica, polegares largos, clinodactilia de 5º dedo bilateral e dedo extranumerário em uma das mãos, que foi retirado. Tem altura de 174,5cm, peso de 61,8kg e PC de 56,5cm.

Figura 15. (A) e (B) Características faciais do paciente 2b (II.5), que incluem braquicefalia à direita, hipertelorismo e estrabismo convergente à esquerda, orelhas pequenas, filtro nasal curto, má-oclusão dentária e mandíbula proeminente. (C) Imagem de corpo inteiro do paciente. (D) e (E) Clinodactilia de 5º dedo e polegares largos.

Resultados

A análise do sequenciamento de exoma completo dos pacientes resultou na identificação em ambos de uma variante em hemizigose no exon 15 do gene *MED12* [NM_005120:c.2179G>A:p.(Asp727Asn)]. Esta variante afeta um aminoácido evolutivamente conservado e foi reportada em bancos de dados populacionais como ExAC e 1K Genomas com frequências de 0,00002 e 0,0003, respectivamente (rs192331892). No banco de dados gnomAD (que contém 123.136 exomas) há informação da variante em um único indivíduo hemizigoto. Nestes bancos populacionais, não há informação clínica sobre portadores e essa variante foi reportada exclusivamente em populações africanas. A patogenicidade da variante foi avaliada por análises *in silico* e os algoritmos FATHMM MKL e MutationTaster a classificaram como “danosa” ou prejudicial à função proteica, enquanto o VEP atribui “impacto moderado”.

A presença da variante no gene *MED12* foi confirmada por sequenciamento Sanger nos irmãos e na mãe, heterozigota portadora obrigatória (Figura 16A). O sequenciamento Sanger foi também realizado nas três irmãs dos afetados, clinicamente normais; duas delas são portadoras da variante em heterozigose (Figura 16B).

Conforme revelado pela análise do padrão de metilação do loco do gene *AR*, as mulheres portadoras da variante apresentam desvio parcial de inativação do cromossomo X em linfócitos de sangue periférico. A Figura 17 mostra o padrão de inativação do cromossomo X da mãe dos afetados, estimado na proporção de 70:30; na irmã portadora II.1 o desvio foi avaliado em 80:20 e na irmã II.2 não foi possível determinar a proporção entre os alelos, pois ela é homozigota para o loco *AR*. A irmã não portadora (II.3) não tem desvio de inativação do cromossomo X (49:51). O desvio de inativação parcial encontrado em duas das portadoras está de acordo com o que já foi descrito na literatura, pois portadoras de variante *missense* do gene *MED12* podem apresentar desvio parcial de inativação do cromossomo X (Bouazzi et al. 2015; Risheg et al. 2007; Schwartz et al. 2007).

Figura 16. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante *missense* c.2179G>A encontrada no exon 15 do gene *MED12*. (A) presença da variante em hemizigose nos irmãos afetados (II.4 e II.5) e em heterozigose na mãe portadora (I.2) e (B) presença da variante nas irmãs portadoras (II.1 e II.2) e detecção apenas do alelo normal na irmã II.3.

Figura 17. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados, evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de apenas um dos alelos (1), mostrando inativação preferencial de um cromossomo X em aproximadamente 70% das células de sangue periférico (razão de inativação 70:30).

Discussão

Nesta família, foi detectada em dois afetados por DI sindrômica uma variante *missense* no gene *MED12*, que codifica a subunidade mediadora 12 de transcrição da RNA polimerase II e desempenha um papel importante na ativação da transcrição gênica em eucariotos (Wang et al. 2013). A proteína MED12 faz parte de um complexo proteico envolvido em uma variedade de atividades celulares.

Diferentes mutações *missense* no gene *MED12* foram associadas a três diferentes síndromes (Opitz-Kaveggia, Lujan-Fryns e Ohdo) (Risheg et al. 2007; Rump et al. 2011; Schwartz et al. 2007; Vulto-van Silfhout et al. 2013). Duas diferentes mutações no exon 21 foram detectadas em pacientes com a síndrome de Opitz-Kaveggia, ao passo uma única mutação no exon 22 foi identificada em pacientes com síndrome de Lujan. Três mutações nos exons 24, 25 e 37 foram identificadas na síndrome de Ohdo. Outras mutações *missense* foram detectadas nos exons 16, 22 e 41 e os indivíduos não apresentavam quadro clínico compatível com o diagnóstico das síndromes já associadas ao gene *MED12* (Bouazzi et al. 2015; Prontera et al. 2016; Yamamoto and Shimojima 2015). As síndromes de Lujan-Fryns e Opitz-Kaveggia têm várias características clínicas em comum: hipoplasia maxilar, má-oclusão dentária, palato alto e estreito, dentes amontoados e micrognatia/retrognatia. Defeitos anais são observados apenas nos pacientes com a síndrome de Opitz-Kaveggia. Tais malformações craniofaciais e do sistema digestivo podem derivar de defeitos endodérmicos primários. No trabalho de Shin e col. (2008), estudos realizados com *zebrafish* mutantes para *med12* revelaram defeitos do desenvolvimento da endoderme que podem explicar as alterações craniofaciais e no sistema digestivo observadas nos afetados pelas síndromes associadas ao gene *MED12* (Shin et al. 2008).

Recentemente, Charzewska e col. (2018) reportaram seis famílias com diferentes mutações *missense* no gene *MED12*. Uma dessas famílias foi identificada com uma mutação já descrita em pacientes do sexo masculino com a síndrome de Ohdo (c.3443G>A) e, pela primeira vez, reportada em duas mulheres afetadas que herdaram a variante de sua mãe portadora assintomática. Foi realizado estudo de inativação do X, porém os autores não encontraram uma

correlação clara entre o quadro clínico das portadoras e seu padrão de inativação do X em células do sangue, semelhante aos resultados publicados na literatura (Charzewska et al. 2018).

Os pacientes do presente estudo, cuja variante foi identificada no exon 15 do gene *MED12*, compartilham algumas características clínicas relatadas nos pacientes com síndromes de Lujan-Fryns ou Opitz-Kaveggia, destacando-se a DI e a disgenesia de corpo caloso. A blefarofimose, que é uma característica frequente na síndrome de Ohdo ligada ao X, não foi detectada em nossos pacientes. A comparação das características clínicas dos irmãos da Família 2 com os sinais já descritos nos afetados pelas síndromes relacionadas ao *MED12* é apresentada na Tabela 6.

Tabela 6. Comparação das características clínicas descritas em afetados pelas síndromes associadas a mutações no gene *MED12* (Lujan-Fryns, Opitz-Kaveggia e Ohdo) e nos irmãos afetados da Família 2.

	Lujan-Fryns	Opitz-Kaveggia	Ohdo	II.4	II.5
Neurológico					
Deficiência intelectual	+	+	+	+	+
Distúrbios comportamentais	+	+	+	+	+
Agenesia do corpo caloso	+	+	–	+	+
Hipotonia (infância)	+	+	+	+	–
Convulsões	+	+	–	–	–
Crescimento					
Estatura elevada	+	–	–	–	–
Macrocefalia	+	+	–	+	–
Craniofacial					
Face alongada e estreita	+	–	–	–	–
Face triangular	–	–	+	–	–
Testa proeminente/alta	+	+	+	–	–
Hipertelorismo/telecanto	–	+	–	–	–
Estrabismo	+	+	+	+	+
Face grosseira	–	–	+	–	–
Inclinação ascendente das fissuras palpebrais	+	+	+	+	+
Blefarofimose	–	–	+	–	–
Raiz nasal alta	+	–	–	–	–
Palato estreito/alto	+	+	–	–	+
Dentes amontoados	+	+	–	–	+
Hipoplasia maxilar	+	+	–	+	+
Micrognatia /retrognatia	+	+	+	+	+
Má-oclusão dentária	+	+	+	–	+
Implantação frontal alta do cabelo	–	+	–	–	–
Orelhas pequenas	–	+	+	+	+
Mãos					
Dedos longos e hiperextensíveis	+	–	+	–	–
Polegares largos	+	+	–	–	+
Pregas palmares horizontais	–	+	–	–	–
Sindactilia	–	+	–	–	–
Gastrointestinal					
Defeitos anais	–	+	–	–	–
Defeitos genitais	–	+	–	–	–
Constipação crônica	–	+	+	–	+
Fala					
Voz anasalada	+	–	–	–	–
Pobre ou ausente	–	–	+	–	–

A variante c.2179G>A identificada na família 2 foi considerada uma nova variante provavelmente patogênica pelas seguintes características: apresenta frequência muito baixa na população em geral, com apenas um hemizigoto relatado; avaliada como prejudicial à função proteica por três diferentes algoritmos; segregação com o fenótipo; e grande superposição do quadro clínico dos pacientes com o espectro clínico dos distúrbios associados a mutações em *MED12*, em especial com as síndromes de Lujan-Fryns e Opitz-Kaveggia.

Uma questão a ser considerada é que os raros casos de portadores dessa variante foram documentados em populações africanas, que coincide com a ancestralidade da família. Entretanto, mesmo nessas populações a frequência é muito baixa. Também é importante ressaltar que a segregação da variante na família revelou que a mãe e duas irmãs são portadoras assintomáticas e apresentam desvio parcial de inativação do cromossomo X, compatível com o descrito na literatura para mutações *missense* nesse gene. Embora o gene *MED12* esteja sujeito à inativação do X, na literatura não encontramos uma correlação clara entre os fenótipos clínicos das mulheres portadoras de mutações *missense* do gene *MED12* e o padrão de inativação em sangue periférico; obviamente, diferentes padrões de inativação do X podem ser encontrados em outros tecidos.

Um aspecto importante a ser mencionado é o fato de que apenas mutações *missense* foram documentadas no gene *MED12* em associação a distúrbios de desenvolvimento e DI síndrômica e não-sindrômica, com exceção de um único caso. Isso levanta a questão referente ao efeito de tais mutações *missense* na função gênica, sugerindo um efeito possível de ganho de função, e não de perda ou haploinsuficiência. Também podemos supor que a variabilidade fenotípica encontrada seria causada pelas diferentes localizações das mutações no gene. Os pacientes relatados em nosso estudo não se enquadram perfeitamente em nenhuma das três síndromes associadas a *MED12*, mas compartilham de muitos sinais clínicos nelas descritos.

Esse é o primeiro relato de DI síndrômica associada a uma nova variante *missense* no exon 15 de *MED12*, em quadro clínico intermediário entre as síndromes de Opitz-Kaveggia (FG) e de Lujan-Fryns.

VI.3.3. FAMÍLIA 3 – NOVAS VARIANTES PATOGÊNICAS NO GENE *CDK5RAP2* ASSOCIADAS À MICROCEFALIA PRIMÁRIA AUTOSSÔMICA RECESSIVA 3 (MCPH3)

Resumo

A irmandade aqui descrita é composta por dois irmãos do sexo masculino afetados por DI leve, microcefalia congênita perda auditiva e manchas café-com-leite distribuídas pelo corpo. A análise de sequenciamento de exoma revelou duas variantes no gene *CDK5RAP2* (OMIM* 608201), que codifica uma proteína localizada no centrôssomo e no complexo de Golgi, também descrita como reguladora da atividade do gene *CDK5*. A sequência do gene *CDK5RAP2* está localizada em 9q33.2 e mutações em homozigose ou heterozigose composta foram associadas à microcefalia primária autossômica recessiva 3 (MCPH3; OMIM #309520 – Microcephaly 3, Primary, Autosomal Recessive).

As duas variantes identificadas na irmandade são do tipo *frameshift*: c.564_565dupGA (p.Lys189fs) e c.1246delG (p.Glu416fs), ocasionando perda de função na proteína. Cada uma delas foi herdada de um dos genitores, indicando heterozigose composta e herança recessiva. Não estão documentadas em nenhum dos bancos consultados de variantes genéticas populacionais, tampouco foram descritas em indivíduos já identificados com mutações em *CDK5RAP2*. Os dados em conjunto, como a ausência dessas variantes em bancos populacionais e seu efeito de perda função em gene já associado a DI e microcefalia de herança recessiva, indicam fortemente sua patogenicidade.

O gene *CDK5RAP2* como causa de microcefalia primária autossômica recessiva

A microcefalia é caracterizada pela redução de 3-4 desvios-padrão (SD) da circunferência da cabeça em comparação a indivíduos de mesmo sexo e idade, apresentando principalmente a diminuição do córtex cerebral (Ahmad et al. 2017; Faheem et al. 2015; Tan et al. 2014). O termo microcefalia primária autossômica recessiva (MCPH) abrange um conjunto de condições clínicas diferentes, primordialmente assim classificadas por apresentarem microcefalia congênita, em indivíduos que em decorrência irão apresentar DI leve a moderada, em geral sem outras co-morbidades associadas (não-sindrômicas). As estimativas de

incidência de MCPH são muito variáveis, a depender da população estudada e do critério de inclusão dos pacientes nessa categoria, variando de 1:30.000-1:250.000 nascidos vivos (Zaqout et al. 2017). Tais condições clínicas são doenças raras e não-progressivas do neurodesenvolvimento. Por outro lado, a microcefalia secundária refere-se a condições clínicas nas quais o perímetro cefálico é normal ao nascimento e a microcefalia se desenvolve no período pós-natal. Como acima mencionado, o grupo MCPH é geneticamente heterogêneo e, de acordo com revisão de Zaqout e col. (2017) (Tabela 7), 17 *loci* foram associados a essa condição (Zaqout et al. 2017); muitos dos genes descritos têm sido implicados na divisão celular e na regulação do ciclo celular (Hussain et al. 2012) durante o desenvolvimento do cérebro.

Tabela 7. Genes conhecidos de microcefalia primária autossômica recessiva (MCPH) de acordo com a revisão de Zaqout e col. (2017).

Locus	Proteína	Gene	Localização	OMIM
MCPH1	Microcephalin 1	<i>MCPH1</i>	8p23.1	607117
MCPH2	WD-repeat-containing protein 62	<i>WDR62</i>	19q13.12	613583
MCPH3	Cyclin-dependent kinase 5 regulatory subunit-associated protein 2	<i>CDK5RAP2</i>	9q33.2	608201
MCPH4	Kinetochore scaffold 1	<i>KNL1</i>	15q15.1	609173
MCPH5	Abnormal spindle-like, microcephaly associated protein	<i>ASPM</i>	1q31.3	605481
MCPH6	Centromeric protein J	<i>CENPJ</i>	13q12.2	609279
MCPH7	SCL/TAL1-interrupting locus protein	<i>STIL</i>	1p33	181590
MCPH8	Centrosomal protein 135 kD	<i>CEP135</i>	4q12	611423
MCPH9	Centrosomal protein 152 kD	<i>CEP152</i>	15q21.1	613529
MCPH10	Zinc finger protein 335	<i>ZNF335</i>	20q13.12	610827
MCPH11	Polyhomeotic-like 1 protein	<i>PHC1</i>	12p13.31	602978
MCPH12	Cyclin-dependent kinase 6	<i>CDK6</i>	7q21.2	603368
MCPH13	Centromeric protein E	<i>CENPE</i>	4q24	117143
MCPH14	SAS-6 centriolar assembly protein	<i>SASS6</i>	1p21.2	609321
MCPH15	Major facilitator superfamily domain-containing protein 2A	<i>MFSD2A</i>	1p34.2	614397
MCPH16	Ankyrin repeat- and LEM domain-containing protein 2	<i>ANKLE2</i>	12q24.33	616062
MCPH17	Citron rho-interacting serine/threonine kinase	<i>CIT</i>	12q24.23	605629

O gene *CDK5RAP2*, que quando mutado causa a MCPH3, codifica uma proteína reguladora da atividade de *CDK5*, localizada no complexo de Golgi e centrôssomos em células do córtex cerebral (Faheem et al. 2015). A proteína contribui para regular a função dos microtúbulos e mutações inativadoras de *CDK5RAP2* diminuem o número de neurônios, assim como reduzem a sobrevivência celular (Faheem et al. 2015; Issa et al. 2013a; Kraemer et al. 2011).

Em 2005, Bond e col. identificaram pela primeira vez mutações em homozigose (p.S81X e p.E385fsX4) no gene *CDK5RAP2* em duas famílias consanguíneas com MCPH, do Paquistão (Bond et al. 2005). Uma nova mutação em homozigose (p.E234X) foi encontrada em uma família consanguínea da Somália; a criança afetada apresentava microcefalia primária associada a perda de audição neurosensorial (Pagnamenta et al. 2012). Tan e col. (2014) relataram a primeira família com mutações em heterozigose composta no gene *CDK5RAP2* que segregavam com o fenótipo de microcefalia.

Até o presente, 15 mutações diferentes no gene *CDK5RAP2* foram descritas, todas de perda de função: oito mutações em homozigose em famílias consanguíneas (Abdullah et al. 2017; Ahmad et al. 2017; Bond et al. 2005; Hassan et al. 2007; Issa et al. 2013b; Pagnamenta et al. 2012; Yigit et al. 2015) e sete em heterozigose composta (Lancaster et al. 2013; Li et al. 2015; Pagnamenta et al. 2016; Tan et al. 2014).

Neste trabalho, nós relatamos uma família com dois irmãos do sexo masculino afetados por DI leve, microcefalia congênita, manchas café-com-leite e deficiência auditiva, heterozigotos compostos para duas novas variantes patogênicas, do tipo *frameshift*, mapeadas nos exons 7 e 12 do gene *CDK5RAP2*.

Descrição da Família 3 (pacientes 3a e 3b)

A Família 3 foi encaminhada pela geneticista clínica Dra. Gisele Santos de Oliveira, de Curitiba - PR. Os pacientes são nascidos de pais saudáveis e não consanguíneos. O heredograma desta família é mostrado na Figura 18.

Figura 18. Heredograma da Família 3, na qual duas variantes diferentes do tipo *frameshift* foram detectadas no gene *CDK5RAP2* em heterozigose composta, segregando com DI num padrão de herança autossômica recessiva.

Paciente 3a (I.1)

Nascimento de parto normal, com 35 semanas de gestação. Ao nascimento apresentava PC de 31cm, que evoluiu aos 9 anos e 9 meses para 44cm. Examinado aos 15 anos de idade (Figura 19), apresenta DI leve, perda de audição neurosensorial, microcefalia (PC de 48,5cm), nariz bicudo com ponte alta, lábios volumosos, diastema superior e dentes amontoados, palato alto e estreito e micrognatia. Apresentava dez manchas café-com-leite maiores que 0,5cm.

Figura 19. Face do paciente 3a (I.1).

Paciente 3b (I.2)

Nascimento a termo de parto normal. Examinado com 1 anos e 4 meses de idade apresentava PC de 40,5cm. Examinado aos 6 anos e 10 meses de idade (Figura 20), apresentava DI leve, microcefalia (PC de 44,5cm), nariz bicudo com ponte alta, diastema superior, palato alto e estreito, micrognatia; adicionalmente, tinha 20 manchas café-com-leite maiores que 0,5cm.

Figura 20. Face do paciente 3b (I.2).

Resultados

A análise dos dados do exoma resultou na identificação em ambos os irmãos de duas diferentes variantes do tipo *frameshift* no gene *CDK5RAP2*, localizadas nos exons 7 [NM_001011649.2:c.564_565dupGA:p.(Lys189fs)] e 12 [NM_001011649.2:c.1246delG:p.(Glu416fs)], respectivamente. Estas variantes não estão relatadas em nenhum dos bancos de dados populacionais pesquisados.

As variantes no gene *CDK5RAP2* foram confirmadas por sequenciamento Sanger nos irmãos. Cada um dos genitores é heterozigoto para uma das variantes, confirmando que se encontram nos afetados em heterozigose composta: variante detectada no exon 7 foi herdada da mãe e a variante no exon 12 foi herdada do pai (Figura 21).

Figura 21. Sequenciamento Sanger em amostras de DNA genômico mostrando as variantes identificadas em heterozigose composta no gene *CDK5RAP2*. (A) variante *frameshift* em heterozigose c.564_565dupGA no exon 7 do gene *CDK5RAP2*, presente nos irmãos (II.1 e II.2) e na mãe (I.2). (B) variante *frameshift* em heterozigose c.1246delG no exon 12 do gene *CDK5RAP2*, presente nos irmãos (II.1 e II.2) e no pai (I.1).

Discussão

Nesta família, identificamos em uma irmandade afetada por DI e microcefalia primária duas novas variantes *frameshift* no gene *CDK5RAP2*. Os pacientes têm em comum microcefalia congênita (<<2º percentil) associada a DI e dismorfismos faciais. A proteína CDK5RAP2 tem um importante papel na divisão celular pois regula a atuação dos microtúbulos, em especial na segregação dos cromossomos; no cérebro, altos níveis de expressão da proteína são observados

em regiões de rápida proliferação celular, particularmente no neuroepitélio (Issa et al. 2013a; Kraemer et al. 2011; Lizarraga et al. 2010). Mutações em homozigose ou heterozigose composta no gene *CDK5RAP2* causam microcefalia primária autossômica recessiva tipo 3 (MCPH3), sendo uma das mais raras causas de MCPH. Os indivíduos com mutação no gene *CDK5RAP2* têm como principal característica fenotípica a microcefalia congênita associada a DI leve-moderada.

Bond e col. (2005) descreveram os primeiros pacientes com mutações em *CDK5RAP2* em duas famílias consanguíneas do Paquistão (Bond et al. 2005). Todos os indivíduos afetados apresentavam microcefalia congênita e variados graus de DI. Em 2007, uma terceira família consanguínea do Paquistão com fenótipo semelhante foi descrita (Hassan et al. 2007). Outros relatos foram descritos em famílias com microcefalia primária, consanguíneas ou não (Issa et al. 2013b; Jouan et al. 2016; Lancaster et al. 2013; Li et al. 2015; Pagnamenta et al. 2016; Pagnamenta et al. 2012; Tan et al. 2014).

O quadro clínico de nossos pacientes é compatível com o descrito na literatura, uma vez que mutações em homozigose ou heterozigose composta do gene *CDK5RAP2* estão associadas a DI, microcefalia congênita e distúrbios faciais. Alguns pacientes com mutações em *CDK5RAP2* apresentam perda de audição neurosensorial (Bond et al. 2005; Lancaster et al. 2013), característica verificada em um dos indivíduos desta família. As manchas café-com-leite, presente em grande número nos irmãos aqui estudados, foram também descritas em outros pacientes (Pagnamenta et al. 2016; Yigit et al. 2015); outros estudos também relatam pacientes apresentando áreas do corpo com alterações de pigmentação (Bond et al. 2005; Issa et al. 2013b).

Portanto, relatamos aqui duas novas variantes (heterozigose composta) patogênicas de perda de função em *CDK5RAP2* associadas à MCPH3, em padrão de herança autossômica recessiva.

VI.3.4. FAMÍLIA 6 – NOVA VARIANTE PROVAVELMENTE PATOGENICA NO GENE *SLC6A8* QUE CAUSA DI RECESSIVA LIGADA AO CROMOSSOMO X

Resumo

A irmandade aqui estudada é composta por três irmãos, dois do sexo masculino e uma do sexo feminino, todos apresentando DI moderada, dificuldade de fala, convulsões e comportamento agressivo. No sequenciamento de exoma foi identificada uma variante *missense* no gene *SLC6A8* (MIM *300036), localizado em Xq28, que codifica o transportador de creatina 1 dependente de sódio e cloro. Mutações e deleções do gene *SLC6A8* foram associadas a quadros clínicos variados em homens, que podem incluir DI, atraso de desenvolvimento, atraso de fala, epilepsia e comportamento agressivo, dentre muitos outros sinais (OMIM #300352 – Cerebral Creatine Deficiency Syndrome 1; CCDS1). As mulheres heterozigotas para tais mutações *SLC6A8* podem apresentar quadro clínico, porém mais leve.

