



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA
PROGRAMA DE PÓS-GRADUAÇÃO MESTRADO EM GENÉTICA

**ANÁLISE DA ORIGEM PARENTAL DA VARIAÇÃO NO
NÚMERO DE CÓPIAS *de novo* PATOGÊNICAS EM
PACIENTES COM DEFICIÊNCIA INTELECTUAL**

Goiânia – GO

© 2018

SAMARA SOCORRO SILVA PEREIRA

ANÁLISE DA ORIGEM PARENTAL DA VARIAÇÃO NO
NÚMERO DE CÓPIAS *de novo* PATOGÊNICA EM PACIENTES
COM DEFICIÊNCIA INTELECTUAL

Dissertação apresentada ao Programa de Pós-Graduação
Mestrado em Genética (MGene) da Pontifícia Universidade
Católica de Goiás, como parte das exigências para a obtenção do
título de Mestre em Genética

Orientador: Prof. Dr. Cláudio Carlos da Silva
Co-Orientadora: Prof^a. Dr^a. Lysa Bernardes Minasi

Goiânia – GO

© 2018

P436a

Pereira, Samara Socorro Silva

Análise da origem parental da variação no número de cópias de novo patogênicas em pacientes com deficiência intelectual[recurso eletrônico]/ Samara Socorro Silva Pereira.-- 2018.

54 f.; il.

Texto em português com resumo em inglês
Dissertação (mestrado) - Pontifícia Universidade Católica de Goiás, Programa de Pós-Graduação Stricto Sensu em Genética, Goiânia, 2018
Inclui referências f. 44-49

1. Genética. 2. Incapacidade intelectual. 3. Genoma humano. I.Silva, Cláudio Carlos da. II.Pontifícia Universidade Católica de Goiás. III. Título.

CDU: 616.89-008.446(043)

ATA COMPLEMENTAR Nº 141/2018

MESTRADO EM GENÉTICA DA PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS

DISCENTE: SAMARA SOCORRO SILVA PEREIRA

DEFENDIDA EM 14 de março de 2018 e Aprovada COM CONCEITO..... A

O título foi alterado não () sim _____

BANCA EXAMINADORA



Prof. Dr. Claudio Carlos da Silva
PUC Goiás (Presidente)



Prof. Dr. Aparecido Divino da Cruz
PUC Goiás



Profa. Dra. Elisângela de Paula Silveira Lacerda/
Membro externo-UFG

Agradecimento

“A gratidão é a memória do coração,” frase atribuída ao filósofo Antístenes, 440 - 365 a.C.

Agradecer sempre foi a minha melhor oração. Sou grata a todos que passaram por mim e deixou um pouco de si. Assim, agradeço a Deus pelo dom da vida e pelas pessoas que cruzaram meu caminho e que contribuíram com minha formação moral e profissional.

Agradeço imensamente tudo o que os meus pais, José Mariano Pereira e Maurane Silva Pereira, fizeram por mim durante a minha vida. O meu caminho até chegar aqui tem mais esforços da parte dos meus pais do que da minha própria parte. Para mim, as mãos calejadas dos meus pais são sagradas, e representam a minha fortaleza. Por vocês eu cheguei até aqui e desejo ir mais adiante.

Agradeço a presença constante dos meus tios na minha formação. Agradeço à tia Imadalva Maria dos Santos em nome dos meus tios paternos. E agradeço *in memoriam* minha tia Maurene Silva Camargo e *in memoriam* meu tio Mauro Silva por terem sido em vida pessoas que tanto me amaram e torceram por minhas conquistas. Carregarei por toda vida a ternura de cada abraço e zelo que cada tio teve comigo. Ainda no cenário da minha família, agradeço pela alegria e por toda força que meu irmão, Weffersson dos Reis Silva Pereira, e meus primos me proporcionam, em nome de todos os meus primos, eu agradeço Josino Silva Neto, Sara Cristina Fonseca da Silva e Sabrina Amersur da Silva *in memoriam*.

Agradeço a todos os meus professores, em nome da professora Daniela Cristina Gonçalves de Araújo, eu agradeço meus professores do Colégio Estadual Santa Luzia (atual Centro de Ensino em Tempo Integral Santa Luzia). Agradeço também aos queridos professores Dr. Alex Silva da Cruz e Dr^a. Caroline Oliveira de Melo Araújo, pelo apoio durante a graduação e o mestrado, assim eu estendo meus agradecimento para todos os meus professores universitários.

Agradeço à Pontifícia Universidade Católica de Goiás por ter proporcionado toda a estrutura para a conclusão deste mestrado, assim estendo os agradecimentos a todos os professores, técnicos e alunos do Núcleo de Pesquisas Replicon (NPR), por terem me oferecido todo suporte durante os anos que estive no NPR para estudar sobre genética, possibilitando minha caminhada desde a iniciação científica até o mestrado. Agradeço aos técnicos Eduardo Rocha Pedrosa, Ma. Damiana Mírian da Cruz e Cunha e Me. Cristiano Luiz

Ribeiro, e estendo o agradecimento para a dona Orlandina Barbosa dos Santos. Agradeço também a secretária do mestrado em Genética, Alessandra Malta de Olivera pela disponibilidade e prontidão.

Agradeço ao professor Dr. Aparecido Divino da Cruz (Peixoto), coordenador do MGene, em nome de todos os professores do mestrado. Agradeço imensamente ao professor por ter me acolhido tão bem no NPR e por ter sido o melhor exemplo de profissional que eu conheci. Estendo meus agradecimentos ao professor Dr. Cláudio Carlos da Silva por ter me acompanhado até aqui na missão de ser meu orientador. Agradeço por ter me acolhido como orientada desde a iniciação científica e durante o mestrado, sou muito grata pela confiança a mim depositada.

Agradeço à professora Dr^a Lysa Bernardes Minasi por ter aceitado o convite para ser minha co-orientadora. Agradeço imensamente a paciência e ajuda para a conclusão deste trabalho. Agradeço também a Irene Plaza Pinto MSc pela imensa contribuição durante todo o processo de preparação deste documento. Vocês duas são muito especiais!

Seguir a vida acadêmica requer muitos desafios e às vezes gera estresse. Porém a amizade torna este caminho mais leve. Assim, agradeço aos amigos que fiz ao longo da minha caminhada na vida acadêmica. Agradeço à Lilian Teodoro de Souza, Lorrynne Guimarães de Oliveira e Nayara Lopes de Sousa por terem me oferecido sempre um ombro amigo. Eu estarei sempre torcendo pelo sucesso de vocês. Estendo meus agradecimentos também para meus amigos Calebe Bertolino Marins de Campos e Lucas Henrique Nascimento Silva Rodrigues.

Enfim, agradeço à professora Nicole de Leeuw, PhD, da *Radboud University Medical Center, do Department of Human Genetics*, localizado em Nijmegen na *Netherlands*, pela disponibilidade na ajuda das análises da origem parental das CNVs. A contribuição da professora Nicole foi essencial para a conclusão deste estudo.

Agradeço com muito carinho a família de todos os pacientes com deficiência intelectual que participaram deste estudo.

Enfim, agradeço a Capes pela bolsa me concedida realização do mestrado.

Obrigada a todos!

Sumário

Tabelas e figuras.....	VIII
Resumo.....	IX
Abstract.....	X
1. INTRODUÇÃO.....	11
2. REFERENCIAL TEÓRICO.....	13
2.1 Deficiência intelectual e seus aspectos etiológicos.....	13
2.2 Alterações numéricas e estruturais no genoma humano.....	15
2.3 Variação no número de cópia - CNV.....	17
2.3.1 Mecanismo de formação de CNVs.....	18
2.3.1.1 Recombinação homóloga não Alélica - NAHR.....	20
2.4. Análise cromossômica por microarranjo - CMA.....	25
3. OBJETIVOS.....	26
3.1 Objetivo Geral.....	26
3.2 Objetivos Específicos.....	26
4. MATERIAIS E MÉTODOS.....	27
4.1 Grupo amostral.....	27
4.2 Cariótipo por bandeamento GTG.....	28
4.3 Extração e quantificação das amostras de DNA.....	28
4.4 Análise cromossômica por microarranjo - CMA.....	29
4.5 Descrição do método de CMA.....	29
4.6 Tratamento e análise dos dados obtidos na genotipagem dos SNPs.....	30
4.7 Análise das CNVs	30
4.8 Análise da origem através da taxa de erro mendeliano.....	31
4.9 Análise de LCR.....	31
5. RESULTADOS E DISCUSSÃO	33
7. CONCLUSÃO.....	43

8.	REFERÊNCIAS BIBLIOGRÁFICAS.....	42
9.	ANEXOS	50
9.1	Anexo 1: <i>Termo de Consentimento Livre e Esclarecido.....</i>	50
9.2	Anexo 2: <i>DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources</i>	53

Lista de quadros, tabelas e figuras

	Pág
Quadro	
Quadro 1 Incapacidade intelectual aferida a partir da aplicação do teste do Quociente de Inteligência e o comportamento de domínios e habilidades dos pacientes.....	13
Figuras	
Figura 1. Via de reparo de DSB através do processo de recombinação homóloga.....	21
Figura 2. Representação esquemática de duplicações e deleções recíprocas de recombinação homóloga não alélica.....	22
Figura 3-a. Rearranjo cromossômico resultante da recombinação entre LCRs na mesma direção tendo como resultado a deleção e a duplicação.....	23
Figura 3-b. Rearranjo cromossômico resultante da recombinação entre LCRs invertidas tendo como resultado a inversão.....	23
Figura 4. Fluxograma apresentando a dinâmica da seleção dos trios incluídos para o estudo da origem das CNVs <i>de novo</i> patogênicas em pacientes com indicação clínica de DI selecionados no período de 2013 a 2015.....	28
Figura 5. Gráfico com a porcentagem da quantidade de CNV perda (tonalidade vermelha) e de ganho (tonalidade azul) com a subdivisão do tamanho do segmento, menor ou maior que 5Mb.....	34
Figura 6. Representação esquemática do ideograma humano representado as CNVs <i>de novo</i> patogênicas identificadas pelo presente estudo.....	35
Figura 7-a e 7-b. Representação da origem parental das CNVs de perda e de ganho em relação ao tamanho em Mb.....	37
Tabelas	
Tabela 1. Características clínicas e moleculares dos pacientes que apresentaram CNV <i>de novo</i> patogênica e análise da origem da CNV.....	39
Tabela 2. Análise da presença de LCRs flanqueando as CNVs.....	40