A variante *missense* c.205G>C detectada nos irmãos aqui descritos não está relatada em nenhum dos bancos consultados de variantes genéticas populacionais. As análises *in silico* da variante classificaram esta variante como “danosa” (algoritmos SIFT, FATHMM MKL e Mutation Taster), “provavelmente danosa” (Polyphen-2) ou de impacto moderado (VEP). Os dados em conjunto, variante rara em gene já associado a DI ligada ao X (*SLC6A8*) e com previsão *in silico* consistente de provável patogenicidade, detectada em irmandade com quadro clínico compatível com o gene, sugerem fortemente a patogenicidade da variante em questão.

O gene *SLC6A8* e a deficiência de creatina resultando em DI sindrômica

As síndromes de deficiência cerebral de creatina (CCDS) são caracterizadas por defeitos congênitos no metabolismo da creatina. Das três síndromes descritas, duas apresentam padrão de herança autossômica recessiva e são causadas por genes que afetam a biossíntese de creatina (*AGAT*, arginina-glicina amidinotransferase, OMIM *602360; *GAMT*, metiltransferase do ácido guanidinoacético, OMIM *601240), enquanto a terceira é causada por defeitos no transportador de creatina 1, sendo ligada ao cromossomo X (*SLC6A8*; MIM

*300036) e afetando o transporte de creatina para o cérebro e músculo (Stockler-Ipsiroglu and van Karnebeek 2014). Localizado em Xq28 (Gregor et al. 1995), o gene *SLC6A8* é expresso na maioria dos tecidos, com níveis mais altos verificados em cérebro e trato gastrointestinal (The human protein atlas; <https://www.proteinatlas.org/ENSG00000130821-SLC6A8/tissue>).

Essa proteína transportadora de creatina localiza-se na membrana celular e é responsável pela captação da creatina do meio extracelular para o citoplasma e vice-versa. A creatina é conhecida principalmente por seu papel essencial no metabolismo energético celular, sendo convertida no compartimento intracelular não enzimaticamente em creatinina, que se difunde para ser excretada na urina (van de Kamp et al. 2014).

Salomons e col. (2001) relataram pela primeira vez que mutações no gene *SLC6A8* causavam a síndrome de deficiência de creatina cerebral 1 (Cerebral Creatine Deficiency Syndrome 1; CCDS1; MIM #300352). Foi observado nos afetados a ausência do sinal de creatina nas imagens de espectroscopia de ressonância magnética de prótons cerebrais (H-MRS) e a elevação da razão creatina/creatinina na urina. Neste trabalho, a mutação c.1539C>T que gera um códon de parada prematura UAG (R514>X) foi detectada em dois homens e em três mulheres de uma mesma família; os homens apresentavam DI leve e grave atraso de fala, duas mulheres portadoras, avó e mãe do propósito, com dificuldade de aprendizado, e uma tia portadora era assintomática. Os autores não conseguiram associar o fenótipo das portadoras com a expressão do alelo selvagem ou mutado, visto que em linhagem linfoblastoide da mãe o alelo selvagem e o alelo mutado eram igualmente expressos, indicando estabilidade do alelo com a mutação *nonsense*. Entretanto, não foi detectada a expressão do alelo mutante em fibroblastos das três portadoras (Salomons et al. 2001). Posteriormente, outros estudos reforçaram o papel de mutações no gene *SLC6A8* em quadros clínicos similares, indicando este gene como causa frequente de DI recessiva ligada ao X.

Neste trabalho, nós relatamos uma família com três irmãos, dois do sexo masculino e uma do sexo feminino, todos afetados em graus diferentes por DI moderada, dificuldade de fala, convulsões e comportamento agressivo, em

quadro clínico provavelmente causado pela variante *missense* identificada no gene *SLC6A8*.

Descrição da Família 6 (Pacientes 6a, 6b e 6c)

A família aqui apresentada foi convidada a participar do projeto por ser composta por três irmãos afetados por DI, nascidos de pais saudáveis e não consanguíneos. O heredograma desta família é mostrado na Figura 22.

Figura 22. Heredograma da Família 6, na qual foi identificada uma variante *missense* no gene *SLC6A8* classificada como provavelmente patogênica, segregando com DI num padrão de herança ligada ao cromossomo X.

Paciente 6a (II.1)

Nasceu a termo, de parto normal. Apresentou convulsões desde 8 meses de idade (controladas com medicação), andou aos 13 meses e iniciou a fala aos 10 anos. Examinado aos 25 anos de idade, apresentava DI moderada associada a comportamento agressivo, com dificuldade de fala (troca de letras e fala repetitiva). Exibia dismorfismos faciais tais com cristas supraorbitárias proeminentes, pálpebras encobertas, nariz alongado e filtro nasal curto. Tem altura de 174,5cm, peso de 85kg e PC de 56,8cm (Figura 23A), cabelos grisalhos a despeito da idade.

Paciente 6b (II.2)

Nasceu prematuro aos 8 meses de parto cesárea e teve septicemia bacteriana do recém-nascido. Apresentou convulsões desde 8 meses de idade (controladas com medicação) e atraso de DNMP, tendo andado aos 2 anos e 6 meses e com início de fala aos 11 anos. Examinado aos 20 anos de idade, apresentava DI moderada associada a comportamento agressivo, fala muito comprometida e hiperatividade. Tem dismorfismos faciais tais com cristas supraorbitárias proeminentes, pálpebras encobertas e nariz alongado e bulboso com filtro nasal curto, altura de 186cm, peso de 105kg e PC de 58cm (Figura 23B).

Paciente 6c (II.3)

Nasceu a termo, de parto normal. Sentou com quase 1 ano e andou aos 2 anos e 6 meses, aos 7 anos falava palavras isoladas; teve menarca com 11 anos e meio. Examinada aos 16 anos de idade, apresentava dismorfismos faciais tais com cristas supraorbitárias proeminentes, sobrancelhas com pelos esparsos (arrancados pela paciente), pálpebras encobertas e nariz alongado e bulboso com filtro nasal curto. É portadora de DI leve associada a comportamento agressivo e hiperfagia, fala apenas palavras isoladas, mas aprendeu a ler com limitações e escreve o próprio nome. Tem altura de 158,4cm, peso de 105kg e PC de 55,5cm (Figura 23C).

Figura 23. Face dos irmãos afetados. (A) paciente 6a (II.1), (B) paciente 6b (II.2) e (C) paciente 6c (II.3).

Resultados

A análise de exoma resultou na identificação de uma variante do tipo *missense* [NM_005629.3:c.205G>C:p.(Ala69Pro)] no exon 1 do gene *SLC6A8*, em hemizigose nos irmãos e em heterozigose na irmã afetada. Esta variante modifica um aminoácido evolutivamente conservado em diferentes espécies e não está relatada em nenhum dos bancos consultados de dados populacionais. A variante foi avaliada por análises *in silico* que a classificaram como “danosa” (SIFT, FATHMM MKL e MutationTaster) ou “provavelmente danosa” (Polyphen-2) à função proteica. De acordo com o banco de dados RefSeq, este gene apresenta três diferentes transcritos e essa variante afetaria dois deles, NM_005629.3 (isoforma 1) e NM_001142805.1 (isoforma 2). Adicionalmente, a variante está mapeada em uma região considerada promotora do gene, de acordo com o banco de dados Ensembl.

Em estudo de segregação da família, a presença da variante foi confirmada nos irmãos e em sua mãe, portadora obrigatória. Nas Figuras 24 e 25 podemos verificar na imagem do arquivo BAM a presença da variante em heterozigose na mãe (I.2) e filha (II.3) e em hemizigose nos irmãos (II.1 e II.2).

A mãe e a filha, conforme revelado pela análise do padrão de metilação do loco do gene *AR* (*Androgen Receptor* – receptor de andrógenos), não apresentavam desvio de inativação do cromossomo X em linfócitos de sangue periférico (Allen et al. 1992). Na Figura 26 temos o padrão de inativação do cromossomo X em células de sangue periférico da mãe dos afetados (57:43) e na irmã portadora (43:57).

Figura 24. Imagem do arquivo BAM mostrando nas *reads* a posição da variante de nucleotídeo único na posição g.152954234G>C (c.205G>C) no exon 1 do gene *SLC6A8* visualizada no *software* VarSeq (Golden Helix). (A) I.2 (mãe) e (B) II.3 (afetada): observa-se a presença do nucleotídeo mutado (em amarelo) em heterozigose; a estrela laranja na sequência proteica indica a posição do aminoácido trocado (Ala69).

Figura 25. Imagem do arquivo BAM mostrando nas *reads* a posição da variante de nucleotídeo único na posição g.152954234G>C (c.205G>C) no exon 1 no gene *SLC6A8* visualizada no *software* VarSeq (Golden Helix). (C) II.1 (afetado) e (D) II.3 (afetado), observa-se a presença do nucleotídeo mutado (em amarelo) em hemizigose e a estrela laranja indica o aminoácido onde a troca ocorreu na proteína (Ala69).

Figura 26. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (A e B) e na filha afetada (C e D), evidenciando desvio parcial de inativação. (A e C) amplificação de dois alelos do loco *AR*, 2 e 3 na mãe e 1 e 2 na filha, mostrando heterozigose. (B e D) após o tratamento com a enzima sensível à metilação *HpaII*. Na mãe, inativação do alelo 2 do cromossomo X em aproximadamente 57% das células de sangue periférico (razão de inativação 57:43) e na filha, inativação do alelo 1 do cromossomo X em aproximadamente 43% das células de sangue periférico (razão de inativação 43:57).

Discussão

Nesta família foi identificada em três irmãos uma variante *missense* no gene *SLC6A8*, segregando com fenótipo de DI, atraso de fala e agressividade. Este gene codifica uma proteína transportadora de creatina, molécula que desempenha um papel importante no metabolismo energético celular. A creatina pode ser obtida na alimentação ou pode ser sintetizada pelo organismo. Duas enzimas são necessárias para a síntese de creatina, a arginina-glicina amidinotransferase (*AGAT*) e a metiltransferase do ácido guanidinoacético (*GAMT*). A proteína AGAT catalisa a transferência de um grupo guanidina de arginina para glicina, para formar ornitina e guanidinoacetato; a GAMT catalisa a transferência de um grupo metil de S-adenosilmetionina para guanidinoacetato para formar S-adenosil-homocisteína e creatina. A creatina presente no plasma, sintetizada ou obtida pela alimentação, é introduzida nas células por meio do

transportador de creatina (*SLC6A8*), localizado na membrana plasmática. O funcionamento ineficaz do transportador de creatina ocasiona diminuição da captação de creatina pelas células. No cérebro, a creatina é responsável por manter os altos níveis de ATP; Marc e col. (2009) estudaram cérebros de ratos e os resultados mostram altos níveis da proteína em regiões cerebrais específicas, principalmente naquelas envolvidas no processamento motor e sensorial e nas funções límbicas (aprendizagem e memória) (Mak et al. 2009). Mutações no gene *SLC6A8* causam a síndrome de deficiência cerebral de creatina 1.

Em estudo retrospectivo, que avaliou dados clínicos, bioquímicos e moleculares de 85 famílias cujas mutações patogênicas no gene *SLC6A8* segregavam em 101 homens afetados, as principais características encontradas foram DI associada a grave comprometimento da fala (100% dos afetados), anormalidades comportamentais (85%) e convulsões (59%), além de déficit de atenção e/ou hiperatividade (55%) e comportamento autista (41%) (van de Kamp et al. 2013). As mutações mais comumente detectadas foram do tipo *missense* (23 diferentes mutações, em 26/85 famílias) e deleções *inframe* (nove mutações, em 20/85 famílias) (van de Kamp et al. 2013). Este gene apresenta 10 diferentes transcritos e seis deles são codificadores de proteína.

Rosenberg e col. (2004) sugeriram que a frequência de mutações em *SLC6A8* em afetados por DI ligada ao X seria elevada, próxima daquela encontrada nos afetados pela síndrome do cromossomo X frágil; seis mutações patogênicas, das quais cinco eram novas, foram identificadas em um total de 288 pacientes, prevalência de pelo menos 2,1% (Rosenberg et al. 2004). Em 2014, Van de Kamp e col. reavaliaram os vários estudos quanto a prevalência da deficiência do transportador de creatina; segundo os autores estima-se que a prevalência em homens com DI esteja entre 0,4 e 1,4% (dependendo da coorte estudada), corroborando a frequência de cerca de 2% em homens com DI ligada ao X (van de Kamp et al. 2014). Desroches e col. (2015) realizaram um levantamento das variantes do gene *SLC6A8* reportadas no banco de dados EVS (The Exome Variant Server – <http://evs.gs.washington.edu/EVS/>) e estimaram em mulheres uma frequência aproximada de 0,024% de portadoras de variantes

possivelmente patogênicas no gene da deficiência do transportador de creatina (uma mutação patogênica em 4060 mulheres) (DesRoches et al. 2015).

O espectro fenotípico encontrado nas mulheres portadoras é muito variável e em geral mais leve do que o encontrado em homens (Mancini et al. 2005; Valayannopoulos et al. 2012; van de Kamp et al. 2011; van de Kamp, Mancini and Salomons 2014), incluindo quadros de DI leve a moderada, problemas comportamentais, convulsões, atraso de aprendizado e também mulheres portadoras assintomáticas. Heterozigotas apresentando quadro clínico leve foram diagnosticadas por meio da investigação de seus filhos afetados mais gravemente (Salomons et al. 2003; Valayannopoulos et al. 2012; van de Kamp et al. 2011; van de Kamp, Mancini and Salomons 2014). Vários trabalhos realizaram estudo de inativação do cromossomo X em portadoras de mutação em *SLC6A8* visando esclarecer a variabilidade de manifestações clínicas, porém não foi possível associar um padrão de desvio de inativação com o fenótipo das portadoras. Desvio total de inativação do cromossomo X não foi detectado nos diferentes tecidos de mulheres heterozigotas, indicando que não há seleção contra a mutação no gene, ao menos nos tecidos estudados; entretanto, tecido cerebral dessas mulheres não foi avaliado.

Os afetados desta família apresentam quadro clínico compatível com a síndrome de deficiência de creatina cerebral 1 que segrega com uma nova variante *missense* no gene *SLC6A8*, provavelmente patogênica. A mãe portadora assintomática e a filha afetada não apresentam desvio de inativação do cromossomo X, o que é compatível com a variabilidade fenotípica já descrita na literatura em mulheres para mutações *missense*. Vários trabalhos na literatura estudaram a inativação do cromossomo X nas mulheres portadoras de mutações no gene *SLC6A8* e não foi estabelecida a associação fenótipo e desvio de inativação, não estando clara ainda a base genética para a variabilidade fenotípica em mulheres.

A variante c.205G>C identificada na família 6 foi considerada uma variante provavelmente patogênica pelas seguintes características: não está presente nos bancos populacionais; foi consistentemente avaliada *in silico* como prejudicial à função proteica; está localizada no exon 1, sobreposta a região considerada

promotora do gene, afetando duas de suas principais isoformas; mutações *missense* são um mecanismo comum associado à patogênese de *SLC6A8*; segregação com o fenótipo e grande sobreposição do mesmo com o espectro clínico dos distúrbios associados a mutações em *SLC6A8*. Assim, nossas descobertas sugerem fortemente que a variante detectada é a causa do quadro clínico desses pacientes.

VI.3.5. FAMÍLIA 8 – MOSAICISMO EM DI COM PADRÃO DE HERANÇA DOMINANTE

Resumo

Neste estudo foi investigada a causa genética de quadro clínico de duas irmãs com DI moderada e características faciais dismórficas. A análise do sequenciamento de exoma revelou uma variante *frameshift* no gene *SETBP1* (OMIM *611060), que codifica a proteína de ligação a SET1, ubiquamente detectada no corpo. Este gene está localizado em 18q12.3 e suas mutações e deleções foram associadas à deficiência intelectual autossômica dominante 29 (OMIM #616078 – Mental Retardation, Autosomal Dominant 29; MRD29) e à síndrome de Schinzel-Giedion (OMIM #269150 – Schinzel-Giedion Midface Retraction Syndrome; SGS).

A variante *frameshift* c.628delC (p.His210Thrfs) detectada nas irmãs aqui descritas não está relatada em nenhum dos bancos consultados de variantes genéticas populacionais e tem o efeito de perda de função da proteína. A mãe, que apresenta atraso cognitivo leve, foi identificada como portadora da variante em mosaico. Os dados em conjunto, variante rara com efeito de perda função em gene já associado a DI autossômica dominante e segregando com o fenótipo de DI leve, indicam fortemente que essa variante é patogênica. Esse é o primeiro caso familiar de deficiência intelectual autossômica dominante 29, associado a herança materna com presença de mosaicismo.

Espectro de manifestações clínicas de mutações no *SETBP1*

O gene *SETBP1*, localizado em 18q12.3, codifica a proteína de ligação a SET1, que recebe esse nome por ser capaz de se ligar a outra proteína chamada

SET. Recentemente foi demonstrado que a enzima SETBP1 atua como ativadora e moduladora de expressão gênica em uma rede de genes de desenvolvimento (revisão em Piazza et al. 2018). Nesse estudo recente de Piazza et al. (2018), foi demonstrada a capacidade da SETBP1 humana de se ligar diretamente ao DNA genômico preferencialmente, mas não exclusivamente, em regiões promotoras ricas em AT. Essa proteína atuaria recrutando para a região de sua ligação várias proteínas de um complexo epigenético (HCF1/KMT2A/PHF8), resultando em maior acessibilidade à cromatina nesse complexo multiproteico ativador da transcrição.

A síndrome de Schinzel-Giedion (SGS) foi descrita pela primeira vez em dois irmãos que apresentavam dismorfismos faciais, defeitos renais e cardíacos e DI grave, nascidos de pais saudáveis e não consanguíneos. A ocorrência do quadro clínico em dois irmãos de sexo diferente foi sugerida herança autossômica recessiva para a síndrome (Schinzel and Giedion 1978). Em 2010, Hoischen e col. foram os primeiros a associar mutações germinativas em heterozigose no gene *SETBP1* à SGS, em quadro clínico caracterizado por atraso de crescimento pós-natal, convulsões, DI grave, dismorfismos faciais e malformações congênitas múltiplas. Os autores sugeriram que as mutações no gene *SETBP1*, identificadas em heterozigose, seriam de ganho de função com efeito dominante negativo; todas as mutações detectadas estavam localizadas no exon 4, fora do domínio de ligação a SET (Hoischen et al. 2010). Até o momento, oito diferentes mutações em *SETBP1* foram relatadas em pacientes com SGS (Carvalho et al. 2015; Herenger et al. 2015; Hoischen et al. 2010; Ko et al. 2013; Lopez-Gonzalez et al. 2015; Suphapeetiporn et al. 2011; Takeuchi et al. 2015), todas no *hotspot* de mutação (exon 4) e *de novo*.

Adicionalmente, o gene *SETBP1* é considerado um oncogene, uma vez que mutações somáticas foram associadas a diversas neoplasias hematológicas, como mielodisplasia (MDS) e síndromes mielodisplásicas/mieloproliferativas (MDS/MPN) (Makishima 2017; Makishima et al. 2013; Yao et al. 2018), sendo consideradas mutações de ganho de função. Essas mutações somáticas em neoplasias hematológicas ocorrem no mesmo *hotspot* no qual estão localizadas as mutações germinativas na SGS; de fato, a propensão para desenvolver

tumores neuroepiteliais está incluída no espectro de manifestações clínicas da síndrome.

Diferentemente do efeito clínico identificado para as mutações somáticas e germinativas de ganho de função, deleções e mutações do gene *SETBP1* que ocasionam perda de função da proteína foram descritas em pacientes com DI leve a moderada em associação a dimorfismos faciais e atraso ou ausência de fala (Coe et al. 2014; Filges et al. 2011; Hamdan et al. 2014; Marseglia et al. 2012).

Neste trabalho, nós relatamos uma família com duas irmãs que apresentam DI moderada e dismorfismos faciais, causada por variante *frameshift* em *SETBP1*, no *hotspot* de mutação exon 4, herdada de mãe que apresenta mosaïcismo germinativo para essa variante e quadro de atraso cognitivo leve, num padrão de herança autossômico dominante.

Descrição da Família 8 (pacientes 8a e 8b)

A Família 8 foi encaminhada pela Dra. Fernanda S. Jehee, pesquisadora docente do Instituto de Ensino e Pesquisa da Santa Casa de Belo Horizonte. Os pacientes foram triados na APAE de Santa Luzia, MG. As duas irmãs nasceram de genitores descritos não consanguíneos, a mãe relatou atraso de aprendizado (durante a consulta evidência de atraso cognitivo leve). O heredograma desta família é mostrado na Figura 27.

Figura 27. Heredograma da Família 8, na qual foi detectada uma variante patogênica no gene *SETBP1* do tipo *frameshift* que segrega com DI num padrão de herança autossômica dominante.

Paciente 8a (II.1)

Nascimento a termo de parto normal. Sentou aos 10 meses e andou com 1 ano e 5 meses de idade. Examinada aos 17 anos de idade (Figura 28), apresentava DI moderada, dolicocefalia e dismorfismos faciais como rosto longo, sinofre, sobrancelhas arqueadas, fenda palpebral com inclinação ascendente e ponte nasal baixa. Apresenta ausência de fala e não aprendeu a ler. Tem altura de 165cm, peso de 58kg e PC de 58,5cm.

Figura 28. Face da paciente 8a (II.1).

Paciente 8b (II.3)

Nascimento a termo de parto normal. Examinada aos 11 anos de idade (Figura 29), apresenta DI moderada e dismorfismos faciais como rosto longo, sobrancelhas arqueadas, ponte nasal baixa e epicanto. Mãe relata craniossinostose, corrigida cirurgicamente aos 8 meses. Fala pouco desenvolvida (poucas palavras) e não aprendeu a ler. Tem altura de 147cm, peso de 33,1kg e PC de 52cm.

Figura 29. Face da paciente 8b (II.3).

Resultados

A análise do exoma das pacientes resultou na identificação em ambas as irmãs de uma variante em heterozigose no exon 4 do gene *SETBP1* [NM_015559.2:c.628delC:p.(His210Thrfs)]. Esta variante é do tipo *frameshift*, levando à perda de função da proteína, e não está relatada em nenhum dos bancos de dados populacionais consultados. A presença da variante no gene *SETBP1* foi confirmada por sequenciamento Sanger nas irmãs e não foi detectada nos genitores (Figura 30).

Figura 30. Sequenciamento Sanger em amostras de DNA genômico mostrando a detecção da deleção em heterozigose da variante g.42529933delC (c.628delC), encontrada no exon 4 do gene *SETBP1* nas irmãs afetadas (II.1 e II.3) e apenas a presença da sequência normal nos pais (I.1 e I.2) e controle.

Considerando que ambas as pacientes são portadoras da variante, ela não poderia ser *de novo*. Para investigar a presença de mosaicismos germinativos nos genitores, foi realizado o ensaio de Droplet Digital PCR (ddPCR) com sonda específica para a variante c.628delC. O resultado revelou que a mãe é portadora da mesma em 0,13% das células de sangue periférico, mostrando a existência de mosaicismos (Figura 31).