Resumo

A variação no número de cópias (CNV) no genoma é um dos fatores etiológicos que pode desencadear a condição da deficiência intelectual (DI), sendo que esta condição atinge cerca de 2% da população mundial. A metodologia de análise cromossômica por microarranjo (CMA) além de ser indicada como teste de primeira escolha para pacientes com DI, tem ajudado também na compreensão da formação de CNVs e classificação destes rearranjos, quanto à patogenicidade, o tipo de recorrência e sua origem. E este estudo objetivou inferir a origem cromossômica parental das CNVs *de novo* patogênicas em pacientes com DI e seu mecanismo de formação. Os pacientes com indicação clínica de DI foram encaminhados ao Núcleo de Pesquisas Replicon/LaGene para realização do cariótipo com bandeamento GTG, e subsequentemente, os que não tiveram alteração numérica e/ou estrutural no cariótipo foram convidados a continuar a investigação em nível genômico, pela metodologia de CMA. Após realização do CMA e classificação das CNVs, foram realizadas a análise da origem parental das CNVs *de novo* patogênicas pela análise do erro mendeliano usando os marcadores de SNPs disponibilizado pelo *software ChAS*. Adicionalmente, foi usado o *UCSC Genome Browser* para detectar Repetições De Poucas Cópias (LCR) circundantes as CNVs para inferir o mecanismo de formação das mesmas. Foi realizado o cariótipo em 290 pacientes com indicação clínica de DI entre os anos de 2013 a 2015 e em 193/290 (66,5%) foram diagnosticados pelo cariótipo. Do conjunto de pacientes que não foram diagnosticados usando o cariótipo, apenas 76/97 (78,3%) aceitaram continuar a investigação pelo CMA. Após realizar o CMA, foi observado 15 CNVs *de novo* patogênicas, 10 CNVs de perda e 5 CNVs de ganho, em 13/76 (17,1%) pacientes. Na análise da origem parental, observou-se que 60% das CNVs são de origem materna e 40% de origem paterna. Não foi possível detectar a influência da idade parental na formação das CNVs. Ao analisar a presença de LCRs circundantes, observou-se que 46,7% das CNVs *de novo* patogênicas são recorrentes e o mecanismo de formação foi a Recombinação Homóloga Não Alélica (NAHR), e 71,4% dessas CNVs recorrentes são de origem materna. Esses dados corroboram com os estudos que afirmam que a maioria das CNVs de origem paterna são não recorrentes devido às células germinativas replicarem inúmeras vezes o seu material genético na fase pré-meiótica, sendo possível inferir sobre o mecanismo de formação que pode ter sido por Replicação Induzida por Quebra e Mediada por Microhomologia (MMBIR) ou Junção de Extremidade Não Alélica (NHEJ).

Palavras-Chave: CMA; LCR; CNV recorrente; NAHR.

Abstract

Copy Number Variation (CNV) has been associated with intellectual disability (ID) and this condition occur in approximately 2% of world population. Chromosomal Microarray Analysis (CMA) is being indicated as first-tier test for individuals with ID and has also helped to understand the mechanisms of CNV formation, classification of these rearrangements, type of recurrence, and its origin. The aim of this study was to infer the parental chromosome origin of *de novo* pathogenic CNV in patients with ID and their mechanisms of formation. Patients with clinical indications of ID were referred to Replicon Research Group/LaGene for G-band karyotyping. CMA approach was done for patients without numerical and/or structural rearrangements results in karyotype. After performing CMA and classification of CNVs, the parental origin of pathogenic CNVs was done using Mendelian error check based on SNPs markers available by ChAs software. In addition, the UCSC Genome Browser website was used to detect Low Copy Repeats (LCR) surrounding the CNVs to infer the mechanisms of their formation. In the period from 2013 to 2015 was performed G-band karyotyping in 290 patients with clinical indication of ID and a total of 193/290 (66.5%) were diagnosed by Karyotype. The group of patients who were not diagnosed using the karyotype, only 76/97 (78.3%) agreed to continue the investigation by CMA's approach. After performing CMA, a total of 15 *de novo* pathogenic CNVs were observed, 10 CNV of loss and 5 CNV of gain, in 13/76 (17.1%) patients. The analysis of the parental origin showed 60% of CNVs are of maternal origin and 40% of paternal origin. It was not possible to detect the influence of parental age in the formation of CNVs. After analyzing the presence of surrounding LCRs, it was observed that 46.7% are recurrent CNVs and the mechanism of formation was Non-Allelic Homologous Recombination (NAHR), and 71.4% of these recurrent CNVs are of maternal origin. These data are in agreement with studies that affirm that the majority of CNVs of paternal origin are nonrecurrent due to germ cells replicate many times their genetic material in the pre-meiotic phase, being possible to infer the mechanism of formation of CNV that may have been by Microhomology-mediated break-induced replication (MMBIR) or Non-homologous end joining (NHEJ).

Keywords: CMA; LCR; recurrent CNV; NAHR.

1 – INTRODUÇÃO

A deficiência intelectual (DI) foi definida em 2011 pela Organização Mundial da Saúde no documento Relatório Mundial sobre Deficiência (do inglês “*World Report on Disability*”) como a dificuldade de entender, aprender e recordar coisas novas e aplicar estas como forma de aprendizagem em novas situações. A DI atinge 1% dos nascidos vivos (MAULIK et al. 2011) e estes podem apresentar grau variável de intensidade, a saber: leve, moderada, grave e profunda (VISSERS et al. 2016). Para definir se o paciente realmente possui deficiência intelectual e seu nível é aplicado o teste de quociente de inteligência (QI) relacionado com a idade e a realidade sociocultural do indivíduo e, a partir da pontuação obtida no teste, uma pessoa é diagnosticada com ou sem deficiência intelectual (SCHALOCK et al. 2010).

Os fatores etiológicos da DI podem ser ambientais e genéticos; Entre os ambientais, destaca-se o consumo de álcool durante a gravidez, as infecções, as complicações perinatais e desnutrição materna e da criança. Dentre os fatores genéticos destaca-se o comprometimento de genes do desenvolvimento e metabolismo neural. Por causa do seu comportamento genético, quadros de DI têm tornado um motivo frequente de encaminhamento destes pacientes para os serviços genéticos pediátricos para fins de diagnóstico clínico e laboratorial (MEFFORD et al. 2012).

O cariótipo com bandeamento GTG é um teste robusto que tem a capacidade de visualizar de 400 a 550 bandas nos cromossomos, detectando alterações maiores que 5-10 Mb. Dos pacientes com DI que são encaminhados para a realização de testes genéticos, cerca de 5% deles possuem alterações numéricas e/ou estruturais visíveis microscopicamente no cariótipo bandeado (SCHAAF et al. 2011).

A Análise Cromossômica por Microarranjo (CMA) é a metodologia recomendada pelos consensos internacionais como o teste de primeira escolha para diagnóstico dos casos com indicação clínica de deficiência intelectual, Transtorno do Espectro Autista (TEA) e Anomalias Congênitas Múltiplas (ACM). A CMA é capaz de detectar alterações genômicas do tipo Variação no Número de Cópia (CNV) de ganhos e/ou perdas menores que 5 Mb, que não é detectado pelo cariótipo com bandeamento GTG. O CMA se destaca por conseguir detectar perdas e ganhos submicroscópicos diagnosticando 15 a 20% de pacientes com DI, TEA e ACM (MILLER et al. 2010; NOWAKOWSKA, 2017).

As CNVs são segmentos de perdas ou ganhos submicroscópicas resultando na mudança da diploidia do *locus* onde ela se encontra (GIRIRAJAN et al. 2011). Devido à capacidade de modular o fenótipo humano, as CNVs são alvos de estudo (ESCARAMÍS et al. 2015). Na tentativa de identificar e classificar as CNVs, atualmente existem muitas plataformas de *arrays* disponíveis para auxiliar em diferentes tipos de pesquisas. Com o aumento da resolução das matrizes de SNP para a CMA, aumentou-se também o rendimento do diagnóstico e possibilitou detectar também variações benignas ou variantes de significância clínica incerta (VOUS) (BORLOT et al. 2017). Com o avanço das metodologias ao longo de 20 anos de estudo sobre as desordens genômicas, foi possível classificar diferentes tipos de CNVs para melhor compreender cada uma delas, como os mecanismos de formação das CNVs recorrentes e não recorrentes que afetam a distribuição destas variações nas populações (HAREL; LUPSKI, 2017).

A partir de 2010, têm sido propostas para a classificação das CNVs quanto à patogenicidade. Em geral, a patogenicidade de uma CNV é aferida levando-se em consideração a posição onde ela se encontra no genoma, se está envolvendo regiões de síndromes conhecidas, se envolve genes importantes para o desenvolvimento e/ou funcionamento neurológico e sua ocorrência em populações saudáveis. Uma importante análise, indicada pelos consensos internacionais, é a testagem dos genitores com intuito de detectar se a CNV foi herdada de pais saudáveis ou se é uma CNV *de novo* (MILLER et al. 2010; NOWAKOWSKA, 2017).

As CNVs *de novo* não são herdadas dos genitores, mas podem ser transmitidas se caso forem formadas nas células germinativas de um dos genitores. Sendo assim, o conhecimento dos fatores de riscos subjacentes à sua formação é importante para compreender o que leva a formação de diferentes tipos de CNVs *de novo*. Para a realização de estudos que envolvam esta temática, têm sido desenvolvidas plataformas de *array* com a utilização de marcadores de Polimorfismo de Nucleotídeo Único (SNP) para a análise da origem parental das CNVs *de novo* (MA et al. 2017).

2 - REFERENCIAL TEÓRICO

2.1 - Deficiência intelectual e seus aspectos etiológicos

O Relatório Mundial Sobre Deficiência, publicado em 2011 pela Organização Mundial da Saúde (OMS), definiu a deficiência intelectual (DI), anteriormente conhecida como retardo mental ou limitação mental, como dificuldades em entender, aprender e recordar coisas novas e em aplicar essa aprendizagem a novas situações. Outra definição para deficiência intelectual foi apresentada pela *American Association on Intellectual and Developmental Disabilities* (AAIDD) como limitações significativas tanto no funcionamento intelectual como no comportamento adaptativo, abrangendo diversas habilidades sociais e práticas diárias, que se originam antes da idade de 18 anos.

A deficiência intelectual, que atinge cerca de 1% dos nascidos vivos (MAULIK et al. 2011), ocorre de forma isolada ou em conjunto com outras malformações congênicas podendo apresentar características neurológicas como deficiência sensorial, epilepsia e sinais do Transtorno do Espectro Autista (TEA), além de exibir gravidade variável: leve, moderada, grave e profunda (VISSERS et al. 2016).

O funcionamento intelectual de pessoas que apresentam características de DI pode ser avaliado pela aplicação de Teste de Quociente de Inteligência (QI). A interpretação dos resultados do QI está apresentado no Quadro 1. A interpretação depende das respostas dos testes, relacionadas com a idade e a realidade sociocultural do indivíduo avaliado. Quando o teste apresenta pontuação inferior a 70, limitação intelectual é identificada (LUCKASSON et al. 2002; SCHALOCK et al. 2010).

Quadro 1. Incapacidade intelectual aferida a partir da aplicação do teste do Quociente de Inteligência e o comportamento de domínios e habilidades dos pacientes

Grau	Faixa de QI	% dos casos	Capacidade adulta
Leve	50-70	80	<ul style="list-style-type: none">• Alfabetização +• Habilidades de autoajuda ++• Boa fala ++• Trabalho semiqualficado +
Moderado	35-50	12	<ul style="list-style-type: none">• Alfabetização +/-• Habilidades de autoajuda +• Fala em casa +• Trabalho não qualificado, com ou sem supervisão +
Grave	20-35	3 a 4	<ul style="list-style-type: none">• Habilidades de autoajuda assistidas +• Fala mínima +• Tarefas domésticas assistidas +
Profundo	Abaixo de 20	1 a 2	<ul style="list-style-type: none">• Fala +/-• Habilidades de autoajuda +/-

Nota: +/- algumas vezes atingível; + atingível; ++ definitivamente atingível.

Fonte: Tratado de Saúde Mental da Infância e Adolescência da IACAPAP (KE; LIU, 2015), com adaptação.

Pessoas que apresentam deficiência intelectual profunda não podem cuidar de si e possuem uma linguagem muito prejudicada, capacidade de expressar emoções limitada e pouco compreendida (ADAMS; OLIVER, 2011). Também é comum apresentarem convulsões, deficiências físicas e expectativa de vida reduzida. As pessoas que possuem DI grave tem dificuldade em pronunciar palavras e possui vocabulário limitado. Precisam de apoio em casa, na escola e na comunidade para poder desenvolver habilidades básicas de autoajuda. Já os indivíduos que apresentam DI moderada, conseguem se comunicar e cuidar de si mesmo com algum apoio, além de poder realizar trabalhos não qualificados ou semiquilificados sob supervisão. As pessoas que possuem DI leve são capazes de se comunicar bem e aprender habilidades básicas, como leitura e informática, além de conseguirem realizar trabalhos domésticos e cuidar de si (KE; LIU, 2015).

A DI pode ser caracterizada como síndrômica e não síndrômica. É classificada como DI síndrômica quando está associada a fenótipos síndrômicos, que incluem malformação congênita, dismorfia, convulsões e entre outras sintomatologias reconhecidas ao exame físico do paciente. Por outro lado, na DI não síndrômica, a única manifestação clínica do paciente é a própria DI. Pacientes com DI síndrômica podem apresentar sinais e sintomas que permitem enquadrá-los em síndromes genéticas já descritas (ROCHA, 2014).

A etiologia da DI tem influência de fatores exógenos, como consumo de álcool pela mãe durante a gravidez, infecções, complicações durante o nascimento e desnutrição extrema. Porém os fatores genéticos são os mais conhecidos por apresentar papel importante no desenvolvimento da deficiência intelectual (KE; LIU, 2015).

Devido à complexidade orgânica e funcional do cérebro, formadas por células altamente especializadas que executam suas funções por estarem interligadas, espera-se que durante o desenvolvimento e o funcionamento diário ao longo da vida de uma pessoa, as proteínas participantes do funcionamento deste sistema estejam ativas, na quantidade certa, no local certo e no momento certo. Portanto, uma mutação pontual, um rearranjo genômico ou uma aberração cromossômica que afeta um ou mais genes pode prejudicar a produção destas proteínas, causando graves consequências para o desenvolvimento cerebral ou funcionamento cognitivo (MEFFORD et al. 2012).

A deficiência intelectual passou a ser o principal motivo do encaminhamento para serviços genéticos pediátricos. Porém as características clínicas e genéticas da DI são muito

heterogêneas e, às vezes, complexas o que dificulta o diagnóstico clínico e laboratorial na maioria dos casos (MEFFORD et al. 2012).

Os indivíduos que apresentam suspeitas de um distúrbio cromossômico como causa da DI são indicados para realizar o teste de cariotipagem com bandeamento GTG. Para aplicação desta técnica, obtém-se cromossomos em metáfase com uma resolução de 400 a 550 bandas. Porém este nível de resolução não detecta alterações cromossômicas que sejam menores de 5 Mb. O teste do cariótipo consegue diagnosticar cerca de 5% dos pacientes que apresentam DI não síndrômica (SCHAAF et al. 2011). A análise cromossômica por microarranjo (CMA), indicada pelos consensos internacionais como teste de primeira escolha para diagnóstico da DI, pode detectar rearranjos menores que 5 Mb. A análise por *microarray* fornece cerca de 15% a 20% de rendimento adicional no diagnóstico quando comparado ao cariótipo (HOCHSTENBACH et al. 2009; MILLER et al. 2010; NOWAKOWSKA, 2017).

2.2 Alterações numéricas e estruturais no genoma humano

O genoma humano está distribuído e compactado em 23 pares cromossomos, sendo que 22 pares são autossômicos e 1 par sexual (MALUF; RIEGEL, 2011). As variações nas sequências do genoma são as principais fontes de investigação científica quando se trata de fenótipos comuns e patológicos na tentativa de desvendar os mecanismos da origem e desenvolvimento destas características. As primeiras pesquisas foram o estudo da ploidia e alterações cariotípicas observadas pela da microscopia óptica, seguido pelo dogma central da biologia molecular até chegar aos modelos atuais mais complexos, envolvendo o estudo e controle da expressão gênica e epigenética (LANDER, 2011).

Além das alterações cromossômicas que podem ser numéricas ou estruturais, as variações no genoma humano consistem em dois principais tipos: (1) variação de um único nucleotídeo, como substituições e deleções, e (2) as variações que afetam vários pares de bases, como inserções, deleções e duplicações que resultam na variação do número de cópias (CNV) (LUPSKI et al. 2011). No início das pesquisas em genética, apenas alterações na forma e no número cromossômico puderam ser avaliados. A partir do desenvolvimento de técnicas para identificação e caracterização do cariótipo humano e variações na normalidade, a base genética de alguns distúrbios começou a ser compreendida (MATTEVI; MIRANDA, 2011).

O grupo de pesquisa liderado pelo cientista francês Jérôme Lejeune realizou as primeiras descrições das alações cromossômicas mais importantes para a medicina, caracterizando citogeneticamente a trissomia do cromossomo 21 como responsável pela Síndrome de Down. O mesmo grupo, em 1963 descreveu a primeira síndrome herdada por deleção, denominada Síndrome de Cri-du-chat, envolvendo uma deleção terminal do braço curto do cromossomo 5, que leva a monossomia parcial da região deletada (LEJEUNE et al. 1959; LEJEUNE et al. 1963).

Na década de 1980, foi desenvolvida a aplicação da técnica de Hibridização Fluorescente *in situ* (FISH), que usa sondas de DNA com fluoróforo na extremidade 3' para hibridizar em sequência específica do DNA (BAUMAN et al. 1980). O FISH representou um importante avanço na detecção de rearranjos cromossômicos que permite de forma mais rápida a confirmação laboratorial de síndromes causadas por microdeleção ou microduplicação. A aplicação do FISH permitiu também investigar as deleções e duplicações subteloméricas, que são responsáveis por 2,5 a 5% dos casos de deficiência intelectual que eram incompreendidos anteriormente à aplicação da técnica de FISH (RAVNAN et al. 2006; BALLIF et al. 2007).

A partir do aperfeiçoamento de novas metodologias, foi possível detectar variações na estrutura do genoma, como forma de deleções, duplicações, inversões, inserções e translocações. As variações estruturais no genoma explicam grande parte da diversidade genômica entre os humanos. É importante também avaliar o impacto funcional destas classes de variações utilizando mapas de variantes e conjuntos de referência de genótipos analisando todo o genoma em diversas populações (CONRAD et al. 2010).

Os Polimorfismos de Nucleotídeo Único (SNPs) são variações que ocorrem com frequência aleatória em pelo menos 1% da população. Os SNPs são fontes de variações importantes em todo o genoma. Porém, devido as CNVs serem maiores e envolver mais material genético que os SNPs, naturalmente as CNVs, principalmente as *de novo* são mais reportadas em eventos patogênicos. Devido a isto, as CNVs *de novo* estão sendo causa de debate científico nos últimos anos (MA et al. 2017)

Conhecer a etiologia genética da deficiência intelectual e de outras doenças impacta positivamente no aconselhamento genético das famílias afetadas por essas alterações, possibilitando orientação clínica e manejo adequados dos pacientes. O conhecimento da

etiologia genética das doenças favorece a compreensão dos mecanismos moleculares fundamentais que, quando alterados, podem desencadear diferentes condições, como a DI, e adicionalmente, oferecem uma oportunidade para a terapia personalizada (NG et al. 2010).

2.3 Variação no número de cópia - CNV

As variações no número de cópias (do inglês, *copy number variations*) são ganhos ou perdas submicroscópicas > 50 Kb de segmentos de DNA ao longo do genoma (SUDMANT et al. 2013). As CNVs mudam a diploidia do *locus* onde se encontram devido a presença de microdeleções ou microduplicações (GIRIRAJAN et al. 2011). As CNVs são alvos de diferentes estudos, pois modulam o fenótipo humano contribuindo para singularidade de cada indivíduo (ESCARAMÍS et al. 2015).

A detecção e estudo das CNVs são importantes para a compreensão da diversidade genética em relação a suscetibilidade a doenças, possibilitando a associação com características fenotípicas importantes para entender a etiologia e os mecanismos patogênicos da doença (XU et al. 2011).

Em regiões com CNVs de ganho, geralmente, o resultado são três cópias, diferente do esperado que são duas cópias. É comum o genoma apresentar também regiões que estejam com mais de três cópias. Estudos de análise de cortes de controle apresentaram que algumas regiões do genoma são toleráveis às CNVs e que cada pessoa carrega muitas mudanças no número de cópias que em sua maioria são benignas (SEBAT et al. 2004).

De acordo com o consenso proposto por Miller e et al. (2010), a classificação das CNVs quanto a patogenicidade é dividida em quatro tipos: patogênica, potencialmente patogênica, neutras e de significado clínico incerto. Para que a CNV seja classificada como patogênica, ela deve envolver regiões relacionadas com síndromes conhecidas, ou envolver genes importantes da via do desenvolvimento e/ou funcionamento neural, além de não ser encontrada em populações saudáveis e não ser herdada de pais saudáveis. As CNVs que são potencialmente patogênicas se encontram em regiões que envolvem genes importantes que participam da via do desenvolvimento neural, porém a região envolvida pela CNV não há associação com a DI. As CNVs potencialmente patogênicas não são herdadas de pais saudáveis e não é encontrada em populações saudáveis (MILLER et al. 2010).

As CNVs neutras são herdadas de pais saudáveis e/ou não envolve genes e são encontradas em banco de dados de populações saudáveis. As CNVs de significado clínico incerto estão envolvendo genes que não fazem parte de via do desenvolvimento neural e não são encontradas em populações saudáveis (MILLER et al. 2010).

As CNVs neutras são variações que não estão associadas a nenhum quadro clínico, podem ser encontradas também em grupos controles (1000 GENOMES PROJECT CONSORTIUM, 2010), comumente estão presentes em regiões polimórfica do genoma e são usadas em estudos evolutivos envolvendo genética de populações (SUDMANT et al. 2010; CAMPBELL et al. 2011).

Frequentemente as CNVs são associadas a uma patologia quando englobam parte de um gene, um gene inteiro, um segmento que abrange vários genes ou regiões regulatórias podendo alterar a dosagem gênica causando síndromes de microdeleção ou microduplicação com variáveis fenótipos no portador (HU et al. 2018).