Figura 31. Detecção de mosaicismo para a variante c.628delC de *SETBP1* em células de sangue periférico da mãe das afetadas. Os gráficos representam os dados referentes às reações-controle e às reações para detecção das sequências-alvo (alelo selvagem e alelo com a variante). Os pontos pretos representam as reações com DNA sem amplificação (gotículas vazias), os pontos verdes, as reações de amplificação do alelo selvagem, os azuis, amplificação do alelo mutado, e os pontos laranjas são os duplo-positivos, que na mesma gotícula tinham os dois alelos (selvagem e mutado). (A) Paciente II.1, detecção de 49,2% do alelo mutado (pontos em azul) em amostra de DNA extraída de sangue periférico. (B) Mãe (I.2), detecção de 0,13% do alelo mutado (em azul) em amostra de DNA extraída de sangue periférico. (C) Pai (I.1), detecção apenas de alelos selvagens quanto à variante testada, em amostra de DNA extraída de sangue periférico.

Discussão

Nesta família, uma nova variante do tipo *frameshift* foi detectada no gene *SETBP1*, em heterozigose, em duas afetadas por DI moderada, ausência de fala e dismorfismos faciais; a variante foi herdada da mãe, que apresenta mosaicismos comprovado e atraso cognitivo. A proteína de ligação SET1 apresenta papel importante na modulação da transcrição gênica. Tipos diferentes de mutações germinativas foram associadas à síndrome de Schinzel-Giedion (SGS) e à deficiência intelectual autossômica dominante 29, enquanto mutações somáticas ocasionam neoplasias hematológicas.

O fenótipo associado à haploinsuficiência de *SETBP1*, causada por microdeleções ou mutações de perda de função, é muito mais leve quando comparado à SGS, resultando em um fenótipo distinto e reconhecível que inclui DI variando de leve a grave, ausência ou grande atraso de fala e dismorfismos faciais (Coe et al. 2014). Esses dados apoiam o conceito de que diferentes tipos de mutações no mesmo gene podem provocar variabilidade fenotípica, já que as mutações em *SETBP1* associadas a SGS são de ganho de função com efeito dominante negativo.

As duas irmãs descritas apresentam quadro clínico compatível com mutações de perda de função de *SETBP1*, devido à nova variante *frameshift* identificada. Na Figura 32 são mostradas imagens faciais dos pacientes descritos por Coe e col., ao lado de fotos das pacientes da Família 8. Em todos os pacientes relatados nos quais a herança da mutação foi determinada, a variante era *de novo* (Coe et al. 2014; Filges et al. 2011; Marseglia et al. 2012; Rauch et al. 2012) e nenhuma irmandade com mutações de perda de função em *SETBP1* havia ainda sido relatada. Chamamos a atenção para o fato de que a variante aqui detectada também teria sido considerada *de novo*, não fosse a repetição do quadro clínico em duas irmãs, que levou à investigação de mosaicismos nos genitores por sequenciamento Sanger e, quando negativo, ao uso de PCR digital. Retrospectivamente, o relato de atraso de aprendizado da mãe sugeria que ela poderia ser portadora da variante patogênica.

Figura 32. (A) Características faciais de pacientes relatados na literatura com variantes de perda de função no gene *SETBP1* (Coe e col. 2014). (B) Características faciais das pacientes da família 8, portadoras da variante *frameshift* aqui identificada.

Esse é o primeiro relato de um caso familiar de variante de perda de função no gene *SETBP1*. O estudo revelou que a mãe era portadora da variante em mosaico de 0,13% em células de sangue periférico. Esse caso ilustra a contribuição, ainda pouco considerada, do mosaicismo germinativo na DI com padrão de herança dominante, geralmente assumidos como *de novo* por serem identificados como casos isolados.

VI.3.6. FAMÍLIA 9 – *ARHGAP4*: UM NOVO GENE DE DEFICIÊNCIA INTELECTUAL NÃO SINDRÔMICA?

Resumo

A família aqui descrita é composta por dois irmãos do sexo masculino afetados por atraso de DNPM sem associação a outros sinais clínicos, nos quais foi identificada uma variante em heterozigose do tipo mudança de nucleotídeo (*missense*) no gene *ARHGAP4* (OMIM *300023), mapeado no cromossomo X, que codifica uma proteína RHO GTPase, potencialmente relacionada a motilidade celular e crescimento de axônios. A família das proteínas Rho GTPase tem padrões distintos de expressão em diferentes células neuronais, indicando funções específicas de áreas do cérebro durante o desenvolvimento e diferenciação. Este gene está localizado no braço longo do cromossomo X (Xq28) e não foi até o momento associado a nenhum quadro clínico, segundo o banco de dados OMIM. A variante *missense* c.1471C>T (p.Arg491Trp) detectada nessa

família ocorre no códon de um aminoácido evolutivamente conservado e não está documentada em nenhum dos bancos de variantes genéticas populacionais consultados. Análises *in silico* classificaram a variante c.1471C>T como provavelmente danosa à função proteica. A análise de segregação detectou a presença da variante na mãe, heterozigota, num padrão que poderia ser de herança recessiva ligada ao cromossomo X. Considerando a raridade da variante e possível efeito na proteína, os dados sugerem essa variante em *ARHGAP4* como uma possível causa de DI não síndrômica nessa irmandade.

O gene *ARHGAP4*

Dados sobre esse gene ainda são escassos na literatura. Identificado em 1996 por Tribioli e col., *ARHGAP4* está localizado em Xq28 e codifica uma proteína que faz parte de uma família de moléculas de sinalização envolvida na regulação de proteínas de *GTP-binding* da superfamília RAS (Tribioli et al. 1996).

A proteína ARHGAP4 é predominantemente expressa em células hematopoiéticas (células T, células B e mieloides) e possivelmente desempenha um importante papel no sistema nervoso, relacionado ao desenvolvimento do cérebro, regulando a dinâmica do citoesqueleto com impacto em motilidade celular e crescimento de axônios (Tribioli et al. 1996; Vogt et al. 2007).

O gene *ARHGAP4* foi previamente sugerido como causa de DI, embora em apenas dois relatos de caso. Portanto, sua efetiva contribuição para alterações cognitivas ainda não foi esclarecida.

Neste trabalho, nós relatamos uma família com dois irmãos apresentando atraso de DNPM, portadores de uma variante *missense* de significado indeterminado (VUS) no exon 11 do gene *ARHGAP4*, identificada por meio do sequenciamento do exoma.

Descrição da Família 9 (pacientes 9a e 9b)

A Família 9 foi encaminhada pela Dra. Fernanda S. Jehee, então pesquisadora docente do Instituto de Ensino e Pesquisa da Santa Casa de Belo Horizonte. Os pacientes foram triados na APAE de Santa Luzia, MG. Os pacientes

são nascidos de pais saudáveis e não consanguíneos. O heredograma desta família é mostrado na Figura 33.

Figura 33. Heredograma da Família 9, na qual segrega uma variante no gene *ARHGAP4* num padrão compatível com herança recessiva ligada ao cromossomo X.

Paciente 9a (II.1)

Nascimento a termo de parto normal, com 3270g. Examinado aos 2 anos e 8 meses de idade (Figura 34A), apresentava atraso de DNPM, orelhas de implantação baixa, ponte nasal deprimida com ponta do nariz elevada e narinas antevertidas, filtro longo e comportamento classificado como transtorno do espectro autista. À época, tinha altura de 102cm, peso de 14,4kg e PC de 48,5cm, parâmetros dentro da normalidade para a idade.

Paciente 9b (II.3)

Nascimento a termo de parto normal, com 3300g. Examinado aos 1 anos e 7 meses de idade (Figura 34B), apresentava atraso de DNPM, fronte proeminente, cílios longos, ponte nasal deprimida com ponta do nariz elevada e narinas antevertidas, filtro longo e lábio superior fino. Tem altura de 84cm, peso de 10,4kg e PC de 48,3cm.

Figura 34. Face dos irmãos afetados. (A) paciente 9a (II.2) e (B) paciente 9b (II.3).

Resultados

A análise dos dados do exoma resultou na identificação de uma variante *missense* em hemizigose no exon 11 do gene *ARHGAP4* [NM_001164741.1:c.1471C>T:p.(Arg491Trp)]. Esta variante *missense* não foi previamente relatada em nenhum dos bancos de dados populacionais consultados.

A variante afeta um aminoácido que é evolutivamente conservado em diferentes espécies de vertebrados e sua patogenicidade foi avaliada *in silico* usando diferentes algoritmos, sendo classificada como “danosa” por SIFT, “possivelmente danosa” por Polyphen-2 e “danosa” por MutationTaster.

A presença da variante foi confirmada por sequenciamento Sanger nos irmãos afetados, assim como em heterozigose na mãe (Figura 35).

Figura 35. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante *missense* c.1471C>T identificada no exon 11 do gene *ARHGAP4*. Presença da variante em hemizigose nos irmãos afetados (II.2 e II.3) e em heterozigose na mãe (I.2); ausência da variante alterada no controle masculino.

Conforme revelado pela análise do padrão de metilação do loco do gene *AR*, a mãe portadora da variante apresenta desvio parcial de inativação do cromossomo X em linfócitos de sangue periférico. A Figura 36 mostra o desvio parcial de inativação do cromossomo X, estimado na proporção 78:22.

Figura 36. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados, heterozigota quanto à variante no gene *ARHGAP4*, evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de um dos alelos (1), mostrando inativação preferencial de um dos cromossomos X em aproximadamente 78% das células de sangue periférico (razão de inativação 78:22).

Discussão

Nesta família, foi identificada em dois afetados por atraso de DNPM uma variante *missense* no gene *ARHGAP4*, que codifica uma proteína citoplasmática que atua na função de microtúbulos em motilidade celular, potencialmente envolvida no desenvolvimento cerebral, uma vez que poderia alterar o crescimento de axônios. Outros genes da família Rho GTPase, como *OPHN1* (OMIM*300127), localizado em Xq12, estão claramente associados a DI por afetar a migração celular e o crescimento de axônios e dendritos *in vivo* (Billuart et al. 1998). Entretanto, o papel do gene *ARHGAP4* não está ainda totalmente elucidado, assim como uma potencial associação com DI.

Em 2012, Huang e col. descreveram um par de gêmeos dizigóticos do sexo masculino afetados por DI e *diabetes insipidus* nefrogênica, ambos portadores de uma deleção genômica de 17,9kb englobando parte das sequências dos genes *AVPR2* e *ARHGAP4*. Outros diversos casos de pacientes portadores de diferentes microdeleções afetando *AVPR2* e *ARHGAP4* haviam sido previamente descritos, a maioria deles exibindo apenas a *diabetes insipidus* nefrogênica, reconhecidamente associada ao gene *AVPR2* (Broides et al. 2006; Demura et al. 2002; Dong et al. 2006; Fujimoto et al. 2008). Entretanto, alguns dos portadores dessas microdeleções em Xq28 também exibiam DI (Schoneberg et al., 1999; Demura et al., 2002), corroborando a hipótese de Huang de que deleções parciais do gene *ARHGAP4*, a depender de sua localização na sequência gênica, estariam relacionadas ao quadro de atraso cognitivo.

Em apoio à hipótese de que o gene *ARHGAP4* é um candidato a causar deficiência intelectual, um trabalho recente relatou a detecção de uma variante *missense* neste gene, em homens com DI; nessa família chinesa, descrita por Liu e col. (2016), a variante em g.153173202G>A (p.Thr941Met) está localizada no exon 22 e segrega com o fenótipo de DI em homens, com mulheres portadoras normais (Liu et al. 2016).

O gene *ARHGAP4* está mapeado na fita reversa e apresenta 24 transcritos, dos quais 12 são codificadores de proteína; dentre estes, seis contêm o exon 11, no qual está localizada a variante *missense* identificada em nosso estudo.

Neste trabalho, nós relatamos uma família com dois irmãos com atraso de DNPM portadores de uma variante *missense* de significado indeterminado (VUS) no gene *ARHGAP4*, herdada de mãe heterozigota com desvio parcial de inativação do cromossomo X. Com a descrição desse novo caso, contribuímos para reforçar a hipótese do gene *ARHGAP4* ser candidato para DI não síndrômica.

VI.3.7. FAMÍLIA 11 – REGIÃO 3’ NÃO-TRADUZIDA (3’ UTR) DE *PHF6*: UMA VARIANTE A SER CONSIDERADA PARA O FENÓTIPO?

Resumo

Descrevemos neste capítulo uma irmandade, um indivíduo do sexo masculino e outro do sexo feminino, ambos apresentando DI síndrômica de grau variável entre eles, nos quais foi identificada uma variante de nucleotídeo único (*single nucleotide variant* - *SNV*) na região 3’ não traduzida do gene *PHF6* (OMIM *300414). Este gene codifica uma proteína com domínio conhecido como dedo de zingo (*zinc finger*), o que indica importante papel na regulação da transcrição, que se localiza em nucléolo. O gene *PHF6* está localizado no braço longo do cromossomo X (Xq26.2) e suas mutações e deleções germinativas foram associadas à síndrome Borjeson-Forssman-Lehmann (OMIM #301900 – Borjeson-Forssman-Lehmann Syndrome; BFLS), assim como mutações somáticas foram descritas em algumas neoplasias hematológicas. A variante detectada nessa família ocorre na região 3’ UTR do gene, que contém sequências-alvo de microRNAs. No banco de dados de genomas GnomAD foram descritos apenas 10 alelos em 20669, sendo um único hemizigoto (rs761139135), todos em população africana.

O gene *PHF6*

Em 2002, Lower e col. associaram o gene *PHF6* à síndrome de Borjeson-Forssman-Lehmann (BFLS; OMIM # 301900). Os autores realizaram um estudo com nove famílias com indivíduos diagnosticados com BFLS; entre os pacientes, foram detectadas oito diferentes mutações no gene *PHF6*, seis *missense* e duas *nonsense*. Apenas uma mutação *missense* foi encontrada duas vezes, em duas

famílias não relacionadas afetadas com BFLS. Esta síndrome é caracterizada por DI moderada a grave, baixa estatura, genitália pequena, epilepsia, hipogonadismo, hipometabolismo, obesidade com ginecomastia, inchaço do tecido subcutâneo da face, fissura pálpebral estreita e orelhas grandes. O atraso no desenvolvimento geralmente aparece antes do primeiro ano de vida. Outras características da síndrome incluem dedos largos, afunilados e encurtados. Os pacientes adultos com BFLS exibem características faciais mais grosseiras, com cristas supraorbitárias proeminentes e olhos fundos (Gecz et al. 2006; Lower et al. 2002).

A proteína PHF6 é altamente conservada em vertebrados e parece desempenhar um papel importante na regulação da transcrição (Liu et al. 2014), em especial de rRNA nos nucleólos. Em 2015, num estudo realizado com camundongos, Franzoni e col. mostraram que a regulação negativa da expressão do gene *PHF6* está correlacionada com aumento nos níveis de expressão de miR-128, sugerindo que miR-128 poderia reprimir a expressão de *PHF6* durante o desenvolvimento do cérebro, atuando para mediar migração neuronal no córtex cerebral (Franzoni et al. 2015).

Neste trabalho, nós relatamos uma família com dois irmãos, um do sexo masculino e uma do sexo feminino, apresentando DI, dificuldade de fala e convulsões, portadores de uma variante *SNV* de significado indeterminado na região 3'UTR do gene *PHF6*, identificada por sequenciamento completo de exoma.

Descrição da Família 11 (pacientes 11a e 11b)

A família aqui apresentada foi averiguada pelos Drs. Angela M. Vianna-Morgante e Paulo A. Otto no Serviço de Aconselhamento Genético do Laboratório de Genética Humana do Departamento de Genética e Biologia Evolutiva, Instituto de Biociências da Universidade de São Paulo (IB-USP), e posteriormente atendidos na Pediatria do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (FMUSP) pela Dra. Rachel Sayuri Honjo Kawahira, médica geneticista. Nesta família foram investigados dois irmãos afetados por DI

sindrômica; são nascidos de pais saudáveis e com consanguinidade remota (primos de 3º grau). O heredograma desta família é mostrado na Figura 37.

Figura 37. Heredograma da Família 11; na qual segrega uma variante de significado indeterminado no gene *PHF6* em mãe clinicamente normal e em um casal de filhos, com padrão compatível com herança recessiva ligada ao X.

Paciente 11a (II.1)

Nascimento de 32 semanas de parto cesáreo devido a hipertensão materna, com 1770g e 40 cm, Apgar de 4/7. Apresentava genitália atípica com escroto bífido, fálus emergindo entre as duas hemibolsas escrotais, meato uretral saindo da glândula do fálus e testículos ausentes. Feito exame de cromatina sexual, que foi negativo; exames subsequentes mostraram cariótipo 46,XY e testosterona aumentada. Examinado aos 2 anos e 7 meses, apresentava atraso de DNPM: sustentou a cabeça aos 8 meses e sentou com oito meses, não andava ou falava. Ao exame, peso de 13,6kg, altura de 92,5cm e PC de 48,5cm, normais para o sexo e idade. Apresentava plágio-braquicefalia com occipício achatado, fendas palpebrais oblíquas para cima, estrabismo divergente, ponta do nariz bulbosa, orelhas de implantação baixa e micrognatia; mamilos hipoplásicos; testículos na região inguinal. Sem alterações em RX crânio, ecocardiograma e exames metabólicos incluindo 7DHC. No exame de ressonância magnética foram

detectados sinais de redução volumétrica supra e infratentorial, com múltiplas áreas de sequelas focais, comprometendo a coroa radiada bilateralmente; redução volumétrica do tronco cerebral; pequeno lipoma de placa quadrigeminal em situação paramediana esquerda. Em eletroencefalograma, foi constatada desorganização difusa da atividade de base, predominando na região temporal posterior direita. Atividade epileptiforme de projeção na região temporal esquerda e predominantemente direita e de forma menos frequente na região frontal e central esquerda. Atividade epileptiforme de projeção generalizada.

Avaliado aos 8 anos, o exame constatou, além dos dismorfismos craniofaciais, DI grave e convulsões; começou a andar aos 4 anos. Não falava e apresentava estereotípias do espectro autista.

Paciente 11b (II.2)

Nascimento a termo de parto normal. Examinada aos 5 anos e 9 meses de idade, apresentava DI leve/moderada, convulsões e não falava. Apresentava fendas palpebrais curtas, oblíquas para cima e ponta do nariz levemente bulbosa. Genitália com discreta hipoplasia de grandes lábios. Avaliada novamente aos 9 anos, falava palavras com ecolalia e apresentava características autistas, estatura normal.

Resultados

A análise do sequenciamento de exoma completo dos dois irmãos resultou na identificação de uma variante de nucleotídeo único na porção 3'UTR do gene *PHF6* [NC_000023.10:g.133562204C>T; NM_032458.2:c.*2844C>T, NM_001015877.1: c.*2514C>T]. Esta variante em gene mapeado no cromossomo X está em hemizigose no irmão e em heterozigose na irmã. No banco de dados gnomAD (que contém 20669 genomas) há informação de 10 alelos, apenas um caso em hemizigose, com frequência global de 0,0005. Neste banco populacional, não há informação clínica sobre portadores e essa variante foi reportada exclusivamente em populações africanas.

A presença da variante foi investigada por sequenciamento Sanger na mãe, que é portadora heterozigota (Figura 38).

Figura 38. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante de nucleotídeo único na posição g.133562204C>T encontrada na 3'UTR do gene *PHF6*. Presença da variante em hemizigose no irmão afetado (II.1) e em heterozigose na mãe (I.2). O pai (I.1) é portador do alelo normal.

Conforme revelado pela análise do padrão de metilação do loco do gene *AR*, a mãe portadora da variante (I.II) apresenta desvio parcial de inativação do cromossomo X em linfócitos de sangue periférico (68:32). A Figura 39 mostra o padrão de inativação do cromossomo X da mãe dos afetados. A filha (II.II) apresenta desvio parcial de inativação do cromossomo X (79:21).

Figura 39. Padrão de inativação do cromossomo X em sangue periférico da mãe dos afetados (I.2), evidenciando desvio parcial de inativação. (A) amplificação de dois alelos do loco *AR*, 1 e 2, mostrando heterozigose; (B) após o tratamento com a enzima sensível à metilação *HpaII*, amplificação majoritária de um dos alelos (1), mostrando inativação preferencial de um cromossomo X em 68% das células de sangue periférico (razão de inativação 68:32).

Discussão

Identificamos em dois afetados por DI uma variante *SNV* na região 3' UTR do gene *PHF6*. Ambos os irmãos manifestaram atraso global de desenvolvimento com evolução para DI com convulsões, embora em graus diferentes, além de dismorfismos faciais em comum e hipogenitália/genitália atípica; o quadro clínico da irmã do probando pode ser considerado bastante mais leve, em especial não havendo alteração marcante de genitália.

Mutações germinativas em *PHF6* causam a síndrome de Börjeson – Forssman – Lehmann (BFLS, MIM # 301900) e também estão associadas a um fenótipo específico do sexo feminino que se sobrepõe à síndrome de Coffin-Siris (MIM #135900), assim como mutações somáticas foram descritas em leucemia linfoblástica aguda de células T (MIM #613065) e leucemia mieloide aguda (MIM #601626). As mutações germinativas descritas em BFLS estão distribuídas pela sequência de *PHF6*, sugerindo que o quadro clínico associado resulte de perda de função da proteína; no entanto, sabendo-se que a função da proteína PHF6 é importante tanto para o desenvolvimento neuronal quanto para supressão tumoral, questiona-se porque apenas poucas mutações germinativas de *PHF6* (~6%) estariam associadas ao câncer (Carter et al. 2009; Chao et al. 2010).

Segundo Todd e col. uma possível explicação seria o fato de que a maioria das mutações germinativas são do tipo *missense*, mais provavelmente reduzindo a expressão do gene (alelos hipomórficos) do que anulando, ao contrário das mutações em câncer, que são *frameshift* e *nonsense* (alelos nulos); adicionalmente, em câncer haveria a necessidade de um segundo evento mutacional no outro alelo do gene, uma vez que este parece ser um supressor tumoral (Todd et al. 2015).

Essa síndrome apresenta extensa heterogeneidade clínica, mas até o presente a identificação de correlações genótipo-fenótipo entre mutações específicas de *PHF6* e a gravidade da deficiência intelectual foram infrutíferas (Todd et al. 2015). Crawford e col. (2006) relataram o primeiro caso de BFLS em mulheres, encontrando uma mutação *de novo* do tipo *frameshift*, no exon 2 do gene *PHF6* (Crawford et al. 2006). Entretanto, vários trabalhos tentaram associar o padrão de desvio de inativação do cromossomo X de mulheres portadoras de mutação em *PHF6* com seu quadro clínico, mas não foi identificada uma associação evidente com o fenótipo; entre as portadoras, foram encontradas mulheres apresentando desvio de inativação de um cromossomo X em 70% das células de sangue periférico, mulheres apresentando desvio superior a 90% e mulheres onde a inativação do X é aleatória. O quadro clínico em mulheres é bastante variável, embora em geral mais leve do que em homens, compatível com o padrão de herança ligado ao X recessivo (Crawford et al. 2006; Gecz et al. 2006; Lower et al. 2002; Turner et al. 2004).