A testagem dos genitores é importante para se detectar a patogenicidade da CNV (PRATTE-SANTOS et al. 2016). As CNVs podem ser classificadas quanto a sua origem: herdada ou *de novo*. Quando se é observado uma CNV no paciente e em um dos progenitores (ou nos dois) no mesmo *loci* e do mesmo tamanho a CNV é classificado como herdada. As CNVs *de novo* são formadas em células germinativas durante a divisão pré-meióticas ou meióticas, ou formadas em células pós-zigóticas. Nestas situações, as variações afetam apenas células da criança. A análise dos genomas parentais, não revela a mesma CNV presente no filho (MA et al. 2017).

O conhecimento científico na área dos distúrbios genômicos, incluindo informações sobre mecanismos de formação de CNVs, sensibilidade à dosagem, traços comuns em portadores, rearranjos complexos completa duas décadas (1998 – 2018). A compreensão dos efeitos das CNVs nos distúrbios genômicos contou essencialmente com o desenvolvimento e aplicação das técnicas de análise cromossômica por microarranjos (HAREL; LUPSKI, 2017).

2.3.1 Mecanismo de formação de CNVs

As CNVs formadas durante a geração das células germinativas são resultado de mecanismos de recombinação homóloga, de erros na segregação cromossômica, de replicação

e de reparo do DNA (MA et al. 2017). O estudo de fatores de risco genéticos e ambientais que influenciam na formação de CNV é importante e necessário para a compreensão dos mecanismos mutacionais, mas ainda são pouco elucidadas. Apesar do conhecimento sobre o papel dos inibidores de replicação na formação das CNVs *de novo*, os diferentes efeitos dos agentes mutagênicos na indução da CNV ainda não são compreendidos (ARLT et al. 2014).

Estudos genômicos de gêmeos monozigóticos com transtornos neuropsiquiátricos buscaram compreender a proporção de formação de CNVs *de novo* pós-zigótica. Dois destes estudos afirmaram não ter encontrado diferença nas CNVs *de novo* em cada um dos pares, sugerindo que a CNV *de novo* em gêmeos monozigóticos tem chance limitadas de serem formadas pré-zigoticamente (BLOOM et al. 2013; STAMOULI et al. 2018).

A compreensão dos mecanismos de formação de variantes estruturais no genoma foi possível a partir de abordagens de modulação utilizando organismos modelos (MA et al. 2017) e do uso de células humanas submetidas a estresses ambientais (COSTANTINO et al. 2014). Além da observação e avaliação direta das alterações genômicas humanas e/ou produtos finais de rearranjos que acarretam em doenças. O estudo de manifestação ou transtorno genômico permite que os eventos mutacionais sejam categorizados nos indivíduos afetados da população (CARVALHO; LUPSKI, 2016).

Durante o processo evolutivo, as células eucarióticas evoluíram mecanismos para reparar diferentes formas de dano na molécula de DNA para conservar com fidelidade a informação genética durante as divisões celulares. Os tipos de danos que a célula tenta reparar são as quebras de dupla-fita (do inglês, *double-strand break* – DSB) que são muito prejudiciais ao genoma pois podem resultar em inserções, deleções ou translocações cromossômicas (PANNUNZIO et al. 2017).

O reparo das DSBs é realizado pelo processo de recombinação de sequência do cromossomo homólogo, sendo usado a informação genética do cromossomo homólogo para reparar o erro da DSB. Outro mecanismo que pode reparar as DSBs é a junção de extremidade não homóloga, onde a maquinaria deste reparo protege e processa as extremidades quebradas para subsequentemente uni-las (WATSON et al. 2015). Estes dois mecanismos de reparo, a recombinação homóloga não alélica (do inglês, *Non-allelic homologous recombination* - NAHR) e a junção de extremidades não homólogas (do inglês, *Non homologous End Joining*

- NHEJ), são considerados os principais mecanismos de formação de rearranjos cromossômicos como as CNVs (STANKIEWICZ; LUPSKI, 2010; ARLT et al. 2011).

2.3.1.1 Recombinação homóloga não Alélica - NAHR

Os rearranjos recorrentes na população, denominados desta forma por serem rearranjos que compartilham tamanho comum, agrupam-se num ponto de interrupção e recorrem em múltiplos indivíduos, geralmente são formados por NAHR devido a presença de repetições de poucas cópias (do inglês, *Low Copy Repeats* – LCRs) que circundam estas regiões de recombinação (SHAW; LUPSKI, 2004). As LCRs compõe 4-5% do genoma humano e possui comprimento entre 10 a 400 Kb, outra característica das LCRs é possuir 95-98% de similaridade entre si (SHAW; LUPSKI, 2004; STANKIEWICZ; LUPSKI, 2010). Os achados na literatura que associam a presença de LCR com rearranjos genômicos relacionados com doenças são explicados devido a presença de grandes LCRs que flanqueiam regiões genômicas importantes e que contribuem negativamente para a estabilidade do genoma tornando esta região mais propensa a rearranjos de DNA via NAHR (SHAW; LUPSKI, 2004; CARVALHO; LUPSKI, 2016).

O reparo por recombinação homóloga normalmente inicia-se quando é reconhecida uma quebra de cadeia dupla (DSB), conforme representado na Figura 1. Na região de DSB realiza-se clivagem na sequência de nucleotídeos para gerar uma lacuna de DNA com extremidade 3' de fita simples (ssDNA) (figura 1-a). Após o reconhecimento da região homóloga, o ssDNA invade a molécula do DNA homólogo (figura 1-b). A fita invasora irá parear com a fita complementar da outra molécula de DNA. A extremidade livre 3' servirá para iniciar a síntese de alongamento do DNA (Figura 1-d). O novo DNA sintetizado usando a molécula homóloga como molde servirá para restaurar as regiões da molécula que foram degradadas durante o processo de reparo (WATSON et al. 2015).

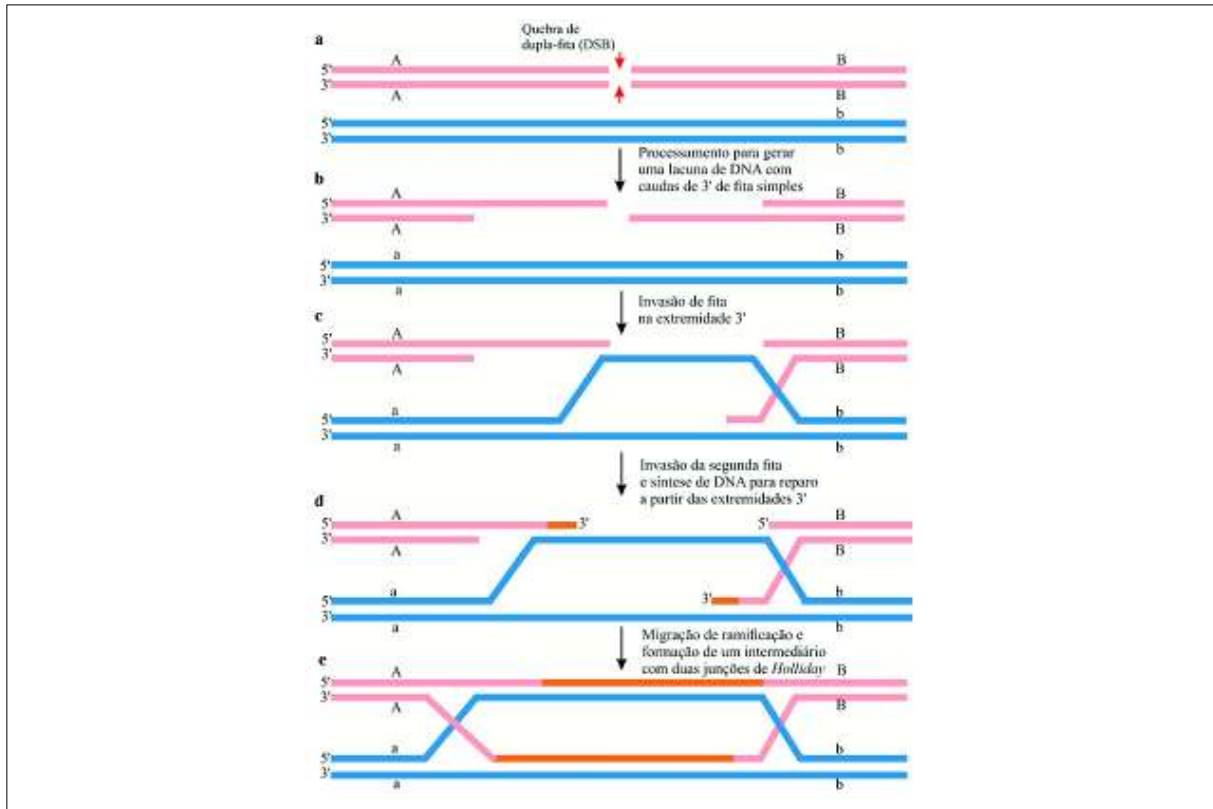


Figura 1. Via de reparo de DSB através do processo de recombinação homóloga (Fonte: Adaptado de Watson et al. 2015).

A região de ssDNA é processada para localizar a sequência homóloga para ser usada como modelo para o reparo. Se a sequência for de ssDNA, ele poderá usar o cromossomo homólogo ou a cromátide irmã para o reparo de recombinação homóloga alélica (*Allelic homologous recombination* - AHR) (HOANG et al. 2010). Devido ao alto grau de similaridade entre as sequências não alélicas das LCRs o reparo poderá recombinar com sequência de homologia não alélica, levando a NAHR e aumentando a chance de ocorrer um rearranjo cromossômico, conforme representado na Figura 2 (GU et al. 2008).

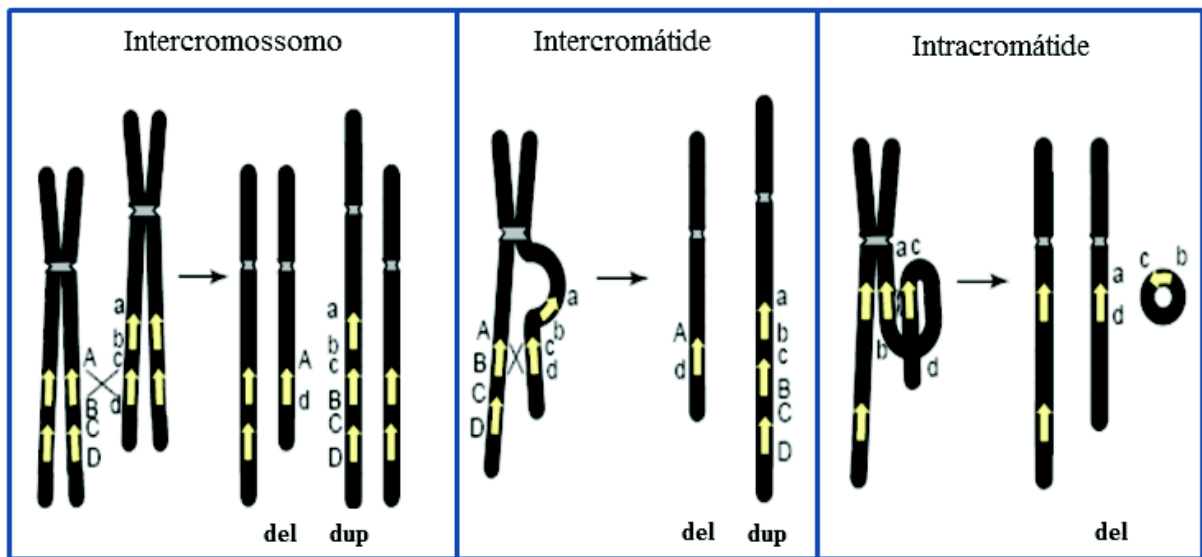


Figura 2 – Representação esquemática de duplicações e deleções recíprocas de recombinação homóloga não alélica. Os cromossomos e as cromátides estão representados de preto e as LCRs pelas setas amarelas. O esquema à direita representa rearranjos entre cromossomos, ao centro intercromátide e à esquerda rearranjo intracromátide (Fonte: Stankiewicz; Lupski 2002, com adaptação).

Estes rearranjos formados por NAHR muitas vezes são resultantes de recombinação entre LCRs que estão na mesma direção ou invertidas. O maior comprimento das LCRs pode aumentar a taxa de NAHR, em contrapartida, o aumento da distância entre as LCRs pode diminuir os eventos de NAHR (STANKIEWICZ; LUPSKI, 2002). Liu e et al. (2011) observaram que o comprimento da LCR correlaciona com a frequência dos rearranjos em linhagens de células germinativas masculinas consistentes com a recombinação ectópica que medeiam NAHR.

A formação dos tipos de rearranjos possíveis em um distúrbio genômico que são delimitados por LCRs são frequentemente observadas no *loci* de origem. Exemplo são o NAHR entre LCR na mesma direção que gera duplicação ou deleção (Figura 3-a) e LCRs invertidas que geram segmentos invertidos como esquematizado na Figura 3-b (CARVALHO; LUPSKI, 2016).

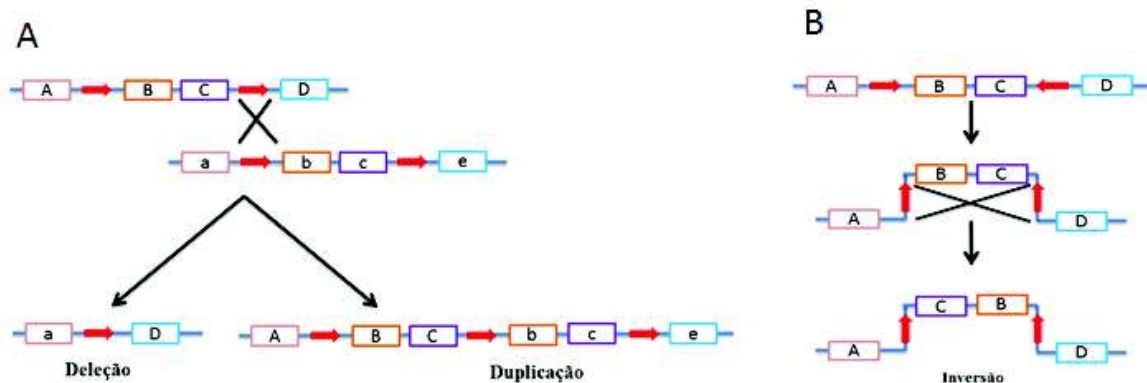


Figura 3-a. Rearranjo cromossômico resultante da recombinação entre LCRs na mesma direção tendo como resultado a deleção e a duplicação. De vermelho as LCRs e dentro de cada retângulo uma sequência única. Figura 3-b. Rearranjo cromossômico resultante da recombinação entre LCRs invertidas tendo como resultado a inversão (Fonte: adaptado de Gu et al. 2008).

2.4. Análise cromossômica por microarranjo - CMA

Os diferentes padrões de variação do genoma podem ajudar a identificar regiões que possuem associação com características biológicas, as quais podem de alguma maneira atuar nos eventos biológicos das populações (LANDER, 2011). Na tentativa de localizar e rastrear as regiões com CNVs, Kellioniemi et al. (1992) propuseram a técnica de hibridação genômica comparativa (do inglês, *Comparative Genomic Hybridisation* - CGH), que possui a capacidade emitir diferentes intensidades de sinais de acordo com a quantidade de fragmento de DNA teste hibridizado, esta intensidade é comparada com uma metáfase referencial (pessoa saudável) e concluindo a quantidade de número de cópias nos *loci*. Segundo Schaaf e colaboradores (2011), os primeiros relatos da aplicação do teste CGH em pacientes com deficiência intelectual e má formação congênita foram realizados em 2004, detectando deleções e duplicações responsáveis pela condição dos pacientes.

Nos anos seguintes a técnica passou a ser chamada de aCGH ou CGH-array devido a utilização de grandes sondas de DNA obtidas através de clones de bactérias (BACs). As sondas são aderidas em uma lâmina de vidro de forma ordenada para o processo de hibridização (SOLINAS-TOLDO et al. 1997; PINKEL et al. 1998). Os trabalhos seguintes, como por exemplo, de Pollack e colaboradores (1999) na tentativa de pesquisar linhagem de células de tumores de câncer de mama, passaram a usar maior resolução através de sondas de cDNAs que são menores e também abrangiam maior parte do genoma. A resolução da técnica de aCGH aumenta de acordo com o encurtamento das sondas e a distância genômica entre cada sonda de DNA (THEISEN, 2008; SUN et al. 2017).

Com o rápido avanço do conhecimento do genoma as plataformas da Illumina e *Affymetrix* criaram matrizes de genotipagem de SNP, sendo usada *array* de SNP ou associada com marcadores não polimórficos, as CNVs (SCHAAF et al. 2011). A adição de SNP *array* acrescentou vantagens para a técnica de CMA, permitindo investigar perda de heterozigosidade, que quando dispersa por todo genoma pode indicar de relação consanguínea ou dissomia uniparental (UPD) quando apresentada em um único cromossomo (HAREL; LUPSKI, 2017). Além disso, os SNPs são usados em pesquisas para inferir a origem parental de CNVs *de novo* observando marcadores polimórficos nos genitores e comparando com o filho (MA et al. 2017).

Atualmente existem várias plataformas de matriz disponíveis para auxiliar o pesquisador em diferentes investigações. Estas matrizes possuem diferença na cobertura, resolução e aplicação do genoma. A resolução do CMA depende do limite entre a distância das sondas na matriz. O aumento da resolução na matriz contribui para um maior rendimento do diagnóstico e também possibilita um maior nível de detecção de CNVs benignas ou variantes de significância clínica desconhecida (VOUS). Portanto, a eficiência na prática clínica depende da escolha correta da plataforma da resolução da matriz que equilibra sensibilidade e especificidade (BORLOT et al. 2017).

As limitações apresentadas pela técnica de análise cromossômica por microarranjos – matrizes de aCGH e de SNP *array* – incluem a incapacidade de detectar rearranjos balanceados, translocação Robertsoniana e inversão, devido não haver perda de segmento genômico nesses rearranjos (SOUTH et al. 2013). Além disso, o CMA não oferece informação sobre a localização e nem a orientação da inserção de sequência de cópia no genoma (HAREL; LUPSKI, 2017). Embora não haja perda e nem ganho de material genético nos rearranjos balanceados, a detecção das translocações balanceadas é importante para o aconselhamento genético devido serem um dos motivos de abortos recorrentes. Neste caso, indica-se realizar a cariótipo (WOU et al. 2016).

Outra limitação da técnica de CMA é a dificuldade de detectar baixo nível mosaïcismo, que segundo o trabalho de Hoang et al. (2010) foi possível detectar até 10% usando uma matriz de aCGH, enquanto o cariótipo detecta 14% (contando 20 células) com 95% de grau de confiança (HOOK, 1977). A detecção de VOUS através da metodologia de CMA é frequentemente criticada pela clínica devido provocar incerteza e trazer ansiedade para os pais durante o aconselhamento genético (WOU et al. 2016).

Apesar das limitações apresentadas pela técnica, o CMA se consagrou como uma excelente metodologia de investigação genômica, sendo indicado como teste de primeira escolha para avaliação de CNV (MILLER et al. 2010). O uso de metodologias de microarranjos tem contribuído não só com a clínica, mas também com a história da investigação dos distúrbios genômicos (HAREL; LUPSKI, 2017).

As investigações genéticas desenvolvidas através da metodologia de CMA têm auxiliado na compreensão dos mecanismos de formação CNVs *de novo*. Neste cenário, pesquisadores buscam investigar e compreender a contribuição de cada genitor para formação de novas mutações (MA et al. 2017).

A contribuição na transmissão de novas mutações da cada genitor para a prole é diferenciada devido o processo da gametogênese masculina ser diferente da gametogênese feminina. Acreditam que as fêmeas contribuem mais com aneuploidias devido a não disjunção cromossômica na meiose, pois as células germinativas femininas são fixadas desde a fase embrionária até a adolescência. Os machos contribuem mais com mutações pontuais geradas na fase pré-meiótica, devido a inúmeros processos de replicação da molécula de DNA (THOMAS et al. 2006). Portanto, pesquisas que visam estudar origem parental de mutações *de novo* são importantes para compreender mecanismos de formação e fatores que influenciam na formação de novas mutações (HEHIR-KWA et al. 2011).

3 – OBJETIVOS

3.1 Objetivo geral

Determinar a origem parental dos cromossomos que apresentam CNVs *de novo* patogênicas em pacientes com indicação clínica da deficiência intelectual.

3.2 Objetivos específicos

- Realizar cariótipo de pacientes encaminhados com indicação clínica de deficiência intelectual para a seleção de pacientes sem alteração numérica e estrutural no cariótipo;
- Realizar a Análise Cromossômica por Microarranjo (CMA) nos pacientes que apresentaram cariótipo sem alteração estrutural e numérica e que os pais aceitaram participar e assinar o TCLE;
- Classificar e identificar as CNVs *de novo* patogênicas;
- Identificar a origem parental das CNVs *de novo* patogênicas e estimar a frequência das CNVs de origem paterna e materna.
- Quantificar a ocorrência de LCR flanqueando as CNVs.

4 – MATERIAIS E MÉTODOS

4.1 Grupo amostral

Os pacientes com indicação clínica de deficiência intelectual foram encaminhados pelos médicos assistentes para o Núcleo de Pesquisas Replicon (NPR) da Escola de Ciências Agrárias e Biológicas da PUC Goiás/Laboratório de Citogenética Humana e Genética Molecular (LaGene) do Laboratório de Saúde Pública Dr. Giovanni Cysneiros (LACEN GO) da Secretaria Estadual de Saúde do Estado de Goiás. O estudo foi aprovado pelo Comitê de Ética e Pesquisa da PUC Goiás sob o protocolo 1721/2011. Os responsáveis pelos pacientes foram esclarecidos quanto ao estudo e assinaram o Termo de Consentimento Livre e Esclarecido (em anexo) Além disso, foram assinados os TCLEs dos bancos de dados DECIPHER como parte de uma colaboração entre a Pontifícia Universidade Católica de Goiás/PUC Goiás com este banco de dados criado para interpretação das variantes genômicas e informações citogenética e clínicas sobre doenças cromossômicas raras.