Estudos recentes revelaram possíveis mecanismos pelos quais mutações de *PHF6* contribuiriam para a patogênese da deficiência intelectual da síndrome Börjeson-Forssman-Lehmann (BFLS), embora ainda não tenha sido elucidada sua atuação para produzir os demais sinais clínicos. Sabe-se que *PHF6* desempenha um papel crítico na migração de neurônios no córtex cerebral de camundongo *in vivo*, e estudos com mutações específicas de pacientes de fato afetam a capacidade da proteína executar esse papel. Adicionalmente, os dados indicam crescentemente que *PHF6* seria um gene da maquinaria epigenética, atuando no controle do status da cromatina por meio da interação com vários complexos para modular a transcrição.

A regulação pós-transcricional do mRNA de *PHF6* parece ser realizada por até 25 microRNAs, incluindo miR-20a, miR-26a, miR-128 e miR-574 (Todd et al. 2015). Estudos realizados em camundongos por Franzoni e col. (2015) mostram que o gene *Phf6* é um alvo regulatório importante para o miR-128, microRNA envolvido no controle da neurogênese e da sinaptogênese (Franzoni et al. 2015). Os autores mostram que a regulação do gene *PHF6* pelo miR-128 é importante para a migração adequada e crescimento dendrítico dos neurônios da camada superior no córtex do camundongo em desenvolvimento. Portanto, concluem que estes resultados devem ter relevância para a compreensão do déficit cognitivo e da suscetibilidade a convulsões em pacientes com mutações no gene *PHF6*.

Neste trabalho, nós relatamos uma família com dois irmãos, um do sexo masculino e uma do sexo feminino afetados por DI, dificuldade de fala e convulsões, sendo o menino notoriamente mais seriamente afetado do que a menina. Os irmãos são portadores de uma variante de significado indeterminado (VUS) na região 3' não traduzida de *PHF6*. Considerando que um dos mecanismos propostos para a patogênese dos sinais clínicos da síndrome Börjeson-Forssman-Lehmann seria por meio de mutações hipomórficas, assim como baseados nos relatos de Franzoni e col. sobre o controle pós-transcricional deste gene pelo miR-128, não podemos descartar a possibilidade de a variante detectada nesta família estar interferindo na regulação do gene *PHF6* por alterar a atuação de microRNAs que alvejam a região. Entretanto, apenas estudos funcionais poderiam esclarecer esta hipótese.

VI.3.8. FAMÍLIA 13 – IDENTIFICAÇÃO DE UMA NOVA VARIANTE NO GENE *MED13L*

Resumo

Estudamos uma família com dois irmãos, um do sexo masculino e outra do feminino, ambos apresentando DI leve a moderada, ausência de fala e características faciais dismórficas. A análise do exoma revelou uma variante sem sentido (*nonsense*) no gene *MED13L* (OMIM* 608771), localizado em 12q24.21. Este gene codifica uma proteína altamente conservada em eucariotos e é um importante regulador da transcrição (Utami et al. 2014), envolvida no

desenvolvimento do coração e cérebro. Haploinsuficiência causada por mutações ou deleções em heterozigose foram associadas a DI moderada, ausência de fala e características faciais distintas, com ou sem defeitos cardíacos (OMIM #616789 – Mental Retardation And Distinctive Facial Features With Or Without Cardiac Defects; MRFACD).

A variante *nonsense* c.4417C>T (p.Q1473X) detectada nos irmãos aqui descritos não está relatada em nenhum dos grandes bancos de variantes genéticas populacionais, introduzindo um códon de parada que ocasiona um efeito de perda de função da proteína. Os dados em conjunto, variante não reportada em bancos populacionais, provavelmente patogênica por ter efeito de perda de função em gene já associado a DI e quadro clínico compatível com a síndrome associada, indicam fortemente a patogenicidade da variante em questão.

O gene *MED13L*

A proteína codificada pelo gene *MED13L*, localizado em 12q24.21, é uma subunidade do complexo mediador que atua como co-ativador para os genes transcritos pela RNA polimerase II. Múltiplas subunidades do complexo mostram especificidade na transmissão de informações de sinais e fatores de transcrição para a maquinaria da RNA polimerase II, permitindo assim o controle da expressão de genes específicos (Yin and Wang 2014). A expressão de *MED13L* é aumentada no cérebro fetal e adulto, sendo especialmente elevada no cerebelo, coração e músculo esquelético (Muncke et al. 2003; Utami et al. 2014).

Em 2003, Muncke e col. sugeriram que o gene *MED13L* seria responsável por causar defeitos cardíacos congênitos, com base no mapeamento entre o primeiro e segundo exons deste gene do ponto de quebra de uma translocação cromossômica detectada em um paciente com DI e defeitos cardíacos (Muncke et al. 2003). Posteriormente, Asadollahi e col. (2013) identificaram em *MED13L* duas deleções intragênicas *de novo* em dois pacientes não relacionados que apresentavam hipotonia, DI moderada, defeito cardíaco e anomalias faciais. Em outro paciente com hipotonia, dificuldades de aprendizagem e malformação cardíaca, foi detectada uma triplicação *de novo* de 1Mb em 12q24.2, que incluía

o gene *MED13L*; os autores concluíram que a haploinsuficiência de *MED13L* era a responsável pelo quadro clínico dos pacientes e que o aumento do número de cópias levava a um quadro clínico mais leve (Asadollahi et al. 2013). Em 2014, Utami e col. realizaram estudos funcionais em camundongos, confirmando nesse modelo que a haploinsuficiência de *MED13L* levava ao quadro clínico reportado nos pacientes (Utami et al. 2014). Neste mesmo ano, outras mutações que levavam à perda de função de *MED13L* foram relatadas em pacientes com DI não associada a defeitos cardíacos (Hamdan et al. 2014; Redin et al. 2014). Estes e outros estudos que se seguiram confirmaram a associação de DI com mutações inativadoras em *MED13L*, também resultando na ampliação do espectro fenotípico dos pacientes (Cafiero et al. 2015; Gordon et al. 2018; van Haelst et al. 2015; Yamamoto and Shimojima 2015).

A Tabela 8, retirada do trabalho de Asadolladi e col. (2017), apresenta uma revisão acerca da frequência das principais características clínicas em pacientes descritos com alterações que levam à perda de função do gene *MED13L*; podemos observar que DI de moderada a grave e dismorfismos faciais são sinais presentes em todos os pacientes, seguido por atraso grave de fala e hipotonia muscular (Asadollahi et al. 2017).

Tabela 8. Frequência das características clínicas em pacientes com mutações ocasionando haploinsuficiência de *MED13L*.

Quadro clínico	Número de pacientes	Frequência
DI ou atraso de desenvolvimento	26/26	100%
Atraso grave de fala	11/22	50%
Características autistas	5/25	20%
Hipotonia	15/25	60%
Problemas de coordenação	9/24	38%
Convulsões ou EEG anormal	4/25	16%
Comportamento agressivo	2/22	9%
MRI anormal	6/16	38%
Estrabismo	3/24	13%
Dismorfismo facial	21/21	100%
<i>testa larga/proeminente</i>	16/21	76%
<i>orelhas de implantação baixa</i>	11/21	52%
<i>estreitamento bitemporal</i>	9/18	50%
<i>sobrancelhas horizontais</i>	6/21	29%
<i>Inclinação ascendente das fissuras palpebrais</i>	13/21	62%
<i>ponte nasal deprimida/plana</i>	18/21	86%
<i>nariz bulboso</i>	15/21	71%
<i>queixo anormal (micro/pro/retrognatia)</i>	12/22	55%
<i>macroglossia</i>	6/21	29%
Defeitos cardíacos congênitos complexos	3/23	13%
Forame oval patente/defeito do septo ventricular	4/23	17%
Parâmetros de crescimento anormais	10/20	50%
<i>microcefalia</i>	3/20	15%
<i>macrocefalia</i>	1/20	5%
<i>peso baixo ao nascimento</i>	3/20	15%
<i>peso baixo pós-natal</i>	3/20	15%
<i>excesso de peso pós-natal</i>	1/20	5%

EEG=eletroencefalograma, MRI=imagem por ressonância magnética.

Neste trabalho, nós relatamos uma família com dois irmãos com DI moderada, ausência de fala e dismorfismos faciais característicos, quadro clínico causado por variante *nonsense* em heterozigose no exon 20 do gene *MED13L* identificada por sequenciamento completo de exoma.

Descrição da Família 13 (pacientes 13a e 13b)

A Família 13 foi atendida no Centro de Assistência Odontológica aos Portadores de Necessidades Especiais (CAOE), Unidade Auxiliar de Estrutura Complexa da Faculdade de Odontologia do Campus de Araçatuba (FOA) da

Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), como parte de *projeto PIPE nº 2012/50981-5*, coordenado pela Ms. Francine Campagnari. Os pacientes são nascidos de pais saudáveis e não consanguíneos. O heredograma desta família é mostrado na Figura 40.

Figura 40. Heredograma da Família 13, na qual uma variante provavelmente patogênica no gene *MED13L* segrega com DI num padrão de herança autossômica dominante.

Paciente 13a (II.2)

Nascimento a termo de parto normal. Apresentou atraso de DNPM: sentou com 1 ano, engatinhou aos 2 anos e andou com 3 anos de idade. Examinado aos 33 anos de idade (Figura 41), apresentava DI moderada e ausência de fala; não lê mas compreende comandos. Apresentava dismorfismos faciais compatíveis com os descritos para a síndrome de haploinsuficiência de *MED13L*: estreitamento bitemporal, face triangular, fissuras palpebrais oblíqua e ponte nasal baixa.

Figura 41. Face do paciente 13a (II.2).

Paciente 13b (II.9)

Nascimento a termo de parto normal. Apresentou atraso de DNPM: sentou com 1 ano, engatinhou com 1 ano e 2 meses e andou com 2 anos de idade. Examinada aos 27 anos de idade (Figura 42), apresentava DI moderada e fala muito comprometida. Apresentava dismorfismos faciais compatíveis com a síndrome de haploinsuficiência de *MED13L*: estreitamento bitemporal, face triangular, fissuras palpebrais oblíquas, ponte nasal baixa e micrognatismo.

Figura 42. Face da paciente 13b (II.9).

Resultados

A análise dos dados do exoma dos irmãos resultou na identificação de uma variante em heterozigose localizada no exon 20 do gene *MED13L* [NM_015335.4:c.4417C>T:p.Q1473X]. Esta variante não está relatada em nenhum dos bancos de dados populacionais consultados e gera um códon de parada (*nonsense*).

Na Figura 43 podemos verificar na imagem do arquivo BAM a presença da variante no gene *MED13L* em heterozigose nos irmãos afetados II.2 e II.9.

Figura 43. Imagem do arquivo BAM mostrando nas *reads* a posição da variante g.116422099G>A no exon 20 do gene *MED13L* visualizada no *software* VarSeq (Golden Helix). (A) I.3 (afetado) e (B) II.9 (afetada): observa-se a presença da base mutada (em vermelho) em heterozigose; a estrela vermelha na sequência proteica indica a posição do aminoácido trocado (Gln1473).

Discussão

Nesta família composta por dois irmãos afetados, um do sexo masculino e uma do feminino, apresentando DI moderada, ausência de fala e características faciais dismórficas, foi detectada uma nova variante *nonsense* no exon 20 do gene *MED13L*, considerada patogênica. Deleções intragênicas e mutações de

perda de função de *MED13L* foram associadas a pacientes com DI e sinais clínicos característicos, com ou sem defeitos cardíacos (OMIM# 616789).

Em trabalho recente de Topping e col. (2018), oito novos pacientes com DI relacionada ao gene *MED13L* são descritos, com ênfase na clínica e revisão da literatura acerca dos aspectos genéticos e fenotípicos. Neste estudo, foram revisados dados de 69 pacientes, todos com mutações *de novo* patogênicas ou provavelmente patogênicas. Foram detectadas microdeleções intragênicas ou englobando o gene inteiro (~27%) e variantes de perda de função (76%), essas últimas ocorrendo ao longo de todo o gene (Topping et al. 2018). Adicionalmente, mutações *missense* também foram descritas, aparentemente relacionadas a um fenótipo mais grave, frequentemente com epilepsia e ausência de habilidade de caminhar. Os autores concluem que, embora a maioria dos sinais clínicos da DI associada a *MED13L* seja inespecífica, atraso de desenvolvimento, comprometimento de fala, ponta do nariz bulbosa e alterações bucais, como macroglossia e boca aberta, ocorrem com frequência, embora as alterações bucais possam ser menos evidentes em indivíduos adultos (Smol et al. 2018).

O quadro clínico de nossos pacientes é compatível com o descrito na literatura, considerando que apresentam os sinais mais frequentes já descritos em associação a mutações em *MED13L* (atraso de desenvolvimento, DI moderada, grave comprometimento de fala e alguns dismorfismos faciais característicos) e que o fenótipo cardíaco apresenta penetrância incompleta. A mutação detectada (p.Q1473X) é *nonsense* e ocorre no exon 20 de *MED13L*, no qual outras duas mutações *nonsense* já foram descritas, confirmando o efeito provável de perda de função de mutações deste tipo nessa região da proteína (Topping et al. 2018). Portanto, é altamente provável que a variante detectada nos irmãos tenha como efeito a perda de função de um dos alelos de *MED13L*, sendo, assim, responsável pelo quadro de DI síndrômica.

O estudo de segregação familiar ainda não pode ser finalizado, uma vez que a família não mora em São Paulo. Estamos em contato para coleta de amostras de DNA da mãe e da irmandade clinicamente normais, visando o estudo de segregação da variante. Por definição, um dos genitores deve ser portador da variante. Tal situação permanece por ser esclarecida, uma vez que penetrância

incompleta não foi descrita para esse gene e os genitores não são afetados. Assim, temos a hipótese de que a variante em *MED13L* estaria em mosaicismo gonadal em um dos genitores. Esse resultado seria compatível com a baixa frequência de afetados na irmandade de 11 irmãos.

Esse é o primeiro relato de um caso familiar de perda de função no gene *MED13L*, a ser finalizado com o estudo de segregação para completo entendimento.

VI.3.9. FAMÍLIA 25 – UM CASO FAMILIAL DE DI SEGREGANDO COM PADRÃO DE HERANÇA AUTOSSÔMICO DOMINANTE ASSOCIADO A UMA VARIANTE NO GENE *MBD5*

Resumo

Descrevemos uma família com cinco irmãos do sexo masculino afetados por DI leve e comportamento agressivo, filhos de mãe com deficiência de aprendizado e de dois pais diferentes. Nesta família, detectamos uma variante não descrita previamente, do tipo mudança de nucleotídeo (*missense*) no gene *MBD5* (OMIM *611472), que segrega com o quadro clínico. Este gene codifica uma proteína da família *methyl-CpG-binding domain* (MBD), que desempenha funções críticas na regulação e desenvolvimento transcricional. Mutações e deleções parciais e totais de *MBD5* foram associadas à deficiência intelectual autossômica dominante 1 (OMIM #156200 - Mental Retardation, Autosomal Dominant 1; MRD1).

A variante *missense* c.80G>A (p.Arg27His) detectada nessa família não está documentada em nenhum dos bancos de variantes genéticas populacionais consultados. Segundo as recomendações do ACMG (Colégio Americano de Genética Médica e Genômica - *The American College of Medical Genetics and Genomics*) e da AMP (Associação de Patologia Molecular - *The Association for Molecular Pathology*) de 2015, a classificação dessa variante seria provavelmente patogênica e, portanto, a variante em *MBD5* aparece como a provável causa de DI nessa família.

O gene *MBD5*

O gene *MBD5*, localizado em 2q23.1, codifica o membro 5 da família de genes que apresenta um domínio de ligação a metil-CpG (MBD); tem 15 exons e codifica no mínimo duas isoformas proteicas. Os dois transcritos principais carregam sempre os exons de 1-5, além de outros seguintes, porém tais exons contém apenas sequências 5'UTR que não estarão, portanto, traduzidas nas isoformas proteicas. A isoforma proteica 1, mais longa, tem 1448 aminoácidos codificados pelos exons 6–15. Esta isoforma apresenta dois domínios conservados, um deles MBD, e outro chamado prolina-triptofano-triptofano-prolina (PWWP); ambos os domínios são encontrados em proteínas que tem a habilidade de se associar à cromatina. A isoforma 2 é codificada pelos exons 6–9, resultando em uma proteína de 851 aminoácidos que não contém o domínio PWWP. As isoformas 1 e 2 compartilham os aminoácidos 1-841, contém um segmento rico em prolina e sinal de localização nuclear. Estudos de expressão mostraram que as duas isoformas são expressas em todos os tecidos e são altamente expressas no cérebro. A proteína MBD5 está provavelmente envolvida na regulação negativa da expressão gênica e no controle da produção de proteínas envolvidas em aprendizado, memória e comportamento. Diferencialmente, a isoforma 1 tem maior expressão nos testículos, enquanto a isoforma 2 nos ovários (Laget et al. 2010).

Na Figura 44, retirada do trabalho de Laget e col. (2010), temos os esquemas mostrando os exons e os domínios das isoformas 1 e 2.

Figura 44. (A) Exons codificadores do gene *MBD5* são mostrados em vermelho nas duas isoformas proteicas. (B) Organização da proteína MBD5. Domínio MBD (ligação a metil-CpG), segmento rico em prolina (P rich); domínio PWWP (prolina-triptofano-triptofano-prolina) e sinal de localização nuclear (NLS).

Em 2007, Wagenstaller e col. foram os primeiros a associar uma deleção do gene *MBD5* a DI leve a grave. Numa coorte de 67 crianças com DI idiopática foi detectada uma deleção *de novo* incluindo os sete primeiros exons do gene *MBD5* em uma criança com DI grave. Neste estudo, amostras adicionais de DNA de 415 crianças encaminhadas para o diagnóstico molecular do gene *FMR1*, mas para as quais não foram encontradas mutações, foram utilizadas para a análise de sequências do gene *MBD5*; foram detectadas quatro variantes *missense* (Wagenstaller et al. 2007).

No estudo colaborativo internacional realizado por Talkowski e col. (2011) foram compilados 65 pacientes com deleções ou translocações envolvendo o segmento cromossômico 2q23.1 e a região crítica para o fenótipo de tais pacientes foi associada à haploinsuficiência do gene *MBD5*. As principais características clínicas encontradas nesses pacientes foram DI, atraso de DNPM, problemas de comportamento, falta de concentração e sinais do espectro autista. Nenhuma alteração no número de cópias de *MBD5* foi observada em 7878 controles, indicando que as alterações no *MBD5* são altamente penetrantes. Os autores também compararam variantes detectadas por sequenciamento na região codificadora de *MBD5* de 747 indivíduos com distúrbios do espectro autista (ASD) com 2043 indivíduos não ASD. Verificaram que a variante *missense* p.79Gly>Glu, localizada no domínio MBD, tem uma frequência significativamente mais alta em pacientes ASD (Talkowski et al. 2011).

Até agora, vários trabalhos foram publicados associando DI, epilepsia e ASD a deleções e variantes do gene *MBD5* (Carvill et al. 2013; Han et al. 2017; Ishizuka et al. 2016; Marco et al. 2018; Mullegama and Elsea 2016; Wang et al. 2017; Woodbury-Smith et al. 2017); os pacientes apresentavam quadro clínico variável, incluindo alterações de comportamento como déficit de atenção, agressividade e distúrbios do sono.

Neste trabalho, nós relatamos uma família cuja nova variante *missense* no exon 6 do gene *MBD5* segrega com padrão de herança dominante em irmandade de homens afetados por DI leve e comportamento agressivo, e sua mãe, portadora obrigatória, que apresenta dificuldade de aprendizado.

Descrição da Família 25

A família aqui apresentada foi averiguada pelas Dras. Ana C. Victorino Krepisch e Angela M. Vianna-Morgante no Serviço de Aconselhamento Genético do Laboratório de Genética Humana do Departamento de Genética e Biologia Evolutiva, Instituto de Biociências da Universidade de São Paulo (IB-USP). Nesta família foram estudadas amostras dos irmãos III.1, III.3 e III.4 e de sua mãe (II.2); examinados aos 25, 20 e 18 anos de idade, respectivamente, os pacientes apresentavam DI moderada, não sabendo ler ou escrever, em associação a um comportamento agressivo. A mãe (II.2), assim como seus filhos, não lê ou escreve, embora apresentasse durante a entrevista boa desenvoltura, indicando ter manifestado dificuldade de aprendizado, mas com um quadro de DI leve; relatou que seu pai também tinha dificuldade de aprendizado e comportamento violento. O heredograma desta família é mostrado na Figura 45, sugerindo inicialmente um padrão de herança dominante, ou ligado ao cromossomo X.

Figura 45. Heredograma da Família 25, na qual uma variante provavelmente patogênica no gene *MBD5* aparentemente segrega com DI e agressividade num padrão de herança autossômica dominante.

Resultados

A análise do exoma de dois meio-irmãos (III.1 e III.4) resultou na identificação de uma variante *missense* em heterozigose no exon 6 do gene *MBD5* [NM_018328:c.80G>A:p.(Arg27His)]. Esta variante *missense* não foi previamente relatada em nenhum dos grandes bancos de dados populacionais e sua patogenicidade foi avaliada usando diferentes algoritmos, sendo classificada como “danosa” por SIFT, “possivelmente danosa” por Polyphen-2 e “danosa” por MutationTaster.

A presença da variante em heterozigose foi confirmada por sequenciamento Sanger nos irmãos afetados (III.1, III.3 e III.4) e na mãe (Figura 46). Os outros afetados dessa família não puderam ser avaliados até o momento.

Figura 46. Sequenciamento Sanger em amostras de DNA genômico mostrando a variante *missense* c.80G>A encontrada no gene *MBD5*. Presença da variante em heterozigose na mãe (II.2) e nos irmãos afetados (II.1, II.3 e II.4); variante ausente no controle masculino.

Discussão

Nesta família, foi identificada uma nova variante *missense* no gene *MBD5* segregando com fenótipo de DI moderada e comportamento agressivo. A presença da variante pode ser averiguada em três dos irmãos afetados e em sua mãe portadora, que apresenta quadro clínico mais leve. Assim, tal variante apresenta um padrão de manifestação dominante, aparentemente com penetrância completa, mas com expressividade variável. Nesta família, os outros dois filhos afetados não compareceram ao atendimento.

A haploinsuficiência de *MBD5* é um distúrbio do neurodesenvolvimento que compreende atraso global do desenvolvimento/deficiência intelectual (DI), comprometimento grave da fala, convulsões, distúrbios do sono e problemas comportamentais com características autistas (Jaillard et al. 2009; Noh and Graham 2012; Talkowski et al. 2011; van Bon et al. 2010). Mais comumente, a haploinsuficiência que causa o fenótipo é decorrente de deleções de 2q23.1 que englobam total ou parcialmente o gene *MBD5* (~90% dos indivíduos relatados), ao passo que deleções intragênicas de *MBD5* afetando 1 ou mais exons correspondem a cerca de 5% dos casos e variantes patogênicas.

Duas variantes *de novo* de perda de função foram reportadas afetando o domínio MBD, c.890_891delTA (p.Ile297Thrfs), detectada numa coorte de pacientes com ASD (O'Roak et al. 2012), e c.150del (p.Thr52Hisfs*31), detectada em paciente com DI (Kleefstra et al. 2012). Outra variante de perda de função, c.254_255delGA (p.Arg85Asnfs*6), foi relatada segregando em um pai portador afetado por DI leve (QI de 73) e também em sua filha, gravemente afetada, exibindo QI de 48 (Han et al. 2017). A nova variante *missense* detectada neste estudo, c.80G>A (p.Arg27His), afeta um domínio muito importante para a função da proteína, o domínio MBD do gene *MBD5*; apenas uma outra variante *missense* (rs34995577) afetando este domínio foi reportada, identificada entre pacientes com ASD, e sua patogenicidade não foi confirmada (Talkowski et al. 2011).