Todos os pacientes que foram encaminhados do SUS para o NPR/LaGene, do ano de 2013 a 2015, com indicação clínica de deficiência intelectual realizaram o cariótipo com banda GTG para seleção de pacientes sem alteração numérica e estrutural no cariótipo. O fluxograma da seleção de pacientes para a pesquisa é representado na Figura 4.

Entre os anos de 2013-2015, foi encaminhado via Sistema Único de Saúde (SUS) 527 pacientes para o Núcleo de Pesquisas Replicon da PUC Goiás/Laboratório de Citogenética Humana e Genética Molecular (LaGene). Destes 527 pacientes, 290 (55%) foram encaminhados com a indicação clínica de deficiência intelectual. Todos os 290 pacientes fizeram cariótipo com bandeamento GTG, e 193/290 (66,5%) tiveram o diagnóstico a partir da cariotipagem. Os outros 97/290 (33,4%) pacientes não apresentaram alterações numéricas e/ou estruturais no exame de cariótipo com bandeamento GTG.

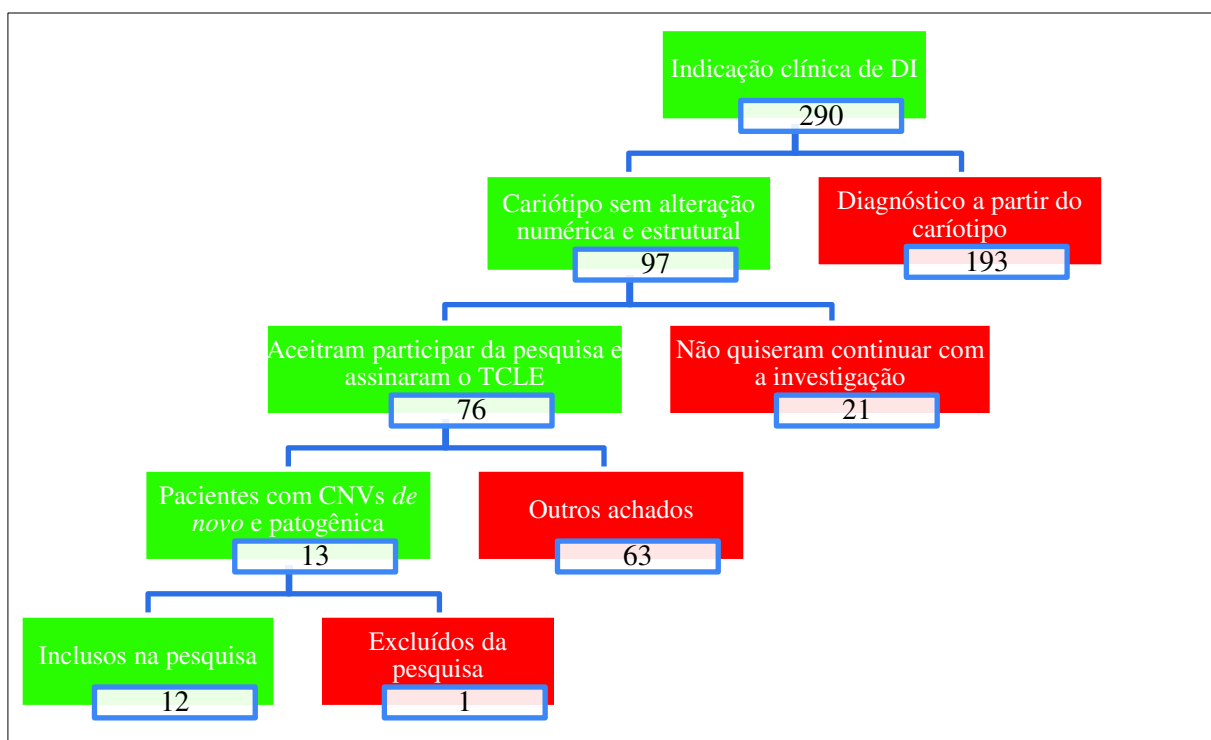


Figura 4. Fluxograma apresentando a dinâmica da seleção dos trios incluídos para o estudo da origem das CNVs de novo patogênicas em pacientes com indicação clínica de DI selecionados no período de 2013 a 2015.

4.2 Cariótipo por bandeamento GTG

Foram coletadas amostras de sangue periférico de 290 pacientes com DI usando seringas heparinizadas. As amostras foram submetidas à cultura de linfócitos T, em curto prazo, para a obtenção de metáfases. Após a cultura, a suspensão de células foi gotejada sobre vapor de banho-maria a 60°C, em lâminas limpas e desengorduradas. O bandeamento GTG com padrão de 550 bandas foi realizado nas lâminas envelhecidas por sete dias e posteriormente tratadas em solução de tripsina, diluída em tampão PBS e coradas em solução 4% de Giemsa, de acordo com protocolo modificado de Verma e Babu (1995). As metáfases foram capturadas usando microscopia de luz e analisadas com auxílio do *software* IKAROS® (*Metasystems Corporation*, EUA).

4.3 Extração e quantificação das amostras de DNA

O material biológico dos 76 pacientes e seus respectivos pais, conservado em EDTA, seguiu para processo de isolamento do DNA usando o kit de extração de DNA *Illustra Blood GenomicPrep® Mini Kit* (GE Healthcare Life Sciences, EUA). O DNA extraído foi

submetido à quantificação (ng/μL) em um espectrofotômetro NanoVue® Plus (GE Healthcare Life Sciences, EUA). Ambos os procedimentos foram executados de acordo com os protocolos sugeridos pelos fabricantes.

4.4 Análise cromossômica por microarranjo

A análise cromossômica por microarranjo (CMA) foi conduzida em um GeneChip® CytoScanHD™ (Affymetrix, Santa Clara, EUA), sendo uma matriz de genotipagem abrangente para o genoma humano. A matriz de genotipagem deste chip apresenta ampla cobertura do genoma e maior desempenho para a análise de alterações cromossômicas humanas, sendo capaz de detectar variações genéticas estruturais. O GeneChip® CytoScanHD™ apresenta sensibilidade de 99% na detecção de CNVs, na determinação de perda de heterozigose (LOH) e baixos níveis de mosaicismos. A matriz do chip possui cerca de 2,6 milhões de cópias de marcadores, incluindo aproximadamente 750 mil SNPs e cerca de 2 milhões de sondas não polimórficas. As sondas que integram o chip apresentavam tamanho de 25 pb e são redundantes.

4.5 Descrição do método de CMA

A metodologia consiste em um ensaio composto por múltiplas sondas de hibridização alelo-específicas que são complementares às regiões de SNP presentes na fração reduzida do genoma amplificado no ensaio. As sondas são constituídas de 25 oligonucleotídeos, com o SNP variável localizado no 13º nucleotídeo. Elas são redundantes e espalhadas ao longo do chip, para atenuar quaisquer efeitos da variação devido à localização física na matriz. Cada sonda tem uma localização fixa no arranjo do chip.

O procedimento se iniciou com uma digestão do DNA amostral com uma enzima de restrição (*NspI*) fornecida pelo fabricante. Este passo requeria aproximadamente 250 ng de DNA. Em seguida, o DNA digerido foi ligado a adaptadores específicos e, posteriormente, amplificados por PCR, com *primers* universais. O DNA amplificado e fragmentado por digestão enzimática foi marcado com biotina e hibridizado no CytoScanHD por 18 horas. Após a hibridização, os chips foram digitalizados no GeneChip® Scanner 3.000 7G (Affymetrix, Santa Clara, EUA). Ao final da digitalização, os sinais luminosos foram lidos e capturados pelo *software Array GeneChip® Command Console®* (AGCC®, Affymetrix, EUA)

gerando arquivos CEL, e as análises dos dados foram realizadas pelo *software Chromosome Analysis Suite*[®] 2.0 (ChAS[®], Affymetrix, EUA). Os arquivos CEL obtidos pelo escaneamento dos chips foram usados para estabelecer os genótipos dos SNPs contidos nas CNVs. As duas principais métricas de controle de qualidade do GeneChip[®] CytoScanHD[™] array foram *Median Absolute Pairwise Difference* (MAPD) e SNP-QC, escores que se aplicam para marcadores de número de cópias e SNP, respectivamente. Para a avaliação das CNVs e liberação do resultado de diagnóstico foram usados os parâmetros de MAPD ≤ 0.25 e de SNP-QC ≥ 15 . Para análise da origem da CNV foi usado o parâmetro do SNP-QC ≥ 15 .

4.6 Tratamento e análise dos dados obtidos na genotipagem dos SNPs

A análise dos resultados foi executada no *software* ChAS[®], que possibilitou investigar alterações estruturais ao longo do genoma, em função das como variações no número de cópias (CNV), é identificadas como ganhos e perdas de segmentos de DNA genômico presentes nas amostras estudadas de modo comparativo, facilitando as interpretações dos resultados. Foram fixados como filtros de diagnóstico: 50 marcadores para ganho e 25 marcadores para perda, distribuídos com uma média de distância entre os marcadores de ≤ 2.000 kb, limitando-se a fragmentos de tamanho ≥ 100 kb.

4.7 Análise das CNVs

Os seguintes critérios de seleção das CNVs foram utilizados:

- Foram incluídas nas análises CNVs que não apareceram nos bancos de dados de população saudável do CytoscanHD (Affymetrix, EUA) e do DGV (*The Database of Genomic Variants*).

- As CNVs foram classificadas, de acordo com o consenso de Miller et al. (2010) e Kearney et al. (2011), em: 1. Patogênicas – CNVs que se sobrepõe a regiões envolvidas em síndromes conhecidas, ou envolve genes importantes das vias do desenvolvimento neural, não ocorre em populações saudáveis, não é herdado de pais saudáveis; 2. Provavelmente Patogênicas - CNVs que não apresentam associação conhecida com a DI, mas envolvem genes que participam de vias do desenvolvimento neural sendo ou não herdada de pais saudáveis; não ocorre em populações saudáveis; 3. Significado Clínico Incerto – CNVs que

não apresentam genes envolvidos em vias do desenvolvimento neural, mas são CNVs que não foram reportadas em populações saudáveis, e que estão em regiões relacionadas com DI.

- As CNVs foram selecionadas de acordo com a origem em *de novo* ou herdadas.
- Para a análise da origem parental foram selecionadas as CNVs patogênicas *de novo*.

4.8 Análise da origem através da taxa de erro mendeliano

Para cada cromossomo que continha CNV *de novo* foram analisados erros no padrão de herança mendeliana por meio da análise de trio (filho-mãe-pai). Esta investigação consiste na detecção da taxa de erro mendeliano mediante comparação entre filho-mãe, filho-pai e entre o trio. Caso o cromossomo do filho não apresente um padrão de herança mendeliano esperado em comparação com o cromossomo de um dos genitores, significa que o filho não herdou o cromossomo daquele genitor e se neste cromossomo estiver presente um rearranjo, a origem deste rearranjo é do genitor que não apresentou o erro mendeliano.