Os dados atuais mostram que os sinais clínicos seriam decorrentes da diminuição da dose da proteína MBD5, independentemente do mecanismo mutacional; de maneira intrigante, os estudos de correlação genótipo-fenótipo mostram não haver distinção clara de quadro clínico entre pacientes com

deleções e pacientes com mutações pontuais, embora, como esperado, deleções afetando muitos genes podem resultar em quadro mais grave. Em alguns casos, os probandos herdaram a variante patogênica de forma autossômica dominante de um genitor com sinais clínicos mais leves, sugerindo a ocorrência de penetrância incompleta ou expressividade variável (Hodge et al. 2014; Talkowski et al. 2011), com grande heterogeneidade fenotípica.

Utilizando os critérios da ACMG para classificação de variantes germinativas, essa variante foi classificada como provavelmente patogênica em virtude de não estar descrita em nenhum banco populacional, segregar com o fenótipo em vários afetados que têm fenótipo compatível com o quadro de DI já reportado para gene conhecido, com expressividade variável, também já averiguada nas famílias relatadas previamente; adicionalmente, o tipo de variante (*missense*) já foi descrito como causa de DI neste gene e a variante está localizada no domínio *methyl-binding* da proteína, no qual outras variantes patogênicas foram identificadas. Outro aspecto que compôs a avaliação de patogenicidade foi a análise *in silico* com algoritmos, mais uma linha que apoia a hipótese de impacto clínico da variante *missense* detectada. Assim, nossas descobertas sugerem fortemente que a variante detectada é a causa do quadro clínico dessa família.

Esse caso é de certa forma incomum por apresentar vários afetados na irmandade, o que pode ser parcialmente atribuído ao fato de que a DI apresentada pela mãe portadora não foi suficiente para reduzir seu *fitness* reprodutivo. O estudo dessa família contribui para ilustrar o espectro de mutações associado ao quadro de DI causado por haploinsuficiência de *MBD5*, assim como evidencia a variabilidade fenotípica associada a tais variantes.

VII. Discussão geral

VII. Discussão geral

Estudos estimaram a prevalência da DI em aproximadamente 1% nos países desenvolvidos (Chiurazzi 2011; Maulik et al. 2011); trata-se de uma condição clínica complexa e altamente heterogênea que resulta de causas genéticas, ambientais ou da combinação de ambas. Estudamos 25 famílias com dois ou mais afetados por DI, síndrômica ou não, com o objetivo de identificar a causa da deficiência intelectual nessas famílias brasileiras. Iniciamos nossos estudos afastando a síndrome do cromossomo X frágil e desequilíbrios genômicos submicroscópicos como causa de DI.

A técnica de sequenciamento do exoma completo (WES) foi o método empregado neste estudo para detectar as variantes patogênicas ou possivelmente patogênicas responsáveis pela DI em cada família. WES é atualmente a plataforma mais popular para a descoberta de genes que causam doenças raras e investiga a porção codificadora que corresponde a aproximadamente 1% do genoma humano (o exoma). Apesar da limitação de excluir alelos não-codificantes da análise, o WES é, por várias razões, uma estratégia bem justificada para identificar variantes de nucleotídeo único (SNVs) responsáveis por fenótipos mendelianos. Uma grande fração de variantes raras, que alteram a proteína, como substituições de base única *missense* ou *nonsense*, *splicing* e pequenas deleções-inserções (*indels*), têm consequências funcionais e podem ser deletérias (Wright et al. 2018). As variantes raras detectadas nas famílias foram classificadas como patogênicas, provavelmente patogênicas ou variantes de significado indeterminado (VUS), de acordo com os critérios da ACMG (Richards et al. 2015).

O estudo de Gilissen e col. (2014) sugere que o uso de sequenciamento de genoma completo seria significativamente mais eficiente do que o de sequenciamento de exoma. No entanto, as variantes encontradas nesse estudo usando sequenciamento de genoma eram, na verdade, exônicas, e os resultados refletem mais um problema de captura e amplificação das sequências-alvo do que de detecção de variantes não codificantes propriamente ditas (Gilissen et al. 2014). Na verdade, ainda é difícil a interpretação de variantes em regiões não

codificadoras. Uma vantagem diagnóstica do uso de sequenciamento de genoma em vez de exoma seria a melhor detecção de variações em número de cópias, embora vários grupos, incluindo o nosso, tenham demonstrado a viabilidade e precisão no uso de sequenciamento alvo para detecção de variantes de número de cópias genômicas (Vetro et al. 2018; Villela et al. 2017). No presente estudo, os pacientes haviam sido previamente triados por *arrays* genômicos para alterações em número de cópias.

Na coorte estudada pudemos observar uma prevalência de afetados do sexo masculino (1,8:1). Na literatura observa-se um desvio da razão sexual de 1,4:1 para DI grave e 1,9:1 para DI moderada, que evidenciam a prevalência sabidamente maior de DI em meninos do que em meninas (Ropers 2010). Nosso objetivo inicial era a busca de variantes em genes com manifestação autossômica recessiva e, por essa razão, optamos pelo estudo preferencial de irmandades em que ao menos um dos irmãos afetados fosse do sexo feminino. Entre as primeiras famílias compostas por irmandades de afetados homem-mulher, duas tiveram que ser excluídas da coorte. Uma das famílias foi excluída porque os irmãos eram portadores da síndrome do cromossomo X frágil. Na outra, foi detectada uma deleção intragênica de 53 kb do gene *SNRPN* (OMIM *182279), região que contém o centro de *imprinting* das síndromes de Prader-Willi e Angelman. A mãe dos afetados é portadora da deleção e, portanto, os irmãos foram diagnosticados com a síndrome de Angelman. Em particular, no nosso centro de atendimento, há um enriquecimento de irmandades do sexo masculino dada a importante linha de pesquisa com a síndrome do cromossomo X frágil e outras condições ligadas ao X da Profa. Angela M. Vianna Morgante. Devido à prevalência de homens afetados por DI e à dificuldade de encontrar famílias com ao menos uma afetada do sexo feminino, nós decidimos incluir no estudo famílias compostas por mais de um afetado independentemente do sexo, não favorecendo particularmente o padrão de herança autossômica, embora ainda teoricamente favorecendo herança recessiva. Assim, foram selecionadas 15 famílias compostas por afetados de ambos os sexos e outras 10 famílias contendo apenas afetados do sexo masculino.

Segundo a revisão de Vissers e col. (2016), mais de 700 genes foram identificados em estudos de DI apresentando diversos padrões de herança (Vissers et al. 2016). De acordo com Kochinke e col. (2016), 62% dos genes mutados em DI apresentam padrão de herança autossômica recessiva, 21% autossômica dominante e 16% ligada ao X, sendo que 101 genes (16%) estão envolvidos em mais de uma doença (Kochinke et al. 2016). Na Tabela 9 temos um resumo da distribuição dos diferentes tipos de padrão de herança identificados na nossa coorte. No total, tivemos uma taxa de diagnóstico molecular para DI de 36% (9/25); excluindo-se as duas famílias nas quais encontramos uma VUS e uma variante em um gene ainda candidato a causar DI, essa taxa diminui para 28% (7/25). Esses valores são compatíveis com os estabelecidos na literatura, que variam de 16% a 42% (de Ligt et al. 2012; Deciphering Developmental Disorders 2015, 2017; Rauch et al. 2012; Vissers et al. 2016; Wright et al. 2018).

Tabela 9. Número de famílias com variantes que potencialmente explicam a deficiência intelectual e respectivos padrões de herança.

Tipo de herança	Nº de famílias (%)
Autossômica dominante	3 (12%)
Autossômica recessiva	1 (4%)
Ligada ao cromossomo X	5 (20%)
Total	9 (36%)

Para a filtragem das variantes, selecionamos as que co-segregavam no par de irmãos estudados e que não apareciam frequentemente na população ou nas amostras sequenciadas. Não incluímos a filtragem por número mínimo de *reads* porque, embora a média de cobertura vertical *on-target* tenha sido de 36x, a cobertura horizontal não pôde ser considerada satisfatória e a filtragem adicional por número de *reads* comprometeria ainda mais nossos resultados. Apesar dessa avaliação, das nove famílias positivas (Tabela 9), em oito as variantes detectadas tiveram uma cobertura vertical ≥ 20 e todas as variantes identificadas foram confirmadas por Sanger. Não tivemos resultados falso-positivos, porém é provável que tenhamos tido falso-negativos. A cobertura baixa em nossas

amostras pode significar que o número de casos com identificação de potenciais mutações apresentado na tabela 9 represente uma subestimativa.

As 16 famílias nas quais não detectamos variantes que expliquem o quadro clínico associado a DI serão reanalisadas. A princípio, devemos analisar cada irmão do par individualmente e verificar visualmente no arquivo BAM se eventuais variantes relevantes estão presentes no outro. Adicionalmente, aumentaremos o sequenciamento dos pares de irmãos para minimizar a chance de que variantes sejam “perdidas” por não estarem adequadamente sequenciadas.

É importante notar que parte das variantes que consideramos patogênicas ou provavelmente patogênicas eram *missense* e ainda não descritas. O estudo de segregação foi essencial para a classificação dessas novas variantes como possivelmente patogênicas/patogênicas. Na ausência de repetição na família, as variantes teriam que ser classificadas como de significado incerto, comprometendo o diagnóstico e aconselhamento genético da família. Na família 1, estudos funcionais demonstraram que a mutação *missense* que segregava com a DI afetava o funcionamento da proteína; estudos funcionais são o padrão-ouro para se constatar a função comprometida do gene, porém seu uso na rotina diagnóstica é ainda inviável.

Quando examinamos a distribuição dos padrões de herança das variantes identificadas podemos notar que genes ligados ao cromossomo X tiveram a maior taxa de detecção, especificamente 20% (5/25); retirando as famílias nas quais encontramos uma VUS e uma variante em gene ainda candidato, a taxa de genes do cromossomo X na nossa amostra diminui para 12% (3/25). Este valor se aproxima da taxa de diagnóstico de 10% para a herança ligada ao X encontrada na literatura (Musante and Ropers 2014; Raymond et al. 2009; Ropers and Hamel 2005; Tarpey et al. 2009). Também devemos observar que três das famílias cujas variantes foram encontradas no cromossomo X são compostas apenas por afetados do sexo masculino e nas outras duas, as mulheres apresentam um quadro mais leve de DI.

A taxa de detecção de famílias com variantes que segregam com DI de acordo com padrão de herança autossômica dominante foi de 12% (3/25). Esse valor foi inesperadamente alto, já que o estudo de irmandades deveria favorecer

o padrão de herança recessiva. É importante salientar que nas duas famílias nas quais a segregação da variante foi investigada, as mães são portadoras da variante e apresentam quadro clínico mais leve. Na mãe portadora da Família 8, o quadro clínico mais leve é explicado pela presença da variante em mosaico muito baixo, 0,13% em células do sangue periférico. Já na mãe da Família 25, não temos evidência de mosaicismo e o fenótipo mais leve poderia ser atribuído a homens apresentarem um limiar inferior às mulheres para manifestação de distúrbios de neurodesenvolvimento (Jacquemont et al. 2014). O fato de genitores serem afetados aumenta a taxa de detecção de variantes patogênicas em genes de herança autossômica dominante (Wright et al. 2015). Na literatura, a DI autossômica dominante quase sempre se deve a mutações *de novo* que podem explicar de 13-35% dos casos de pacientes com DI grave (Vissers et al. 2016); entretanto, chamamos a atenção para o fato do mosaicismo na mãe da Família 8 ter sido detectado apenas pelo uso da técnica de PCR digital, após a ocorrência de duas filhas afetadas evidenciarem a existência de um portador obrigatório; se fosse um caso isolado, teria provavelmente sido considerado *de novo*. Essa família ilustra a contribuição, ainda pouco considerada, do mosaicismo germinativo como causa de DI com padrão de herança dominante. A presença de genitores afetados nas famílias estudadas tem impacto claro na detecção de genes de DI de herança autossômica dominante na coorte estudada. Dadas as condições sócio-culturais desfavoráveis às quais parte significativa da população brasileira está sujeita, a DI leve pode passar despercebida e não representar um impeditivo para a constituição de famílias.

Os estudos realizados com a população do Oriente Médio têm revelado um aumento de genes de herança recessiva associados à DI (Musante and Ropers 2014; Najmabadi et al. 2011; Najmabadi et al. 2007), embora cada gene explique a DI em um número pequeno de famílias. Em populações nas quais os casamentos consanguíneos são comuns, as alterações autossômicas recessivas são responsáveis pela maioria dos casos de DI, e mesmo em populações nas quais não ocorre casamento consanguíneo, a herança recessiva estaria associada a 13-24% dos casos. No presente estudo, embora formulado para privilegiar o padrão de herança recessiva, detectamos apenas uma família cuja DI foi

associada a esse padrão, apresentando variantes em heterozigose composta. A taxa de diagnóstico foi inferior ao relatado na literatura mesmo com a investigação de irmandades com mais de um afetado. Provavelmente, uma estratégia mais eficiente para a detecção de genes associados à DI sindrômica em irmandades seja a busca em afetados que apresentem características dismórficas ou anomalias congênitas similares e específicas, mas nem sempre esses casos estão disponíveis.

Apontamos nessa discussão motivos técnicos como uma das causas para a falha na identificação da base molecular da DI em 64% das famílias; outro fator relevante para diminuir a detecção de variantes que possam explicar doenças de herança mendeliana é o mosaicismo. Variantes em mosaico somático podem surgir de erros que ocorrem durante a replicação do DNA mitótico. Essas variantes surgem de eventos mutacionais não corrigidos e se acumulam continuamente à medida que as células se dividem durante o desenvolvimento (Campbell et al. 2015). O reconhecimento de uma variante em mosaico é relacionado à quantidade de *reads* que estão cobrindo uma determinada base sequenciada. Apesar do WES aumentar a capacidade de identificar variantes em mosaico, discernir variantes heterozigotas de erros de sequenciamento é um desafio, pois o mosaicismo pode estar presente numa frequência baixa. As variantes em mosaico somático podem ter consequências clínicas para o portador ou, como detectado na mãe da Família 8, contribuir para o risco de distúrbios genômicos nos filhos.

Outros fatores que explicam a não detecção de variantes em parte das famílias estudadas são: a existência de genes ainda não associados a fenótipos e padrões de herança mais complexos, não mendelianos. Já foi demonstrado que alterações em regiões não codificantes, como a expansão de trinucleotídeos ou a perda de regiões regulatórias (Chen et al. 2009; Colak et al. 2014; Lecointre et al. 2009) e elementos regulatórios dentro de sequências intrônicas e UTRs (Chatterjee and Pal 2009; Chen et al. 2006a, b) podem impactar o fenótipo. Estudos têm mostrado variantes genéticas dentro das 3'UTRs, em locais alvo de microRNA, causadoras ou associadas a um risco aumentado de doença (Annala et al. 2018; Bandiera et al. 2010; Chen et al. 2018; Mir et al. 2018). A variante

de significado incerto detectada na região 3'UTR do gene *PHF6* que segrega com DI na Família 11, se confirmada a patogenicidade, seria um exemplo desse evento. Porém, além do custo ainda elevado de sequenciamento de exoma completo, nosso conhecimento é bastante limitado para interpretar variantes não codificantes. O uso mais intenso das tecnologias de sequenciamento de fragmentos longos permitirá averiguar não apenas expansões de trinucleotídeos, mas também rearranjos estruturais. O uso de transcriptomas também pode ser muito informativo no caso de genes rearranjados, embora a detecção do RNA de interesse seja dependente do tipo de tecido usado e padrão de expressão do gene. Essencial, como em todas as outras áreas de genômica atuais, é a constituição de consórcios que compilam e disponibilizam dados genômicos e clínicos, além de desenvolver ferramentas de bioinformática para a análise desses dados.

VIII. Conclusão

VIII. Conclusão

1. A taxa de diagnóstico molecular na coorte estudada foi de 28-36%. Embora esses valores estejam de acordo com o que é encontrado na literatura com o emprego do WES, foi abaixo do esperado considerando que a amostra era de casos familiares, enquanto esses valores da literatura foram estimados sobretudo em casos isolados. A cobertura sub-ótima dos exomas de nossas amostras deve ter contribuído para uma taxa de detecção menor.
2. A frequência inesperadamente alta de DI de herança autossômica dominante em casos familiares mostra a importância de mosaicismismo gonadal nesse tipo de herança e evidencia que, na ausência de recorrência, muitos casos isolados podem erroneamente serem considerados como mutações novas.
3. A presença de mães afetadas por DI leve ou dificuldade de aprendizado na nossa coorte é outro fator que contribui para a alta taxa de DI de herança autossômica dominante.
4. Identificamos um gene candidato, o *ARHGAP4*, que pertence a uma família de genes já relacionados a DI. Estudos funcionais ou identificação de outros afetados são necessários para avaliar a contribuição deste gene como causa de DI.
5. Dez variantes novas em genes conhecidos de DI foram descritas nesse trabalho.
6. Em uma família, ilustramos a contribuição do estudo funcional para a determinação da patogenicidade de uma variante *missense* que segrega com a DI.

IX. Referências bibliográficas

IX. Referências bibliográficas

- Abbasi-Moheb L, Mertel S, Gonsior M, Nouri-Vahid L, Kahrizi K, Cirak S, Wieczorek D, Motazacker MM, Esmaeeli-Nieh S, Cremer K, Weissmann R, Tzschach A, Garshasbi M, Abedini SS, Najmabadi H, Ropers HH, Sigrist SJ, Kuss AW (2012) Mutations in NSUN2 cause autosomal-recessive intellectual disability. *Am J Hum Genet* 90: 847-55. doi: 10.1016/j.ajhg.2012.03.021
- Abdullah U, Farooq M, Mang Y, Marriam Bakhtiar S, Fatima A, Hansen L, Kjaer KW, Larsen LA, Faryal S, Tommerup N, Mahmood Baig S (2017) A novel mutation in CDK5RAP2 gene causes primary microcephaly with speech impairment and sparse eyebrows in a consanguineous Pakistani family. *Eur J Med Genet* 60: 627-630. doi: 10.1016/j.ejmg.2017.07.017
- Ahmad I, Baig SM, Abdulkareem AR, Hussain MS, Sur I, Toliat MR, Nurnberg G, Dalibor N, Moawia A, Waseem SS, Asif M, Nagra H, Sher M, Khan MMA, Hassan I, Rehman SU, Thiele H, Altmuller J, Noegel AA, Nurnberg P (2017) Genetic heterogeneity in Pakistani microcephaly families revisited. *Clin Genet* 92: 62-68. doi: 10.1111/cge.12955
- Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW (1992) Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. *Am J Hum Genet* 51: 1229-39.
- Annala M, Taavitsainen S, Vandekerckhove G, Bacon JVW, Beja K, Chi KN, Nykter M, Wyatt AW (2018) Frequent mutation of the FOXA1 untranslated region in prostate cancer. *Commun Biol* 1: 122. doi: 10.1038/s42003-018-0128-1
- Asadollahi R, Oneda B, Sheth F, Azzarello-Burri S, Baldinger R, Joset P, Latal B, Knirsch W, Desai S, Baumer A, Houge G, Andrieux J, Rauch A (2013) Dosage changes of MED13L further delineate its role in congenital heart defects and intellectual disability. *Eur J Hum Genet* 21: 1100-4. doi: 10.1038/ejhg.2013.17
- Asadollahi R, Zweier M, Gogoll L, Schiffmann R, Sticht H, Steindl K, Rauch A (2017) Genotype-phenotype evaluation of MED13L defects in the light of a novel truncating and a recurrent missense mutation. *Eur J Med Genet* 60: 451-464. doi: 10.1016/j.ejmg.2017.06.004
- Ballif BC, Rosenfeld JA, Traylor R, Theisen A, Bader PI, Ladda RL, Sell SL, Steinraths M, Surti U, McGuire M, Williams S, Farrell SA, Filiano J, Schnur RE, Coffey LB, Tervo RC, Stroud T, Marble M, Netzloff M, Hanson K, Aylsworth AS, Bamforth JS, Babu D, Niyazov DM, Ravnan JB, Schultz RA, Lamb AN, Torchia BS, Bejjani BA, Shaffer LG (2012) High-resolution array CGH defines critical regions and candidate genes for microcephaly, abnormalities of the corpus callosum, and seizure phenotypes in patients with microdeletions of 1q43q44. *Hum Genet* 131: 145-56. doi: 10.1007/s00439-011-1073-y
- Bandiera S, Hatem E, Lyonnet S, Henrion-Caude A (2010) microRNAs in diseases: from candidate to modifier genes. *Clin Genet* 77: 306-13. doi: 10.1111/j.1399-0004.2010.01370.x
- Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP, White TA, Stojanov P, Van Allen E, Stransky N, Nickerson E, Chae SS, Boysen G, Auclair D, Onofrio RC, Park K, Kitabayashi N, MacDonald TY, Sheikh K, Vuong T, Guiducci C, Cibulskis K, Sivachenko A, Carter SL, Saksena G, Voet D, Hussain WM, Ramos AH, Winckler W, Redman MC, Ardlie K, Tewari AK, Mosquera JM, Rupp N, Wild PJ, Moch H, Morrissey C, Nelson PS, Kantoff PW, Gabriel SB, Golub TR, Meyerson M, Lander ES, Getz G, Rubin MA, Garraway LA (2012) Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* 44: 685-9. doi: 10.1038/ng.2279

- Bentley DR (2000) The Human Genome Project--an overview. *Med Res Rev* 20: 189-96.
- Billuart P, Bienvenu T, Ronce N, des Portes V, Vinet MC, Zemni R, Roest Crollius H, Carrie A, Fauchereau F, Cherry M, Briault S, Hamel B, Fryns JP, Beldjord C, Kahn A, Moraine C, Chelly J (1998) Oligophrenin-1 encodes a rhoGAP protein involved in X-linked mental retardation. *Nature* 392: 923-6. doi: 10.1038/31940
- Bittel DC, Theodoro MF, Kibiryeva N, Fischer W, Talebizadeh Z, Butler MG (2008) Comparison of X-chromosome inactivation patterns in multiple tissues from human females. *J Med Genet* 45: 309-13. doi: 10.1136/jmg.2007.055244
- Bittles A (2001) Consanguinity and its relevance to clinical genetics. *Clin Genet* 60: 89-98.
- Bittles AH, Neel JV (1994) The costs of human inbreeding and their implications for variations at the DNA level. *Nat Genet* 8: 117-21. doi: 10.1038/ng1094-117
- Bochtler M, Ditzel L, Groll M, Hartmann C, Huber R (1999) The proteasome. *Annu Rev Biophys Biomol Struct* 28: 295-317. doi: 10.1146/annurev.biophys.28.1.295
- Bond J, Roberts E, Springell K, Lizarraga SB, Scott S, Higgins J, Hampshire DJ, Morrison EE, Leal GF, Silva EO, Costa SM, Baralle D, Raponi M, Karbani G, Rashid Y, Jafri H, Bennett C, Corry P, Walsh CA, Woods CG (2005) A centrosomal mechanism involving CDK5RAP2 and CENPJ controls brain size. *Nat Genet* 37: 353-5. doi: 10.1038/ng1539
- Borck G, Hog F, Dentici ML, Tan PL, Sowada N, Medeira A, Gueneau L, Thiele H, Kousi M, Lepri F, Wenzek L, Blumenthal I, Radicioni A, Schwarzenberg TL, Mandriani B, Fischetto R, Morris-Rosendahl DJ, Altmuller J, Reymond A, Nurnberg P, Merla G, Dallapiccola B, Katsanis N, Cramer P, Kubisch C (2015) BRF1 mutations alter RNA polymerase III-dependent transcription and cause neurodevelopmental anomalies. *Genome Res* 25: 155-66. doi: 10.1101/gr.176925.114
- Bouazzi H, Lesca G, Trujillo C, Alwasiyah MK, Munnich A (2015) Nonsyndromic X-linked intellectual deficiency in three brothers with a novel MED12 missense mutation [c.5922G>T (p.Glu1974His)]. *Clin Case Rep* 3: 604-9. doi: 10.1002/ccr3.301
- Bouyacoub Y, Zribi H, Azzouz H, Nasrallah F, Abdelaziz RB, Kacem M, Rekaya B, Messaoud O, Romdhane L, Charfeddine C, Bouziri M, Bouziri S, Tebib N, Mokni M, Kaabachi N, Boubaker S, Abdelhak S (2013) Novel and recurrent mutations in the TAT gene in Tunisian families affected with Richner-Hanhart syndrome. *Gene* 529: 45-9. doi: 10.1016/j.gene.2013.07.066
- Budny B, Badura-Stronka M, Materna-Kiryluk A, Tzschach A, Raynaud M, Latos-Bielenska A, Ropers HH (2010) Novel missense mutations in the ubiquitination-related gene UBE2A cause a recognizable X-linked mental retardation syndrome. *Clin Genet* 77: 541-51. doi: 10.1111/j.1399-0004.2010.01429.x
- Buffalo V (2011) Scythe <https://github.com/vsbuffalo/scythe>. <https://github.com/vsbuffalo/scythe>.
- Cafiero C, Marangi G, Orteschi D, Ali M, Asaro A, Ponzi E, Moncada A, Ricciardi S, Murdolo M, Mancano G, Contaldo I, Leuzzi V, Battaglia D, Mercuri E, Slavotinek AM, Zollino M (2015) Novel de novo heterozygous loss-of-function variants in MED13L and further delineation of the MED13L haploinsufficiency syndrome. *Eur J Hum Genet* 23: 1499-504. doi: 10.1038/ejhg.2015.19
- Campbell IM, Shaw CA, Stankiewicz P, Lupski JR (2015) Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet* 31: 382-92. doi: 10.1016/j.tig.2015.03.013
- Carter MT, Picketts DJ, Hunter AG, Graham GE (2009) Further clinical delineation of the Borjeson-Forssman-Lehmann syndrome in patients with PHF6 mutations. *Am J Med Genet A* 149A: 246-50. doi: 10.1002/ajmg.a.32624
- Carvalho E, Honjo R, Magalhaes M, Yamamoto G, Rocha K, Naslavsky M, Zatz M, Passos-Bueno MR, Kim C, Bertola D (2015) Schinzel-Giedion syndrome in two Brazilian

- patients: Report of a novel mutation in SETBP1 and literature review of the clinical features. *Am J Med Genet A* 167A: 1039-46. doi: 10.1002/ajmg.a.36789
- Carvill GL, Heavin SB, Yendle SC, McMahon JM, O'Roak BJ, Cook J, Khan A, Dorschner MO, Weaver M, Calvert S, Malone S, Wallace G, Stanley T, Bye AM, Bleasel A, Howell KB, Kivity S, Mackay MT, Rodriguez-Casero V, Webster R, Korczyn A, Afawi Z, Zelnick N, Lerman-Sagie T, Lev D, Moller RS, Gill D, Andrade DM, Freeman JL, Sadleir LG, Shendure J, Berkovic SF, Scheffer IE, Mefford HC (2013) Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. *Nat Genet* 45: 825-30. doi: 10.1038/ng.2646
- Chao MM, Todd MA, Kontny U, Neas K, Sullivan MJ, Hunter AG, Picketts DJ, Kratz CP (2010) T-cell acute lymphoblastic leukemia in association with Borjeson-Forssman-Lehmann syndrome due to a mutation in PHF6. *Pediatr Blood Cancer* 55: 722-4. doi: 10.1002/pbc.22574
- Charzewska A, Maiwald R, Kahrizi K, Oehl-Jaschkowitz B, Dufke A, Lemke JR, Enders H, Najmabadi H, Tzschach A, Hachmann W, Menzel C, Bienek M, Poznanski J, Nawara M, Chilarska T, Obersztyn E, Hoffman-Zacharska D, Gos M, Bal J, Kalscheuer VM (2018) The power of the Mediator complex - expanding the genetic architecture and phenotypic spectrum of MED12-related disorders. *Clin Genet*. doi: 10.1111/cge.13412
- Chatterjee S, Pal JK (2009) Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell* 101: 251-62. doi: 10.1042/BC20080104
- Chelly J, Mandel JL (2001) Monogenic causes of X-linked mental retardation. *Nat Rev Genet* 2: 669-80. doi: 10.1038/35088558
- Chen DH, Ma M, Scavina M, Blue E, Wolff J, Karna P, Dorschner MO, Raskind WH, Bird TD (2018) An 8-generation family with X-linked Charcot-Marie-Tooth: Confirmation Of the pathogenicity Of a 3' untranslated region mutation in GJB1 and its clinical features. *Muscle Nerve* 57: 859-862. doi: 10.1002/mus.26037
- Chen J, Wildhardt G, Zhong Z, Roth R, Weiss B, Steinberger D, Decker J, Blum WF, Rappold G (2009) Enhancer deletions of the SHOX gene as a frequent cause of short stature: the essential role of a 250 kb downstream regulatory domain. *J Med Genet* 46: 834-9. doi: 10.1136/jmg.2009.067785
- Chen JM, Ferec C, Cooper DN (2006a) A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes I: general principles and overview. *Hum Genet* 120: 1-21. doi: 10.1007/s00439-006-0180-7
- Chen JM, Ferec C, Cooper DN (2006b) A systematic analysis of disease-associated variants in the 3' regulatory regions of human protein-coding genes II: the importance of mRNA secondary structure in assessing the functionality of 3' UTR variants. *Hum Genet* 120: 301-33. doi: 10.1007/s00439-006-0218-x
- Chiurazzi P (2011) Mental retardation: is naming the real issue? *Am J Med Genet A* 155A: 974-5. doi: 10.1002/ajmg.a.33950
- Chiurazzi P, Oostra BA (2000) Genetics of mental retardation. *Curr Opin Pediatr* 12: 529-35.
- Chiurazzi P, Pirozzi F (2016) Advances in understanding - genetic basis of intellectual disability. *F1000Res* 5. doi: 10.12688/f1000research.7134.1
- Chong JX, Buckingham KJ, Jhangiani SN, Boehm C, Sobreira N, Smith JD, Harrell TM, McMillin MJ, Wiszniewski W, Gambin T, Coban Akdemir ZH, Doheny K, Scott AF, Avramopoulos D, Chakravarti A, Hoover-Fong J, Mathews D, Witmer PD, Ling H, Hetrick K, Watkins L, Patterson KE, Reinier F, Blue E, Muzny D, Kircher M, Bilguvar K, Lopez-Giraldez F, Sutton VR, Tabor HK, Leal SM, Gunel M, Mane S, Gibbs RA, Boerwinkle E, Hamosh A, Shendure J, Lupski JR, Lifton RP, Valle D, Nickerson DA, Centers for Mendelian G, Bamshad MJ (2015) The Genetic Basis of Mendelian

- Phenotypes: Discoveries, Challenges, and Opportunities. *Am J Hum Genet* 97: 199-215. doi: 10.1016/j.ajhg.2015.06.009
- Coe BP, Witherspoon K, Rosenfeld JA, van Bon BW, Vulto-van Silfhout AT, Bosco P, Friend KL, Baker C, Buono S, Vissers LE, Schuurs-Hoeijmakers JH, Hoischen A, Pfundt R, Krumm N, Carvill GL, Li D, Amaral D, Brown N, Lockhart PJ, Scheffer IE, Alberti A, Shaw M, Pettinato R, Tervo R, de Leeuw N, Reijnders MR, Torchia BS, Peeters H, O'Roak BJ, Fichera M, Hehir-Kwa JY, Shendure J, Mefford HC, Haan E, Gecz J, de Vries BB, Romano C, Eichler EE (2014) Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat Genet* 46: 1063-71. doi: 10.1038/ng.3092
- Colak D, Zaninovic N, Cohen MS, Rosenwaks Z, Yang WY, Gerhardt J, Disney MD, Jaffrey SR (2014) Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* 343: 1002-5. doi: 10.1126/science.1245831
- Conaway RC, Brower CS, Conaway JW (2002) Emerging roles of ubiquitin in transcription regulation. *Science* 296: 1254-8. doi: 10.1126/science.1067466
- Cooper GM, Coe BP, Girirajan S, Rosenfeld JA, Vu TH, Baker C, Williams C, Stalker H, Hamid R, Hannig V, Abdel-Hamid H, Bader P, McCracken E, Niyazov D, Leppig K, Thiese H, Hummel M, Alexander N, Gorski J, Kussmann J, Shashi V, Johnson K, Rehder C, Ballif BC, Shaffer LG, Eichler EE (2011) A copy number variation morbidity map of developmental delay. *Nat Genet* 43: 838-46. doi: 10.1038/ng.909
- Crawford DC, Acuna JM, Sherman SL (2001) FMR1 and the fragile X syndrome: human genome epidemiology review. *Genet Med* 3: 359-71. doi: 10.1097/00125817-200109000-00006
- Crawford J, Lower KM, Hennekam RC, Van Esch H, Megarbane A, Lynch SA, Turner G, Gecz J (2006) Mutation screening in Borjeson-Forssman-Lehmann syndrome: identification of a novel de novo PHF6 mutation in a female patient. *J Med Genet* 43: 238-43. doi: 10.1136/jmg.2005.033084
- Czeschik JC, Bauer P, Buiting K, Dufke C, Guillen-Navarro E, Johnson DS, Koehler U, Lopez-Gonzalez V, Ludecke HJ, Male A, Morrogh D, Riess A, Tzschach A, Wieczorek D, Kuechler A (2013) X-linked intellectual disability type Nascimento is a clinically distinct, probably underdiagnosed entity. *Orphanet J Rare Dis* 8: 146. doi: 10.1186/1750-1172-8-146
- de Leeuw N, Bulk S, Green A, Jaekle-Santos L, Baker LA, Zinn AR, Kleefstra T, van der Smagt JJ, Vianne Morgante AM, de Vries BB, van Bokhoven H, de Brouwer AP (2010) UBE2A deficiency syndrome: Mild to severe intellectual disability accompanied by seizures, absent speech, urogenital, and skin anomalies in male patients. *Am J Med Genet A* 152A: 3084-90. doi: 10.1002/ajmg.a.33743
- de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, Kroes T, Vulto-van Silfhout AT, Koolen DA, de Vries P, Gilissen C, del Rosario M, Hoischen A, Scheffer H, de Vries BB, Brunner HG, Veltman JA, Vissers LE (2012) Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 367: 1921-9. doi: 10.1056/NEJMoa1206524
- Deciphering Developmental Disorders S (2015) Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 519: 223-8. doi: 10.1038/nature14135
- Deciphering Developmental Disorders S (2017) Prevalence and architecture of de novo mutations in developmental disorders. *Nature* 542: 433-438. doi: 10.1038/nature21062
- DesRoches CL, Patel J, Wang P, Minassian B, Salomons GS, Marshall CR, Mercimek-Mahmutoglu S (2015) Estimated carrier frequency of creatine transporter deficiency in females in the general population using functional characterization

- of novel missense variants in the SLC6A8 gene. *Gene* 565: 187-91. doi: 10.1016/j.gene.2015.04.011
- Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S, Boyer TG (2008) Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. *Mol Cell* 31: 347-59. doi: 10.1016/j.molcel.2008.05.023
- Dornelas Lde F, Duarte NM, Magalhaes Lde C (2015) [Neuropsychomotor developmental delay: conceptual map, term definitions, uses and limitations]. *Rev Paul Pediatr* 33: 88-103. doi: 10.1016/j.rpped.2014.04.009
- Dover J, Schneider J, Tawiah-Boateng MA, Wood A, Dean K, Johnston M, Shilatifard A (2002) Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J Biol Chem* 277: 28368-71. doi: 10.1074/jbc.C200348200
- Durkin MS, Hasan ZM, Hasan KZ (1998) Prevalence and correlates of mental retardation among children in Karachi, Pakistan. *Am J Epidemiol* 147: 281-8.
- Ellison JW, Rosenfeld JA, Shaffer LG (2013) Genetic basis of intellectual disability. *Annu Rev Med* 64: 441-50. doi: 10.1146/annurev-med-042711-140053
- Faheem M, Naseer MI, Rasool M, Chaudhary AG, Kumosani TA, Ilyas AM, Pushparaj P, Ahmed F, Algahtani HA, Al-Qahtani MH, Saleh Jamal H (2015) Molecular genetics of human primary microcephaly: an overview. *BMC Med Genomics* 8 Suppl 1: S4. doi: 10.1186/1755-8794-8-S1-S4
- Fernell E (1998) Aetiological factors and prevalence of severe mental retardation in children in a Swedish municipality: the possible role of consanguinity. *Dev Med Child Neurol* 40: 608-11.
- Filges I, Shimojima K, Okamoto N, Rothlisberger B, Weber P, Huber AR, Nishizawa T, Datta AN, Miny P, Yamamoto T (2011) Reduced expression by SETBP1 haploinsufficiency causes developmental and expressive language delay indicating a phenotype distinct from Schinzel-Giedion syndrome. *J Med Genet* 48: 117-22. doi: 10.1136/jmg.2010.084582
- Flint J, Wilkie AO, Buckle VJ, Winter RM, Holland AJ, McDermid HE (1995) The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nat Genet* 9: 132-40. doi: 10.1038/ng0295-132
- Franzoni E, Booker SA, Parthasarathy S, Rehfeld F, Grosser S, Srivatsa S, Fuchs HR, Tarabykin V, Vida I, Wulczyn FG (2015) miR-128 regulates neuronal migration, outgrowth and intrinsic excitability via the intellectual disability gene Phf6. *Elife* 4. doi: 10.7554/eLife.04263
- Gecz J, Turner G, Nelson J, Partington M (2006) The Borjeson-Forssman-Lehman syndrome (BFLS, MIM #301900). *Eur J Hum Genet* 14: 1233-7. doi: 10.1038/sj.ejhg.5201639
- Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A, Leach R, Klein R, Tearle R, Bo T, Pfundt R, Yntema HG, de Vries BB, Kleefstra T, Brunner HG, Vissers LE, Veltman JA (2014) Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511: 344-7. doi: 10.1038/nature13394
- Giugliano T, Santoro C, Torella A, Del Vecchio Blanco F, Bernardo P, Nigro V, Piluso G (2018) UBE2A deficiency in two siblings: A novel splicing variant inherited from a maternal germline mosaicism. *Am J Med Genet A* 176: 722-726. doi: 10.1002/ajmg.a.38589
- Gordon CT, Chopra M, Oufadem M, Alibeu O, Bras M, Boddaert N, Bole-Feysot C, Nitschke P, Abadie V, Lyonnet S, Amiel J (2018) MED13L loss-of-function variants in two patients with syndromic Pierre Robin sequence. *Am J Med Genet A* 176: 181-186. doi: 10.1002/ajmg.a.38536

- Graham JM, Jr., Schwartz CE (2013) MED12 related disorders. *Am J Med Genet A* 161A: 2734-40. doi: 10.1002/ajmg.a.36183
- Gregor P, Nash SR, Caron MG, Seldin MF, Warren ST (1995) Assignment of the creatine transporter gene (SLC6A8) to human chromosome Xq28 telomeric to G6PD. *Genomics* 25: 332-3.
- Haddad DM, Vilain S, Vos M, Esposito G, Matta S, Kalscheuer VM, Craessaerts K, Leyssen M, Nascimento RM, Vianna-Morgante AM, De Strooper B, Van Esch H, Morais VA, Verstreken P (2013) Mutations in the intellectual disability gene Ube2a cause neuronal dysfunction and impair parkin-dependent mitophagy. *Mol Cell* 50: 831-43. doi: 10.1016/j.molcel.2013.04.012
- Hamamy HA, Masri AT, Al-Hadidy AM, Ajlouni KM (2007) Consanguinity and genetic disorders. Profile from Jordan. *Saudi Med J* 28: 1015-7.
- Hamdan FF, Srouf M, Capo-Chichi JM, Daoud H, Nassif C, Patry L, Massicotte C, Ambalavanan A, Spiegelman D, Diallo O, Henrion E, Dionne-Laporte A, Foucherat A, Pshezhetsky AV, Venkateswaran S, Rouleau GA, Michaud JL (2014) De novo mutations in moderate or severe intellectual disability. *PLoS Genet* 10: e1004772. doi: 10.1371/journal.pgen.1004772
- Han JY, Lee IG, Jang W, Kim M, Kim Y, Jang JH, Park J (2017) Diagnostic exome sequencing identifies a heterozygous MBD5 frameshift mutation in a family with intellectual disability and epilepsy. *Eur J Med Genet* 60: 559-564. doi: 10.1016/j.ejmg.2017.08.003
- Hassan MJ, Khurshid M, Azeem Z, John P, Ali G, Chishti MS, Ahmad W (2007) Previously described sequence variant in CDK5RAP2 gene in a Pakistani family with autosomal recessive primary microcephaly. *BMC Med Genet* 8: 58. doi: 10.1186/1471-2350-8-58
- Heather JM, Chain B (2016) The sequence of sequencers: The history of sequencing DNA. *Genomics* 107: 1-8. doi: 10.1016/j.ygeno.2015.11.003
- Hehir-Kwa JY, Rodriguez-Santiago B, Vissers LE, de Leeuw N, Pfundt R, Buitelaar JK, Perez-Jurado LA, Veltman JA (2011) De novo copy number variants associated with intellectual disability have a paternal origin and age bias. *J Med Genet* 48: 776-8. doi: 10.1136/jmedgenet-2011-100147
- Herenger Y, Stoetzel C, Schaefer E, Scheidecker S, Maniere MC, Pelletier V, Alembik Y, Christmann D, Clavert JM, Terzic J, Fischbach M, De Saint Martin A, Dollfus H (2015) Long term follow up of two independent patients with Schinzel-Giedion carrying SETBP1 mutations. *Eur J Med Genet* 58: 479-87. doi: 10.1016/j.ejmg.2015.07.004
- Hildebrand MS, Tankard R, Gazina EV, Damiano JA, Lawrence KM, Dahl HH, Regan BM, Shearer AE, Smith RJ, Marini C, Guerrini R, Labate A, Gambardella A, Tinuper P, Lichetta L, Baldassari S, Bisulli F, Pippucci T, Scheffer IE, Reid CA, Petrou S, Bahlo M, Berkovic SF (2015) PRIMA1 mutation: a new cause of nocturnal frontal lobe epilepsy. *Ann Clin Transl Neurol* 2: 821-30. doi: 10.1002/acn3.224
- Hochstenbach R, van Binsbergen E, Engelen J, Nieuwint A, Polstra A, Poddighe P, Ruivenkamp C, Sikkema-Raddatz B, Smeets D, Poot M (2009) Array analysis and karyotyping: workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands. *Eur J Med Genet* 52: 161-9. doi: 10.1016/j.ejmg.2009.03.015
- Hodge JC, Mitchell E, Pillalamarri V, Toler TL, Bartel F, Kearney HM, Zou YS, Tan WH, Hanscom C, Kirmani S, Hanson RR, Skinner SA, Rogers RC, Everman DB, Boyd E, Tapp C, Mullegama SV, Keelean-Fuller D, Powell CM, Elsea SH, Morton CC, Gusella JF, DuPont B, Chaubey A, Lin AE, Talkowski ME (2014) Disruption of MBD5 contributes to a spectrum of psychopathology and neurodevelopmental abnormalities. *Mol Psychiatry* 19: 368-79. doi: 10.1038/mp.2013.42

- Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, Steehouwer M, de Vries P, de Reuver R, Wieskamp N, Mortier G, Devriendt K, Amorim MZ, Revencu N, Kidd A, Barbosa M, Turner A, Smith J, Oley C, Henderson A, Hayes IM, Thompson EM, Brunner HG, de Vries BB, Veltman JA (2010) De novo mutations of SETBP1 cause Schinzel-Giedion syndrome. *Nat Genet* 42: 483-5. doi: 10.1038/ng.581
- Hoischen A, van Bon BW, Rodriguez-Santiago B, Gilissen C, Vissers LE, de Vries P, Janssen I, van Lier B, Hastings R, Smithson SF, Newbury-Ecob R, Kjaergaard S, Goodship J, McGowan R, Bartholdi D, Rauch A, Peippo M, Cobben JM, Wieczorek D, Gillessen-Kaesbach G, Veltman JA, Brunner HG, de Vries BB (2011) De novo nonsense mutations in ASXL1 cause Bohring-Opitz syndrome. *Nat Genet* 43: 729-31. doi: 10.1038/ng.868
- Honda S, Orii KO, Kobayashi J, Hayashi S, Imamura A, Imoto I, Nakagawa E, Goto Y, Inazawa J (2010) Novel deletion at Xq24 including the UBE2A gene in a patient with X-linked mental retardation. *J Hum Genet* 55: 244-7. doi: 10.1038/jhg.2010.14
- Howard RA, Sharma P, Hajjar C, Caldwell KA, Caldwell GA, du Breuil R, Moore R, Boyd L (2007) Ubiquitin conjugating enzymes participate in polyglutamine protein aggregation. *BMC Cell Biol* 8: 32. doi: 10.1186/1471-2121-8-32
- Hussain MS, Baig SM, Neumann S, Nurnberg G, Farooq M, Ahmad I, Alef T, Hennies HC, Technau M, Altmuller J, Frommolt P, Thiele H, Noegel AA, Nurnberg P (2012) A truncating mutation of CEP135 causes primary microcephaly and disturbed centrosomal function. *Am J Hum Genet* 90: 871-8. doi: 10.1016/j.ajhg.2012.03.016
- Ishizuka K, Kimura H, Yoshimi A, Banno M, Kushima I, Uno Y, Okada T, Mori D, Aleksic B, Ozaki N (2016) Investigation of single-nucleotide variants in MBD5 associated with autism spectrum disorders and schizophrenia phenotypes. *Nagoya J Med Sci* 78: 465-474. doi: 10.18999/nagjms.78.4.465
- Issa L, Kraemer N, Rickert CH, Siffringer M, Ninnemann O, Stoltenburg-Didinger G, Kaindl AM (2013a) CDK5RAP2 expression during murine and human brain development correlates with pathology in primary autosomal recessive microcephaly. *Cereb Cortex* 23: 2245-60. doi: 10.1093/cercor/bhs212
- Issa L, Mueller K, Seufert K, Kraemer N, Rosenkotter H, Ninnemann O, Buob M, Kaindl AM, Morris-Rosendahl DJ (2013b) Clinical and cellular features in patients with primary autosomal recessive microcephaly and a novel CDK5RAP2 mutation. *Orphanet J Rare Dis* 8: 59. doi: 10.1186/1750-1172-8-59
- Jacquemont S, Coe BP, Hersch M, Duyzend MH, Krumm N, Bergmann S, Beckmann JS, Rosenfeld JA, Eichler EE (2014) A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. *Am J Hum Genet* 94: 415-25. doi: 10.1016/j.ajhg.2014.02.001
- Jaillard S, Dubourg C, Gerard-Blanluet M, Delahaye A, Pasquier L, Dupont C, Henry C, Tabet AC, Lucas J, Aboura A, David V, Benzacken B, Odent S, Pipiras E (2009) 2q23.1 microdeletion identified by array comparative genomic hybridisation: an emerging phenotype with Angelman-like features? *J Med Genet* 46: 847-55. doi: 10.1136/jmg.2008.058156
- Josh NA (2011) Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33). <http://www.citeulike.org/user/mvermaat/article/13260426>.
- Jouan L, Ouled Amar Bencheikh B, Daoud H, Dionne-Laporte A, Dobrzeniecka S, Spiegelman D, Rochefort D, Hince P, Szuto A, Lassonde M, Barbelanne M, Tsang WY, Dion PA, Theoret H, Rouleau GA (2016) Exome sequencing identifies recessive CDK5RAP2 variants in patients with isolated agenesis of corpus callosum. *Eur J Hum Genet* 24: 607-10. doi: 10.1038/ejhg.2015.156