Para obter a informação da origem parental da região do rearranjo, foi utilizada a ferramenta *Graphs Table*, presente no *software* ChAS[®], que usa os marcadores de SNP para construir os genótipos do filho, da mãe e do pai. Nesta ferramenta, os marcadores de SNP são representados pelo alelo A e alelo B, que simboliza duas possibilidades diferentes para cada marcador de SNP, sendo uma cópia originada da mãe e a outra do pai.

A análise da origem parental do cromossomo que apresenta a CNV *de novo* foi realizada utilizando a coordenada da região da CNV. Esta análise foi feita através da quantificação da taxa de erro mendeliano parental, onde o cromossomo parental que teve maior quantidade de erro não foi o cromossomo que originou a CNV *de novo*.

4.9 Análise de LCRs

A investigação da presença de LCRs circundantes as CNVs foi realizada através do banco de dados da *Genome Browser* da *University of California Santa Cruz* - UCSC (<http://www.genome.ucsc.edu>). Para definir as regiões flanqueadoras foi utilizado o número do cromossomo e a coordenada da CNV. Após a localização da região da CNV foi usado um aumento de 3 vezes do tamanho da CNV, sendo consideradas as LCRs com similaridade \geq

90% de suas sequências de nucleotídeos. Em seguida, utilizamos a plataforma *Galaxy* (<http://usegalaxy.org>) para gerar uma tabela com todas as LCRs observadas que flaqueavam cada CNV. Após gerar os dados de LCRs no *Galaxy*, foram selecionadas as LCRs que tinham a maior porcentagem de similaridade e a mesma direção.

5 – RESULTADOS E DISCUSSÃO

Após ter aplicado a técnica de cariótipo, 97 pacientes não apresentaram alterações numéricas e estruturais no cariótipo. Os pacientes, juntamente com seus pais e/ou responsáveis, foram convidados a participar desta pesquisa, e foram esclarecidos quanto à metodologia utilizada. Do total de convites, 76/97 (78,3%) trios aceitaram participar da avaliação citogenômica pelo teste CMA e assinaram o TCLE. Após a realização do teste CMA, foram detectadas CNVs *de novo* patogênicas em 17,1% (13/76) dos pacientes com DI que possuem cariótipo sem alteração. A taxa de diagnóstico da nossa análise está de acordo com o reportado na literatura que compreende 15-20% de diagnóstico em indivíduos com DI (MILLER et al. 2010; BORLOT et al. 2017; NOWAKOWSKA, 2017).

Para a análise da origem parental foram selecionados 13 trios cuja detecção do CMA foi positiva para CNVs *de novo* patogênicas. Entretanto, apenas 12 trios foram incluídos para a análise da origem parental das CNVs. A exclusão de um dos trios foi devida os dados de marcadores de origem paterna e materna terem sido iguais, portanto não foi possível inferir a origem parental da CNV.

Os 12 probandos tinham a faixa etária entre 1 a 22 anos de idade, sendo que a média de idade foi de 10,2 anos. Entre estes pacientes, 8/12 (66,7%) são do sexo feminino e 4/12 (33,3%) do sexo masculino. Os médicos assistentes destes pacientes relataram outros sinais fenotípicos clinicamente observados além da deficiência intelectual, que incluíram atraso no desenvolvimento global, baixa estatura, microcefalia, múltiplos estigmas, alteração no eletroencefalograma, anomalias congênitas múltiplas e desordem do comportamento. A idade dos pais no momento da concepção foi heterogênea. As mães tinham entre 15 a 37 anos, tendo a média de idade de 27,7 anos. Os pais tinham entre 19 a 49 anos de idade, apresentando uma média superior à das mães, sendo de 32,2 anos de idade.

Pela CMA, foram detectadas 15 CNVs *de novo* patogênicas, em 12 pacientes, sendo que 10/15 foram CNVs de perda (66,7%) e 5/15 foram CNVs de ganho (33,3%). As CNVs *de novo* apresentaram tamanho bastante variado. A média de tamanho das CNVs de perda foi de 5 Mb, sendo que a maior CNV de perda apresentou 15,39 Mb e a menor 0,51 Mb. A média de tamanho das CNVs de ganho foi 18,9 Mb, sendo que a maior foi de 59,51 Mb e a menor foi de 1,43 Mb. A Figura 5 representa a distribuição das CNVs de ganho e perda em função do tamanho.

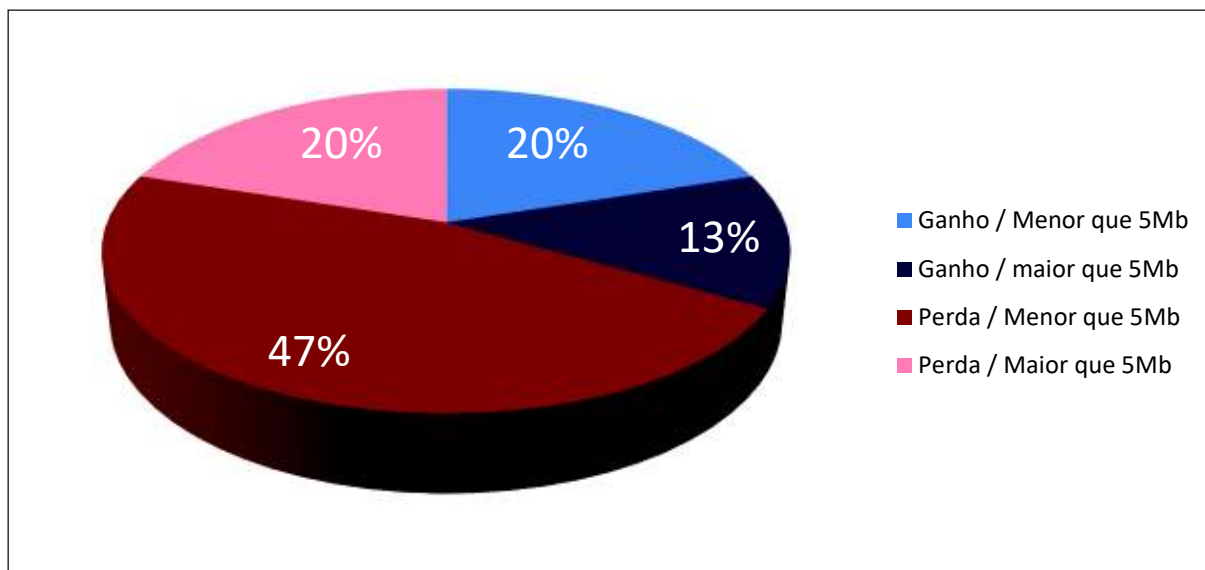


Figura 5. Percentual das CNVs de perda (vermelho) e ganho (azul) distribuídas pelo tamanho do segmento envolvido no rearranjo, menor ou maior que 5 Mb, observadas em pacientes com deficiência intelectual.

No presente estudo, foi observado uma frequência maior de perda genômica associada a CNVs *de novo* patogênicas, sendo 67% de CNVs de perda e 33% de perda, corroborando com outros estudos que também analisaram pacientes com DI. No estudo de Ma et al. (2017) foi detectado 79,3% de CNVs de perda e 38,7 CNVs de ganho. O trabalho de Boggula et al. (2016) reportou que 92,8% das CNVs patogênicas são de perda e 7,1% são de ganho. Adicionalmente, Tao et al. (2014) identificaram 65,9% de CNVs de perda e 34,1% de ganho. Estes achados podem ser explicados devido as CNVs de perda serem geralmente mais patogênicas que as CNVs de ganho, e como o estudo foi realizado em pacientes com DI, um fenótipo relacionado com desordem genômica, espera-se um maior número de CNV de perda (LEE et al. 2007; PRATTER-SANTOS et al., 2016).

A variação no tamanho das CNVs observadas no presente estudo está de acordo com a variação reportada em outros trabalhos (TAO et al. 2014; MA et al. 2017). Algumas regiões do genoma suportam CNVs maiores e outras são mais sensíveis à variação no tamanho. Normalmente, quanto maior uma CNV, mais genes ela irá envolver, sendo assim, as CNVs maiores tem o potencial patogênico aumentado quando comparado as CNVs menores. No entanto deve-se observar a região na qual as CNVs se encontram. Portanto, se avaliar a presença de genes sensíveis à dosagem e genes importantes para o desenvolvimento neurológico na CNV. Normalmente, as pequenas CNVs patogênicas de perda incluem alguns genes haploinsuficientes (LEE et al. 2007). Apesar do número de CNVs de ganho ter sido

menor, a média do tamanho das CNVs de ganho (18,9 Mb) foi superior em relação a CNV de perda (5 Mb) no presente estudo.

As CNVs *de novo* patogênicas detectadas nos pacientes foram apresentadas de forma esquemática no idiograma da Figura 6, destacando-se o tipo da CNV e sua localização cromossômica. As CNVs reportadas neste estudo estão presentes em dez cromossomos (cromossomos 1, 4, 6, 7, 9, 13, 17, 18, 22 e X), sendo que três cromossomos apresentaram duas CNVs (cromossomos 1, 4 e X) e o cromossomo 17 apresentou três CNVs.