- Kalia SS, Adelman K, Bale SJ, Chung WK, Eng C, Evans JP, Herman GE, Hufnagel SB, Klein TE, Korf BR, McKelvey KD, Ormond KE, Richards CS, Vlangos CN, Watson M, Martin CL, Miller DT (2017) Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med* 19: 249-255. doi: 10.1038/gim.2016.190
- Kleefstra T, Kramer JM, Neveling K, Willemsen MH, Koemans TS, Vissers LE, Wissink-Lindhout W, Fenckova M, van den Akker WM, Kasri NN, Nillesen WM, Prescott T, Clark RD, Devriendt K, van Reeuwijk J, de Brouwer AP, Gilissen C, Zhou H, Brunner HG, Veltman JA, Schenck A, van Bokhoven H (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am J Hum Genet* 91: 73-82. doi: 10.1016/j.ajhg.2012.05.003
- Ko JM, Lim BC, Kim KJ, Hwang YS, Ryu HW, Lee JH, Kim JS, Chae JH (2013) Distinct neurological features in a patient with Schinzel-Giedion syndrome caused by a recurrent SETBP1 mutation. *Childs Nerv Syst* 29: 525-9. doi: 10.1007/s00381-013-2047-2
- Kochinke K, Zweier C, Nijhof B, Fenckova M, Cizek P, Honti F, Keerthikumar S, Oortveld MA, Kleefstra T, Kramer JM, Webber C, Huynen MA, Schenck A (2016) Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet* 98: 149-64. doi: 10.1016/j.ajhg.2015.11.024
- Kraemer N, Issa L, Hauck SC, Mani S, Ninnemann O, Kaindl AM (2011) What's the hype about CDK5RAP2? *Cell Mol Life Sci* 68: 1719-36. doi: 10.1007/s00018-011-0635-4
- Kuss AW, Garshasbi M, Kahrizi K, Tzschach A, Behjati F, Darvish H, Abbasi-Moheb L, Puettmann L, Zecha A, Weissmann R, Hu H, Mohseni M, Abedini SS, Rajab A, Hertzberg C, Wiczorek D, Ullmann R, Ghasemi-Firouzabadi S, Banihashemi S, Arzhanghi S, Hadavi V, Bahrami-Monajemi G, Kasiri M, Falah M, Nikuei P, Dehghan A, Sobhani M, Jamali P, Ropers HH, Najmabadi H (2011) Autosomal recessive mental retardation: homozygosity mapping identifies 27 single linkage intervals, at least 14 novel loci and several mutation hotspots. *Hum Genet* 129: 141-8. doi: 10.1007/s00439-010-0907-3
- Laget S, Joulie M, Le Masson F, Sasai N, Christians E, Pradhan S, Roberts RJ, Defossez PA (2010) The human proteins MBD5 and MBD6 associate with heterochromatin but they do not bind methylated DNA. *PLoS One* 5: e11982. doi: 10.1371/journal.pone.0011982
- Lancaster MA, Renner M, Martin CA, Wenzel D, Bicknell LS, Hurles ME, Homfray T, Penninger JM, Jackson AP, Knoblich JA (2013) Cerebral organoids model human brain development and microcephaly. *Nature* 501: 373-9. doi: 10.1038/nature12517
- Larti F, Kahrizi K, Musante L, Hu H, Papari E, Fattahi Z, Bazazzadegan N, Liu Z, Banan M, Garshasbi M, Wienker TF, Ropers HH, Galjart N, Najmabadi H (2015) A defect in the CLIP1 gene (CLIP-170) can cause autosomal recessive intellectual disability. *Eur J Hum Genet* 23: 331-6. doi: 10.1038/ejhg.2014.13
- Lecointre C, Pichon O, Hamel A, Heloury Y, Michel-Calemard L, Morel Y, David A, Le Caignec C (2009) Familial acampomelic form of campomelic dysplasia caused by a 960 kb deletion upstream of SOX9. *Am J Med Genet A* 149A: 1183-9. doi: 10.1002/ajmg.a.32830
- Leonard H, Wen X (2002) The epidemiology of mental retardation: challenges and opportunities in the new millennium. *Ment Retard Dev Disabil Res Rev* 8: 117-34. doi: 10.1002/mrdd.10031

- Lesca G, Moizard MP, Bussy G, Boggio D, Hu H, Haas SA, Ropers HH, Kalscheuer VM, Des Portes V, Labalme A, Sanlaville D, Edery P, Raynaud M, Lespinasse J (2013) Clinical and neurocognitive characterization of a family with a novel MED12 gene frameshift mutation. *Am J Med Genet A* 161A: 3063-71. doi: 10.1002/ajmg.a.36162
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589-95. doi: 10.1093/bioinformatics/btp698
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078-9. doi: 10.1093/bioinformatics/btp352
- Li MH, Arndt K, Das S, Weiss EM, Wu Y, Gwal K, Shekdar KV, Zackai EH (2015) Compound heterozygote CDK5RAP2 mutations in a Guatemalan/Honduran child with autosomal recessive primary microcephaly, failure to thrive and speech delay. *Am J Med Genet A* 167: 1414-7. doi: 10.1002/ajmg.a.36975
- Li Q, Wang K (2017) InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet* 100: 267-280. doi: 10.1016/j.ajhg.2017.01.004
- Liu Z, Li F, Ruan K, Zhang J, Mei Y, Wu J, Shi Y (2014) Structural and functional insights into the human Borjeson-Forssman-Lehmann syndrome-associated protein PHF6. *J Biol Chem* 289: 10069-83. doi: 10.1074/jbc.M113.535351
- Lizarraga SB, Margossian SP, Harris MH, Campagna DR, Han AP, Blevins S, Mudbhary R, Barker JE, Walsh CA, Fleming MD (2010) Cdk5rap2 regulates centrosome function and chromosome segregation in neuronal progenitors. *Development* 137: 1907-17. doi: 10.1242/dev.040410
- Lopez-Gonzalez V, Domingo-Jimenez MR, Burglen L, Ballesta-Martinez MJ, Whalen S, Pinero-Fernandez JA, Guillen-Navarro E (2015) [Schinzel-Giedion syndrome: a new mutation in SETBP1]. *An Pediatr (Barc)* 82: e12-6. doi: 10.1016/j.anpedi.2014.06.017
- Lower KM, Turner G, Kerr BA, Mathews KD, Shaw MA, Gedeon AK, Schelley S, Hoyme HE, White SM, Delatycki MB, Lampe AK, Clayton-Smith J, Stewart H, van Ravenswaay CM, de Vries BB, Cox B, Grompe M, Ross S, Thomas P, Mulley JC, Gecz J (2002) Mutations in PHF6 are associated with Borjeson-Forssman-Lehmann syndrome. *Nat Genet* 32: 661-5. doi: 10.1038/ng1040
- Lubs HA, Stevenson RE, Schwartz CE (2012) Fragile X and X-linked intellectual disability: four decades of discovery. *Am J Hum Genet* 90: 579-90. doi: 10.1016/j.ajhg.2012.02.018
- Lujan JE, Carlin ME, Lubs HA (1984) A form of X-linked mental retardation with marfanoid habitus. *Am J Med Genet* 17: 311-22. doi: 10.1002/ajmg.1320170124
- Mak CS, Waldvogel HJ, Dodd JR, Gilbert RT, Lowe MT, Birch NP, Faull RL, Christie DL (2009) Immunohistochemical localisation of the creatine transporter in the rat brain. *Neuroscience* 163: 571-85. doi: 10.1016/j.neuroscience.2009.06.065
- Makinen N, Mehine M, Tolvanen J, Kaasinen E, Li Y, Lehtonen HJ, Gentile M, Yan J, Enge M, Taipale M, Aavikko M, Katainen R, Virolainen E, Bohling T, Koski TA, Launonen V, Sjoberg J, Taipale J, Vahteristo P, Aaltonen LA (2011) MED12, the mediator complex subunit 12 gene, is mutated at high frequency in uterine leiomyomas. *Science* 334: 252-5. doi: 10.1126/science.1208930
- Makinen N, Vahteristo P, Kampjarvi K, Arola J, Butzow R, Aaltonen LA (2013) MED12 exon 2 mutations in histopathological uterine leiomyoma variants. *Eur J Hum Genet* 21: 1300-3. doi: 10.1038/ejhg.2013.33
- Makishima H (2017) Somatic SETBP1 mutations in myeloid neoplasms. *Int J Hematol* 105: 732-742. doi: 10.1007/s12185-017-2241-1

- Makishima H, Yoshida K, Nguyen N, Przychodzen B, Sanada M, Okuno Y, Ng KP, Gudmundsson KO, Vishwakarma BA, Jerez A, Gomez-Segui I, Takahashi M, Shiraishi Y, Nagata Y, Guinta K, Mori H, Sekeres MA, Chiba K, Tanaka H, Muramatsu H, Sakaguchi H, Paquette RL, McDevitt MA, Kojima S, Sauntharajah Y, Miyano S, Shih LY, Du Y, Ogawa S, Maciejewski JP (2013) Somatic SETBP1 mutations in myeloid malignancies. *Nat Genet* 45: 942-6. doi: 10.1038/ng.2696
- Mandel JL, Chelly J (2004) Monogenic X-linked mental retardation: is it as frequent as currently estimated? The paradox of the ARX (Aristaless X) mutations. *Eur J Hum Genet* 12: 689-93. doi: 10.1038/sj.ejhg.5201247
- Marco EJ, Aitken AB, Nair VP, da Gente G, Gerdes MR, Bologlu L, Thomas S, Sherr EH (2018) Burden of de novo mutations and inherited rare single nucleotide variants in children with sensory processing dysfunction. *BMC Med Genomics* 11: 50. doi: 10.1186/s12920-018-0362-x
- Marseglia G, Scordo MR, Pescucci C, Nannetti G, Biagini E, Scandurra V, Gerundino F, Magi A, Benelli M, Torricelli F (2012) 372 kb microdeletion in 18q12.3 causing SETBP1 haploinsufficiency associated with mild mental retardation and expressive speech impairment. *Eur J Med Genet* 55: 216-21. doi: 10.1016/j.ejmg.2012.01.005
- Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S (2011) Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res Dev Disabil* 32: 419-36. doi: 10.1016/j.ridd.2010.12.018
- Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, Church DM, Crolla JA, Eichler EE, Epstein CJ, Faucett WA, Feuk L, Friedman JM, Hamosh A, Jackson L, Kaminsky EB, Kok K, Krantz ID, Kuhn RM, Lee C, Ostell JM, Rosenberg C, Scherer SW, Spinner NB, Stavropoulos DJ, Tepperberg JH, Thorland EC, Vermeesch JR, Waggoner DJ, Watson MS, Martin CL, Ledbetter DH (2010) Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86: 749-64. doi: 10.1016/j.ajhg.2010.04.006
- Mir A, Sritharan K, Mittal K, Vasli N, Araujo C, Jamil T, Rafiq MA, Anwar Z, Mikhailov A, Rauf S, Mahmood H, Shakoor A, Ali S, So J, Naeem F, Ayub M, Vincent JB (2014) Truncation of the E3 ubiquitin ligase component FBXO31 causes non-syndromic autosomal recessive intellectual disability in a Pakistani family. *Hum Genet* 133: 975-84. doi: 10.1007/s00439-014-1438-0
- Mir R, Javid J, Al Balawi IA, Alkharsah KR, Hadi MA, Rahman MA, Hamoud E, Al Alawi Y, Al Zahrani ABM, Abu-Duhier FM (2018) A Germline Mutation in the BRCA1 3'UTR Variant Predicts Susceptibility to Breast Cancer in a Saudi Arabian Population. *Asian Pac J Cancer Prev* 19: 859-866. doi: 10.22034/APJCP.2018.19.3.859
- Mullegama SV, Elsea SH (2016) Clinical and Molecular Aspects of MBD5-Associated Neurodevelopmental Disorder (MAND). *Eur J Hum Genet* 24: 1376. doi: 10.1038/ejhg.2016.77
- Muncke N, Jung C, Rudiger H, Ulmer H, Roeth R, Hubert A, Goldmuntz E, Driscoll D, Goodship J, Schon K, Rappold G (2003) Missense mutations and gene interruption in PROSIT240, a novel TRAP240-like gene, in patients with congenital heart defect (transposition of the great arteries). *Circulation* 108: 2843-50. doi: 10.1161/01.CIR.0000103684.77636.CD
- Musante L, Ropers HH (2014) Genetics of recessive cognitive disorders. *Trends Genet* 30: 32-9. doi: 10.1016/j.tig.2013.09.008
- Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S, Jamali P, Zecha A, Mohseni M, Puttmann L, Vahid LN, Jensen C, Moheb LA, Bienek M, Larti F, Mueller I, Weissmann R, Darvish H, Wrogemann K, Hadavi V, Lipkowitz B, Esmaeeli-Nieh S, Wieczorek D, Kariminejad R, Firouzabadi SG,

- Cohen M, Fattahi Z, Rost I, Mojahedi F, Hertzberg C, Dehghan A, Rajab A, Banavandi MJ, Hoffer J, Falah M, Musante L, Kalscheuer V, Ullmann R, Kuss AW, Tzschach A, Kahrizi K, Ropers HH (2011) Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 478: 57-63. doi: 10.1038/nature10423
- Najmabadi H, Motazacker MM, Garshasbi M, Kahrizi K, Tzschach A, Chen W, Behjati F, Hadavi V, Nieh SE, Abedini SS, Vazifehmand R, Firouzabadi SG, Jamali P, Falah M, Seifati SM, Gruters A, Lenzner S, Jensen LR, Ruschendorf F, Kuss AW, Ropers HH (2007) Homozygosity mapping in consanguineous families reveals extreme heterogeneity of non-syndromic autosomal recessive mental retardation and identifies 8 novel gene loci. *Hum Genet* 121: 43-8. doi: 10.1007/s00439-006-0292-0
- Nascimento RM, Otto PA, de Brouwer AP, Vianna-Morgante AM (2006) UBE2A, which encodes a ubiquitin-conjugating enzyme, is mutated in a novel X-linked mental retardation syndrome. *Am J Hum Genet* 79: 549-55. doi: 10.1086/507047
- Neri G, Schwartz CE, Lubs HA, Stevenson RE (2018) X-linked intellectual disability update 2017. *Am J Med Genet A* 176: 1375-1388. doi: 10.1002/ajmg.a.38710
- Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, Beck AE, Tabor HK, Cooper GM, Mefford HC, Lee C, Turner EH, Smith JD, Rieder MJ, Yoshiura K, Matsumoto N, Ohta T, Niikawa N, Nickerson DA, Bamshad MJ, Shendure J (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 42: 790-3. doi: 10.1038/ng.646
- Niranjan TS, Skinner C, May M, Turner T, Rose R, Stevenson R, Schwartz CE, Wang T (2015) Affected kindred analysis of human X chromosome exomes to identify novel X-linked intellectual disability genes. *PLoS One* 10: e0116454. doi: 10.1371/journal.pone.0116454
- Noh GJ, Graham JM, Jr. (2012) 2q23.1 microdeletion of the MBD5 gene in a female with seizures, developmental delay and distinct dysmorphic features. *Eur J Med Genet* 55: 59-62. doi: 10.1016/j.ejmg.2011.10.001
- O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD, Turner EH, Stanaway IB, Vernot B, Malig M, Baker C, Reilly B, Akey JM, Borenstein E, Rieder MJ, Nickerson DA, Bernier R, Shendure J, Eichler EE (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* 485: 246-50. doi: 10.1038/nature10989
- Oladnabi M, Musante L, Larti F, Hu H, Abedini SS, Wienker T, Ropers HH, Kahrizi K, Najmabadi H (2015) New evidence for the role of calpain 10 in autosomal recessive intellectual disability: identification of two novel nonsense variants by exome sequencing in Iranian families. *Arch Iran Med* 18: 179-84. doi: 0151803/AIM.008
- Opitz JM, Kaveggia EG (1974) Studies of malformation syndromes of man 33: the FG syndrome. An X-linked recessive syndrome of multiple congenital anomalies and mental retardation. *Z Kinderheilkd* 117: 1-18.
- Pagnamenta AT, Howard MF, Knight SJ, Keays DA, Quaghebeur G, Taylor JC, Kini U (2016) Activation of an exonic splice-donor site in exon 30 of CDK5RAP2 in a patient with severe microcephaly and pigmentary abnormalities. *Clin Case Rep* 4: 952-956. doi: 10.1002/ccr3.663
- Pagnamenta AT, Murray JE, Yoon G, Sadighi Akha E, Harrison V, Bicknell LS, Ajilogba K, Stewart H, Kini U, Taylor JC, Keays DA, Jackson AP, Knight SJ (2012) A novel nonsense CDK5RAP2 mutation in a Somali child with primary microcephaly and sensorineural hearing loss. *Am J Med Genet A* 158A: 2577-82. doi: 10.1002/ajmg.a.35558

- Pajusalu S, Reimand T, Ounap K (2015) Novel homozygous mutation in KPTN gene causing a familial intellectual disability-macrocephaly syndrome. *Am J Med Genet A* 167A: 1913-5. doi: 10.1002/ajmg.a.37105
- Petroski MD (2008) The ubiquitin system, disease, and drug discovery. *BMC Biochem* 9 Suppl 1: S7. doi: 10.1186/1471-2091-9-S1-S7
- Philibert RA, Madan A (2007) Role of MED12 in transcription and human behavior. *Pharmacogenomics* 8: 909-16. doi: 10.2217/14622416.8.8.909
- Pinkel D, Gray JW, Trask B, van den Engh G, Fuscoe J, van Dekken H (1986a) Cytogenetic analysis by in situ hybridization with fluorescently labeled nucleic acid probes. *Cold Spring Harb Symp Quant Biol* 51 Pt 1: 151-7.
- Pinkel D, Straume T, Gray JW (1986b) Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci U S A* 83: 2934-8.
- Prontera P, Ottaviani V, Rogaia D, Isidori I, Mencarelli A, Malerba N, Cocciadiferro D, Rolph P, Stangoni G, Vulto-van Silfhout A, Merla G (2016) A novel MED12 mutation: Evidence for a fourth phenotype. *Am J Med Genet A* 170: 2377-82. doi: 10.1002/ajmg.a.37805
- Rafiq MA, Leblond CS, Saqib MA, Vincent AK, Ambalavanan A, Khan FS, Ayaz M, Shaheen N, Spiegelman D, Ali G, Amin-ud-Din M, Laurent S, Mahmood H, Christian M, Ali N, Fennell A, Nanjiani Z, Egger G, Caron C, Waqas A, Ayub M, Rasheed S, Forgeot d'Arc B, Johnson A, So J, Brohi MQ, Motttron L, Ansar M, Vincent JB, Xiong L (2015) Novel VPS13B Mutations in Three Large Pakistani Cohen Syndrome Families Suggests a Baloch Variant with Autistic-Like Features. *BMC Med Genet* 16: 41. doi: 10.1186/s12881-015-0183-0
- Rafiullah R, Aslamkhan M, Paramasivam N, Thiel C, Mustafa G, Wiemann S, Schlesner M, Wade RC, Rappold GA, Berkel S (2016) Homozygous missense mutation in the LMAN2L gene segregates with intellectual disability in a large consanguineous Pakistani family. *J Med Genet* 53: 138-44. doi: 10.1136/jmedgenet-2015-103179
- Rajab A, Schuelke M, Gill E, Zwirner A, Seifert F, Morales Gonzalez S, Knierim E (2015) Recessive DEAF1 mutation associates with autism, intellectual disability, basal ganglia dysfunction and epilepsy. *J Med Genet* 52: 607-11. doi: 10.1136/jmedgenet-2015-103083
- Rauch A, Hoyer J, Guth S, Zweier C, Kraus C, Becker C, Zenker M, Huffmeier U, Thiel C, Ruschendorf F, Nurnberg P, Reis A, Trautmann U (2006) Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am J Med Genet A* 140: 2063-74. doi: 10.1002/ajmg.a.31416
- Rauch A, Wieczorek D, Graf E, Wieland T, Ende S, Schwarzmayer T, Albrecht B, Bartholdi D, Beygo J, Di Donato N, Dufke A, Cremer K, Hempel M, Horn D, Hoyer J, Joset P, Ropke A, Moog U, Riess A, Thiel CT, Tzschach A, Wiesener A, Wohlleber E, Zweier C, Ekici AB, Zink AM, Rump A, Meisinger C, Grallert H, Sticht H, Schenck A, Engels H, Rappold G, Schrock E, Wieacker P, Riess O, Meitinger T, Reis A, Strom TM (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380: 1674-82. doi: 10.1016/S0140-6736(12)61480-9
- Raymond FL, Whibley A, Stratton MR, Gecz J (2009) Lessons learnt from large-scale exon re-sequencing of the X chromosome. *Hum Mol Genet* 18: R60-4. doi: 10.1093/hmg/ddp071
- Redin C, Gerard B, Lauer J, Herenger Y, Muller J, Quartier A, Masurel-Paulet A, Willems M, Lesca G, El-Chehadeh S, Le Gras S, Vicaire S, Philipps M, Dumas M, Geoffroy V, Feger C, Haumesser N, Alembik Y, Barth M, Bonneau D, Colin E, Dollfus H, Doray B, Delrue MA, Drouin-Garraud V, Flori E, Fradin M, Francannet C, Goldenberg A, Lumbroso S, Mathieu-Dramard M, Martin-Coignard D, Lacombe D, Morin G, Polge A, Sukno S, Thauvin-Robinet C, Thevenon J, Doco-Fenzy M,

- Genevieve D, Sarda P, Edery P, Isidor B, Jost B, Olivier-Faivre L, Mandel JL, Piton A (2014) Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J Med Genet* 51: 724-36. doi: 10.1136/jmedgenet-2014-102554
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehm HL, Committee ALQA (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17: 405-24. doi: 10.1038/gim.2015.30
- Riley BE, Loughheed JC, Callaway K, Velasquez M, Brecht E, Nguyen L, Shaler T, Walker D, Yang Y, Regnstrom K, Diep L, Zhang Z, Chiou S, Bova M, Artis DR, Yao N, Baker J, Yednock T, Johnston JA (2013) Structure and function of Parkin E3 ubiquitin ligase reveals aspects of RING and HECT ligases. *Nat Commun* 4: 1982. doi: 10.1038/ncomms2982
- Risheg H, Graham JM, Jr., Clark RD, Rogers RC, Opitz JM, Moeschler JB, Peiffer AP, May M, Joseph SM, Jones JR, Stevenson RE, Schwartz CE, Friez MJ (2007) A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. *Nat Genet* 39: 451-3. doi: 10.1038/ng1992
- Roos L, Fang M, Dali C, Jensen H, Christoffersen N, Wu B, Zhang J, Xu R, Harris P, Xu X, Gronskov K, Tumer Z (2014) A homozygous mutation in a consanguineous family consolidates the role of ALDH1A3 in autosomal recessive microphthalmia. *Clin Genet* 86: 276-81. doi: 10.1111/cge.12277
- Ropers HH (2008) Genetics of intellectual disability. *Curr Opin Genet Dev* 18: 241-50. doi: 10.1016/j.gde.2008.07.008
- Ropers HH (2010) Genetics of early onset cognitive impairment. *Annu Rev Genomics Hum Genet* 11: 161-87. doi: 10.1146/annurev-genom-082509-141640
- Ropers HH, Hamel BC (2005) X-linked mental retardation. *Nat Rev Genet* 6: 46-57. doi: 10.1038/nrg1501
- Rosenberg C, Knijnenburg J, Bakker E, Vianna-Morgante AM, Sloos W, Otto PA, Kriek M, Hansson K, Krepisch-Santos AC, Fiegler H, Carter NP, Bijlsma EK, van Haeringen A, Szuhai K, Tanke HJ (2006) Array-CGH detection of micro rearrangements in mentally retarded individuals: clinical significance of imbalances present both in affected children and normal parents. *J Med Genet* 43: 180-6. doi: 10.1136/jmg.2005.032268
- Rosenberg EH, Almeida LS, Kleefstra T, deGrauw RS, Yntema HG, Bahi N, Moraine C, Ropers HH, Fryns JP, deGrauw TJ, Jakobs C, Salomons GS (2004) High prevalence of SLC6A8 deficiency in X-linked mental retardation. *Am J Hum Genet* 75: 97-105. doi: 10.1086/422102
- Rosenfeld JA, Fox JE, Descartes M, Brewer F, Stroud T, Gorski JL, Upton SJ, Moeschler JB, Monteleone B, Neill NJ, Lamb AN, Ballif BC, Shaffer LG, Ravnan JB (2015) Clinical features associated with copy number variations of the 14q32 imprinted gene cluster. *Am J Med Genet A* 167A: 345-53.
- Ross CA, Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nat Med* 10 Suppl: S10-7. doi: 10.1038/nm1066
- Rump P, Niessen RC, Verbruggen KT, Brouwer OF, de Raad M, Hordijk R (2011) A novel mutation in MED12 causes FG syndrome (Opitz-Kaveggia syndrome). *Clin Genet* 79: 183-8. doi: 10.1111/j.1399-0004.2010.01449.x
- Salomons GS, van Dooren SJ, Verhoeven NM, Cecil KM, Ball WS, Degrauw TJ, Jakobs C (2001) X-linked creatine-transporter gene (SLC6A8) defect: a new creatine-deficiency syndrome. *Am J Hum Genet* 68: 1497-500. doi: 10.1086/320595

- Sambrook J, Fritsch EF, Maniatis T (1987) *Molecular cloning : a laboratory manual*, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Sanger F, Coulson AR (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol* 94: 441-8.
- Schinzel A, Giedion A (1978) A syndrome of severe midface retraction, multiple skull anomalies, clubfeet, and cardiac and renal malformations in sibs. *Am J Med Genet* 1: 361-75. doi: 10.1002/ajmg.1320010402
- Schuurs-Hoeijmakers JH, Vulto-van Silfhout AT, Vissers LE, van de V, II, van Bon BW, de Ligt J, Gilissen C, Hehir-Kwa JY, Neveling K, del Rosario M, Hira G, Reitano S, Vitello A, Failla P, Greco D, Fichera M, Galesi O, Kleefstra T, Grealis MT, Ockeloen CW, Willemsen MH, Bongers EM, Janssen IM, Pfundt R, Veltman JA, Romano C, Willemsen MA, van Bokhoven H, Brunner HG, de Vries BB, de Brouwer AP (2013) Identification of pathogenic gene variants in small families with intellectually disabled siblings by exome sequencing. *J Med Genet* 50: 802-11. doi: 10.1136/jmedgenet-2013-101644
- Schwartz CE, Tarpey PS, Lubs HA, Verloes A, May MM, Risheg H, Friez MJ, Futreal PA, Edkins S, Teague J, Briault S, Skinner C, Bauer-Carlin A, Simonsen RJ, Joseph SM, Jones JR, Gecz J, Stratton MR, Raymond FL, Stevenson RE (2007) The original Lujan syndrome family has a novel missense mutation (p.N1007S) in the MED12 gene. *J Med Genet* 44: 472-7. doi: 10.1136/jmg.2006.048637
- Shaffer LG, Bejjani BA, Torchia B, Kirkpatrick S, Coppinger J, Ballif BC (2007) The identification of microdeletion syndromes and other chromosome abnormalities: cytogenetic methods of the past, new technologies for the future. *Am J Med Genet C Semin Med Genet* 145C: 335-45. doi: 10.1002/ajmg.c.30152
- Shin CH, Chung WS, Hong SK, Ober EA, Verkade H, Field HA, Huisken J, Stainier DY (2008) Multiple roles for Med12 in vertebrate endoderm development. *Dev Biol* 317: 467-79. doi: 10.1016/j.ydbio.2008.02.031
- Smol T, Petit F, Piton A, Keren B, Sanlaville D, Afenjar A, Baker S, Bedoukian EC, Bhoj EJ, Bonneau D, Boudry-Labis E, Bouquillon S, Boute-Benejean O, Caumes R, Chatron N, Colson C, Coubes C, Coutton C, Devillard F, Dieux-Coeslier A, Doco-Fenzy M, Ewans LJ, Faivre L, Fassi E, Field M, Fournier C, Francannet C, Genevieve D, Giurgea I, Goldenberg A, Green AK, Guerrot AM, Heron D, Isidor B, Keena BA, Krock BL, Kuentz P, Lapi E, Le Meur N, Lesca G, Li D, Marey I, Mignot C, Nava C, Nesbitt A, Nicolas G, Roche-Lestienne C, Roscioli T, Satre V, Santani A, Stefanova M, Steinwall Larsen S, Saugier-Verber P, Picker-Minh S, Thuillier C, Verloes A, Vieville G, Wenzel M, Willems M, Whalen S, Zarate YA, Ziegler A, Manouvrier-Hanu S, Kalscheuer VM, Gerard B, Ghoumid J (2018) MED13L-related intellectual disability: involvement of missense variants and delineation of the phenotype. *Neurogenetics* 19: 93-103. doi: 10.1007/s10048-018-0541-0
- Stevenson RE (2000) Splitting and lumping in the nosology of XLMR. *Am J Med Genet* 97: 174-82. doi: 10.1002/1096-8628(200023)97:3<174::AID-AJMG1034>3.0.CO;2-4
- Stockler-Ipsiroglu S, van Karnebeek CD (2014) Cerebral creatine deficiencies: a group of treatable intellectual developmental disorders. *Semin Neurol* 34: 350-6. doi: 10.1055/s-0034-1386772
- Suphapeetiporn K, Srichomthong C, Shotelersuk V (2011) SETBP1 mutations in two Thai patients with Schinzel-Giedion syndrome. *Clin Genet* 79: 391-3. doi: 10.1111/j.1399-0004.2010.01552.x
- Suresh B, Lee J, Kim KS, Ramakrishna S (2016) The Importance of Ubiquitination and Deubiquitination in Cellular Reprogramming. *Stem Cells Int* 2016: 6705927. doi: 10.1155/2016/6705927

- Takeuchi A, Okamoto N, Fujinaga S, Morita H, Shimizu J, Akiyama T, Ninomiya S, Takanashi J, Kubo T (2015) Progressive brain atrophy in Schinzel-Giedion syndrome with a SETBP1 mutation. *Eur J Med Genet* 58: 369-71. doi: 10.1016/j.ejmg.2015.05.006
- Talkowski ME, Mullegama SV, Rosenfeld JA, van Bon BW, Shen Y, Repnikova EA, Gastier-Foster J, Thrush DL, Kathiresan S, Ruderfer DM, Chiang C, Hanscom C, Ernst C, Lindgren AM, Morton CC, An Y, Astbury C, Brueton LA, Lichtenbelt KD, Ades LC, Fichera M, Romano C, Innis JW, Williams CA, Bartholomew D, Van Allen MI, Parikh A, Zhang L, Wu BL, Pyatt RE, Schwartz S, Shaffer LG, de Vries BB, Gusella JF, Elsea SH (2011) Assessment of 2q23.1 microdeletion syndrome implicates MBD5 as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. *Am J Hum Genet* 89: 551-63. doi: 10.1016/j.ajhg.2011.09.011
- Tan CA, Topper S, Ward Mever C, Stein J, Reeder A, Arndt K, Das S (2014) The first case of CDK5RAP2-related primary microcephaly in a non-consanguineous patient identified by next generation sequencing. *Brain Dev* 36: 351-5. doi: 10.1016/j.braindev.2013.05.001
- Tarpey PS, Smith R, Pleasance E, Whibley A, Edkins S, Hardy C, O'Meara S, Latimer C, Dicks E, Menzies A, Stephens P, Blow M, Greenman C, Xue Y, Tyler-Smith C, Thompson D, Gray K, Andrews J, Barthorpe S, Buck G, Cole J, Dunmore R, Jones D, Maddison M, Mironenko T, Turner R, Turrell K, Varian J, West S, Widaa S, Wray P, Teague J, Butler A, Jenkinson A, Jia M, Richardson D, Shepherd R, Wooster R, Tejada MI, Martinez F, Carvill G, Goliath R, de Brouwer AP, van Bokhoven H, Van Esch H, Chelly J, Raynaud M, Ropers HH, Abidi FE, Srivastava AK, Cox J, Luo Y, Mallya U, Moon J, Parnau J, Mohammed S, Tolmie JL, Shoubridge C, Corbett M, Gardner A, Haan E, Rujirabanjerd S, Shaw M, Vandeleur L, Fullston T, Easton DF, Boyle J, Partington M, Hackett A, Field M, Skinner C, Stevenson RE, Bobrow M, Turner G, Schwartz CE, Gecz J, Raymond FL, Futreal PA, Stratton MR (2009) A systematic, large-scale resequencing screen of X-chromosome coding exons in mental retardation. *Nat Genet* 41: 535-43. doi: 10.1038/ng.367
- Thunstrom S, Sodermark L, Ivarsson L, Samuelsson L, Stefanova M (2015) UBE2A deficiency syndrome: a report of two unrelated cases with large Xq24 deletions encompassing UBE2A gene. *Am J Med Genet A* 167A: 204-10. doi: 10.1002/ajmg.a.36800
- Todd MA, Ivanochko D, Picketts DJ (2015) PHF6 Degrees of Separation: The Multifaceted Roles of a Chromatin Adaptor Protein. *Genes (Basel)* 6: 325-52. doi: 10.3390/genes6020325
- Torring PM, Larsen MJ, Brasch-Andersen C, Krogh LN, Kibaek M, Laulund L, Illum N, Dunkhase-Heinl U, Wiesener A, Popp B, Marangi G, Hjortshoj TD, Ek J, Vogel I, Becher N, Roos L, Zollino M, Fagerberg CR (2018) Is MED13L-related intellectual disability a recognizable syndrome? *Eur J Med Genet*. doi: 10.1016/j.ejmg.2018.06.014
- Tribioli C, Droetto S, Bione S, Cesareni G, Torrisi MR, Lotti LV, Lanfranccone L, Toniolo D, Pelicci P (1996) An X chromosome-linked gene encoding a protein with characteristics of a rhoGAP predominantly expressed in hematopoietic cells. *Proc Natl Acad Sci U S A* 93: 695-9.
- Tsurusaki Y, Ohashi I, Enomoto Y, Naruto T, Mitsui J, Aida N, Kurosawa K (2017) A novel UBE2A mutation causes X-linked intellectual disability type Nascimento. *Hum Genome Var* 4: 17019. doi: 10.1038/hgv.2017.19
- Tucker T, Zahir FR, Griffith M, Delaney A, Chai D, Tsang E, Lemyre E, Dobrzeniecka S, Marra M, Eydoux P, Langlois S, Hamdan FF, Michaud JL, Friedman JM (2014)

- Single exon-resolution targeted chromosomal microarray analysis of known and candidate intellectual disability genes. *Eur J Hum Genet* 22: 792-800. doi: 10.1038/ejhg.2013.248
- Turner G, Lower KM, White SM, Delatycki M, Lampe AK, Wright M, Smith JC, Kerr B, Schelley S, Hoyne HE, De Vries BB, Kleefstra T, Grompe M, Cox B, Gecz J, Partington M (2004) The clinical picture of the Borjeson-Forssman-Lehmann syndrome in males and heterozygous females with PHF6 mutations. *Clin Genet* 65: 226-32.
- Tzschach A, Grasshoff U, Beck-Woedl S, Dufke C, Bauer C, Kehrer M, Evers C, Moog U, Oehl-Jaschkowitz B, Di Donato N, Maiwald R, Jung C, Kuechler A, Schulz S, Meinecke P, Spranger S, Kohlhasse J, Seidel J, Reif S, Rieger M, Riess A, Sturm M, Bickmann J, Schroeder C, Dufke A, Riess O, Bauer P (2015) Next-generation sequencing in X-linked intellectual disability. *Eur J Hum Genet* 23: 1513-8. doi: 10.1038/ejhg.2015.5
- Utami KH, Winata CL, Hillmer AM, Aksoy I, Long HT, Liany H, Chew EG, Mathavan S, Tay SK, Korzh V, Sarda P, Davila S, Cacheux V (2014) Impaired development of neural-crest cell-derived organs and intellectual disability caused by MED13L haploinsufficiency. *Hum Mutat* 35: 1311-20. doi: 10.1002/humu.22636
- Utin GE, Haliloglu G, Volkan-Salanci B, Cetinkaya A, Kiper PO, Alanay Y, Aktas D, Anlar B, Topcu M, Boduroglu K, Alikasifoglu M (2014) Etiological yield of SNP microarrays in idiopathic intellectual disability. *Eur J Paediatr Neurol* 18: 327-37. doi: 10.1016/j.ejpn.2014.01.004
- van Bon BW, Koolen DA, Brueton L, McMullan D, Lichtenbelt KD, Ades LC, Peters G, Gibson K, Moloney S, Novara F, Pramparo T, Dalla Bernardina B, Zoccante L, Balottin U, Piazza F, Pecile V, Gasparini P, Guerci V, Kets M, Pfundt R, de Brouwer AP, Veltman JA, de Leeuw N, Wilson M, Antony J, Reitano S, Luciano D, Fichera M, Romano C, Brunner HG, Zuffardi O, de Vries BB (2010) The 2q23.1 microdeletion syndrome: clinical and behavioural phenotype. *Eur J Hum Genet* 18: 163-70. doi: 10.1038/ejhg.2009.152
- van de Kamp JM, Betsalel OT, Mercimek-Mahmutoglu S, Abulhoul L, Grunewald S, Anselm I, Azzouz H, Bratkovic D, de Brouwer A, Hamel B, Kleefstra T, Yntema H, Campistol J, Vilaseca MA, Cheillan D, D'Hooghe M, Diogo L, Garcia P, Valongo C, Fonseca M, Frints S, Wilcken B, von der Haar S, Meijers-Heijboer HE, Hofstede F, Johnson D, Kant SG, Lion-Francois L, Pitelet G, Longo N, Maat-Kievit JA, Monteiro JP, Munnich A, Muntau AC, Nassogne MC, Osaka H, Ounap K, Pinard JM, Quijano-Roy S, Poggenburg I, Poplawski N, Abdul-Rahman O, Ribes A, Arias A, Yapliito-Lee J, Schulze A, Schwartz CE, Schwenger S, Soares G, Sznajer Y, Valayannopoulos V, Van Esch H, Waltz S, Wamelink MM, Pouwels PJ, Errami A, van der Knaap MS, Jakobs C, Mancini GM, Salomons GS (2013) Phenotype and genotype in 101 males with X-linked creatine transporter deficiency. *J Med Genet* 50: 463-72. doi: 10.1136/jmedgenet-2013-101658
- van de Kamp JM, Mancini GM, Salomons GS (2014) X-linked creatine transporter deficiency: clinical aspects and pathophysiology. *J Inher Metab Dis* 37: 715-33. doi: 10.1007/s10545-014-9713-8
- van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, Giltay JC, van Haaften G (2015) Further confirmation of the MED13L haploinsufficiency syndrome. *Eur J Hum Genet* 23: 135-8. doi: 10.1038/ejhg.2014.69
- Verloes A, Bremond-Gignac D, Isidor B, David A, Baumann C, Leroy MA, Stevens R, Gillerot Y, Heron D, Heron B, Benzacken B, Lacombe D, Brunner H, Bitoun P (2006) Blepharophimosis-mental retardation (BMR) syndromes: A proposed clinical classification of the so-called Ohdo syndrome, and delineation of two new

- BMR syndromes, one X-linked and one autosomal recessive. *Am J Med Genet A* 140: 1285-96. doi: 10.1002/ajmg.a.31270
- Vetro A, Goidin D, Lesende I, Limongelli I, Ranzani GN, Novara F, Bonaglia MC, Rinaldi B, Franchi F, Manolakos E, Lonardo F, Scarano F, Scarano G, Costantino L, Tedeschi S, Giglio S, Zuffardi O (2018) Diagnostic application of a capture based NGS test for the concurrent detection of variants in sequence and copy number as well as LOH. *Clin Genet* 93: 545-556. doi: 10.1111/cge.13060
- Villela D, Costa SS, Vianna-Morgante AM, Krepschi ACV, Rosenberg C (2017) Efficient detection of chromosome imbalances and single nucleotide variants using targeted sequencing in the clinical setting. *Eur J Med Genet* 60: 667-674. doi: 10.1016/j.ejmg.2017.08.020
- Visser LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, van Lier B, Arts P, Wieskamp N, del Rosario M, van Bon BW, Hoischen A, de Vries BB, Brunner HG, Veltman JA (2010) A de novo paradigm for mental retardation. *Nat Genet* 42: 1109-12. doi: 10.1038/ng.712
- Visser LE, Gilissen C, Veltman JA (2016) Genetic studies in intellectual disability and related disorders. *Nat Rev Genet* 17: 9-18. doi: 10.1038/nrg3999
- Vogt DL, Gray CD, Young WS, 3rd, Orellana SA, Malouf AT (2007) ARHGAP4 is a novel RhoGAP that mediates inhibition of cell motility and axon outgrowth. *Mol Cell Neurosci* 36: 332-42. doi: 10.1016/j.mcn.2007.07.004
- Vulto-van Silfhout AT, de Vries BB, van Bon BW, Hoischen A, Ruitkamp-Versteeg M, Gilissen C, Gao F, van Zwam M, Hartevelde CL, van Essen AJ, Hamel BC, Kleefstra T, Willemsen MA, Yntema HG, van Bokhoven H, Brunner HG, Boyer TG, de Brouwer AP (2013) Mutations in MED12 cause X-linked Ohdo syndrome. *Am J Hum Genet* 92: 401-6. doi: 10.1016/j.ajhg.2013.01.007
- Wagenstaller J, Spranger S, Lorenz-Depiereux B, Kazmierczak B, Nathrath M, Wahl D, Heye B, Glaser D, Liebscher V, Meitinger T, Strom TM (2007) Copy-number variations measured by single-nucleotide-polymorphism oligonucleotide arrays in patients with mental retardation. *Am J Hum Genet* 81: 768-79. doi: 10.1086/521274
- Wang H, Shen Q, Ye LH, Ye J (2013) MED12 mutations in human diseases. *Protein Cell* 4: 643-6. doi: 10.1007/s13238-013-3048-3
- Wang Y, Du X, Bin R, Yu S, Xia Z, Zheng G, Zhong J, Zhang Y, Jiang YH, Wang Y (2017) Genetic Variants Identified from Epilepsy of Unknown Etiology in Chinese Children by Targeted Exome Sequencing. *Sci Rep* 7: 40319. doi: 10.1038/srep40319
- Weissman AM (2001) Themes and variations on ubiquitylation. *Nat Rev Mol Cell Biol* 2: 169-78. doi: 10.1038/35056563
- Woodbury-Smith M, Nicolson R, Zarrei M, Yuen RKC, Walker S, Howe J, Uddin M, Hoang N, Buchanan JA, Chrysler C, Thompson A, Szatmari P, Scherer SW (2017) Variable phenotype expression in a family segregating microdeletions of the NRXN1 and MBD5 autism spectrum disorder susceptibility genes. *NPJ Genom Med* 2. doi: 10.1038/s41525-017-0020-9
- Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, van Kogelenberg M, King DA, Ambridge K, Barrett DM, Bayzietnova T, Bevan AP, Bragin E, Chatzimichali EA, Gribble S, Jones P, Krishnappa N, Mason LE, Miller R, Morley KI, Parthiban V, Prigmore E, Rajan D, Sifrim A, Swaminathan GJ, Tivey AR, Middleton A, Parker M, Carter NP, Barrett JC, Hurles ME, FitzPatrick DR, Firth HV, study DDD (2015) Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet* 385: 1305-14. doi: 10.1016/S0140-6736(14)61705-0
- Wright CF, FitzPatrick DR, Firth HV (2018) Paediatric genomics: diagnosing rare disease in children. *Nat Rev Genet* 19: 253-268. doi: 10.1038/nrg.2017.116

- Yamamoto T, Shimojima K (2015) A novel MED12 mutation associated with non-specific X-linked intellectual disability. *Hum Genome Var* 2: 15018. doi: 10.1038/hgv.2015.18
- Yao XY, Zhou JD, Yang J, Zhang W, Ma JC, Wen XM, Yao DM, Xu ZJ, Wu DH, He PF, Qian J, Lin J (2018) SETBP1 mutations in Chinese patients with acute myeloid leukemia and myelodysplastic syndrome. *Pathol Res Pract* 214: 706-712. doi: 10.1016/j.prp.2018.03.010
- Yigit G, Brown KE, Kayserili H, Pohl E, Caliebe A, Zahnleiter D, Rosser E, Bogershausen N, Uyguner ZO, Altunoglu U, Nurnberg G, Nurnberg P, Rauch A, Li Y, Thiel CT, Wollnik B (2015) Mutations in CDK5RAP2 cause Seckel syndrome. *Mol Genet Genomic Med* 3: 467-80. doi: 10.1002/mgg3.158
- Yin JW, Wang G (2014) The Mediator complex: a master coordinator of transcription and cell lineage development. *Development* 141: 977-87. doi: 10.1242/dev.098392
- Zaqout S, Morris-Rosendahl D, Kaindl AM (2017) Autosomal Recessive Primary Microcephaly (MCPH): An Update. *Neuropediatrics* 48: 135-142. doi: 10.1055/s-0037-1601448
- Zhao Y, Alexandrov PN, Jaber V, Lukiw WJ (2016) Deficiency in the Ubiquitin Conjugating Enzyme UBE2A in Alzheimer's Disease (AD) is Linked to Deficits in a Natural Circular miRNA-7 Sponge (circRNA; ciRS-7). *Genes (Basel)* 7. doi: 10.3390/genes7120116