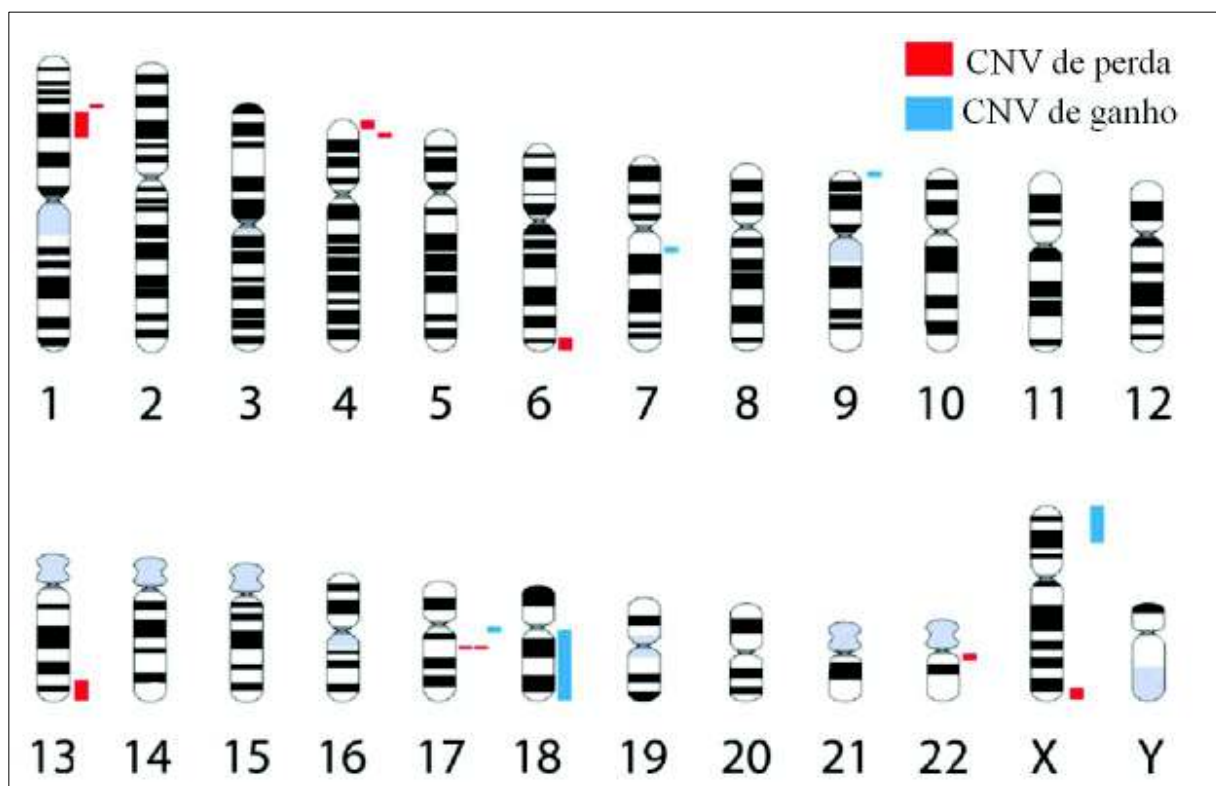


Figura 6. Representação esquemática do idiograma humano representado as CNVs *de novo* patogênicas identificadas pelo presente estudo. As faixas em vermelho ao lado dos cromossomos representam as CNVs de perda e as faixas em azul representam as CNVs de ganho. Em relação ao tamanho das CNVs: a primeira CNV 1p31.3p31.1 é de perda e tem 10.89 Mb, a CNV de perda no 1p32.3 tem 1,59 Mb; a CNV no 4p16.3 possui 3,86 Mb e a CNV no 4p16.3p16.2 possui 1,55 Mb; a CNV de perda no 6q26q27 possui 8,2 Mb; a CNV no 7q11.23 possui 1,43Mb; a CNV no 9p24.3p24.2 possui 4.21 Mb; a CNV no 13q32.3q34 possui 15,39Mb; a CNV no 17q21.31 possui 0,56 Mb (primeira), a segunda CNV no 17q21.31 possui 0,51 Mb e a terceira CNV no 17p11.2 possui 3,68Mb; a CNV de ganho no 18q11.1q23 é uma trissomia parcial, com 59,41Mb; a CNV de perda no 22q11.21 possui 2,88 Mb; a primeira CNV no cromossomo Xq27.3q28 é de perda e possui 4,18Mb e a segunda no Xp22.33p21.3 é de ganho 25,72Mb.

No estudo de Tao et al. (2014), que estudaram também pacientes com DI, foi detectado uma grande variação no tamanho das CNVs, sendo que as duas das maiores CNVs foram de ganho, sendo uma de 31,6 Mb localizada em 18q12.1q22.1 e a outra de 83,9 Mb em Xq13.1q28. E ainda CNV relataram uma grande de perda no 39,1 Mb em Xp22.33–p11.4. No nosso estudo, também foram encontradas duas CNVs grandes de ganho nos mesmos cromossomos envolvendo a mesmas regiões. Por outro lado, a maior CNV de perda observada no presente estudo foi de 15,39 Mb e, 13q32.3q34. A trissomia e monossomia completa do cromossomo X normalmente são compatíveis com a vida em humanos. Estas variações no cromossomo X podem ser toleradas devido aos processos de ativação e inativação de genes próprios deste cromossomo. No entanto, dependendo dos pontos de quebras associados às CNVs de perda e ganho no cromossomo X também podem ser objetos de mecanismos patogênicos que incluem os efeitos de dosagem gênica, de dano direto, de *imprinting* ou de posição subjacentes observadas nas situações de desequilíbrio genômico.

Já é bem esclarecido que as CNVs *de novo* são transmitidas pelos gametas dos pais e a formação destas CNVs ocorre no processo de gametogênese, sendo produtos de reparo de DNA, erro de replicação, recombinação homóloga e erros de segregação cromossômica. Portanto, a instabilidade genômica do tecido germinativo é essencial para a aptidão da reprodução e o desenvolvimento saudável da prole. Apesar desta informação ser clara, ainda pouco se sabe e poucos estudos correlacionam a formação das CNVs *de novo* com a origem parental (MA et al. 2017).

Os processos mutacionais e a contribuição dos genitores é uma importante área para o estudo da formação de novas variações no genoma. E os indivíduos do sexo masculino contribuem para a variação genômica da sua prole, na maior, parte com mutações pontuais devido às inúmeras replicações do material genético para as divisões celulares pré-meióticas que são realizadas pelas espermatogônias. Este contexto é diferente nas mulheres nas quais as células germinativas ainda no período intrauterino do conceito feminino inicia o ciclo meiótico e o interrompem no final da meiose I. O reinício deste ciclo meiótico acontece na puberdade e conclui com a fecundação/fertilização. Neste caso, postula-se que as mães contribuem mais que os pais para as taxas de aneuploidias devido a não disjunção cromossômica dos cromossomos homólogos durante a primeira divisão meiótica, tendo um aumento na frequência de erros de segregação cromossômica com o avançar da idade (1000 GENOMES PROJECT CONSORTIUM, 2010).

Apesar de haver poucos estudos investigando a origem parental das CNVs *de novo*, a tentativa de desvendar a contribuição dos pais para os rearranjos iniciou a muitos anos, quando passou a se investigar citogeneticamente a origem parental dos cromossomos com anomalias estruturais não balanceadas. Um dos estudos pioneiros na busca da origem parental das CNVs *de novo* foi publicado em 1980 por Charmberlin e Mageni, que analisaram 19 pacientes com rearranjos citogenéticos estruturais *de novo*. Usando diferentes colorações com bandeamento para detectar heteromorfismos nos cromossomos como marcador hereditário, foi possível inferir a origem parental dos 19 rearranjos. Os autores reportaram que dos 19 rearranjos, 9 não eram translocações Robertsonianas. Destes 9 rearranjos, 8 eram de origem paterna.

Análise de citogenética bandeamento G de alta resolução (600-700 bandas) e a FISH permitiram que Thomas et al. (2006) reportassem a origem parental de 39 deleções terminais e 35 deleções intersticiais. Desses 74 rearranjos, 84% eram de origem paterna. Nesse mesmo estudo, foi analisado também 12 duplicações, sendo que 7/12 (58%) eram de origem paterna e 5/12 (42%) de origem materna.

No presente estudo, usando a metodologia de CMA, detectou-se a origem parental das 15 CNVs *de novo* patogênicas. Em 10 das CNVs *de novo* patogênicas de perda, 7/10 são de origem materna (70%) e 3/10 são de origem paterna (30%). Na análise das CNVs *de novo* patogênicas de ganho, 3/5 são de origem paterna (60%) e 2/5 (40%) são de origem materna.

A distribuição das CNVs quanto ao seu tamanho e a origem parental está representada na Figura 7-a e 7-b.

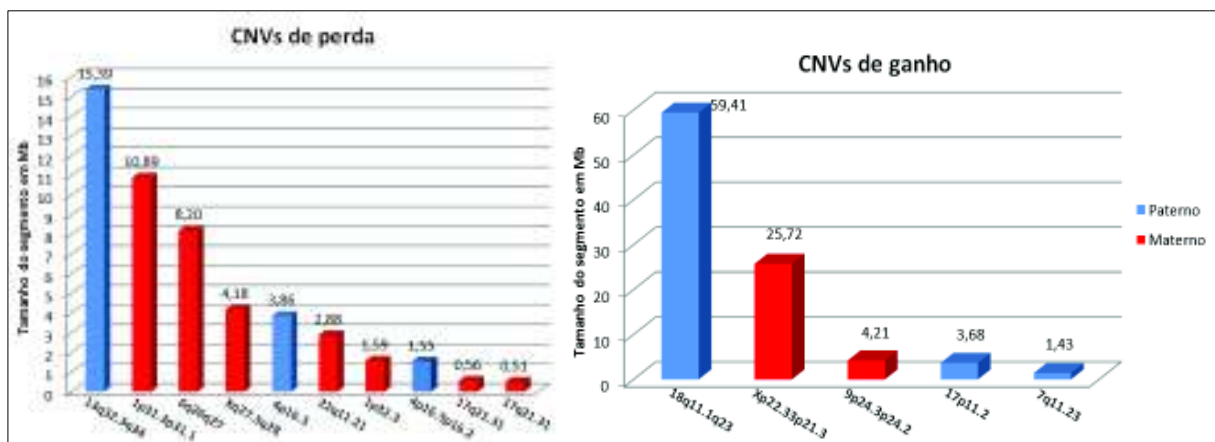


Figura 7-a e 7-b. Representação da origem parental das CNVs de perda e de ganho em relação ao tamanho em Mb. No eixo X é representado a região cromossômica de cada CNV e no eixo Y o tamanho do segmento da CNV em Mb. Cada coluna é uma CNV. As colunas em azul representam que a origem parental da CNV é paterna e as colunas em vermelho representam que a origem parental da CNV é materna.

Hehir-Kwa et al. (2011) fizeram a análise da origem parental de 118 CNVs *de novo* usando a metodologia de CMA com uma matriz de SNP de 250 K (*Affymetrix*, Santa Clara, Califórnia, EUA). O estudo apresentou 76,3% das CNVs de origem paterna (90/118) enquanto as de origem materna representaram 23,7% (28/118). Naqueles estudos, além de detectar que as CNVs *de novo* eram em sua maioria de origem paterna, foi detectado também o incremento de CNV com o aumento da idade dos pais. Estes dados não corroboram com os achados do nosso estudo que demonstrou que 60% das CNVs *de novo* patogênicas foram de origem materna. É importante ressaltar que o tamanho amostral do presente estudo possa contribuir randomicamente para este efeito, que corresponde apenas a um viés amostral.

No estudo de Delio et al. (2013) foi analisado a origem parental de 389 indivíduos com perda no 22q11, condizente com a síndrome de Velocardiofacial e DiGeorge. Neste estudo eles detectaram que em 219 indivíduos (56%) a origem parental da CNV é materna e em 170 a origem é paterna. O viés da origem da CNV de perda na região 22q11 foi estatisticamente significativa ($p=0,0151$). Ao combinar com estudos anteriores, eles observaram que 465 (57%) tinham origem materna e 345 (43%) tinham origem paterna. No nosso estudo, a CNV de perda no 22q11.21 também teve a origem parental materna.

Hehir-Kwa et al. (2011) afirmaram que além da contribuição de mutações pontuais, as CNVs *de novo* em DI são de grande parte de origem paterna, além também de rearranjos cromossômicos visíveis a microscopia (THOMAS et al. 2006). Chamberlin e Magenis (1980), com dados que também corroboram com estes achados, ainda em 1980, afirmaram que o estilo de vida dos genitores deve ser investigado, devido estes rearranjos cromossômicos serem formados por ruptura e junção, assim, agentes que provocam quebras na molécula de DNA, como produtos químicos e radiação, deve ser investigado quanto a formação de novos rearranjos.

Os pais dos pacientes avaliados no presente estudo tinham idades muito heterogêneas, portanto, não conseguimos associar os achados da origem cromossômica das CNVs com a idade parental. Na Tabela 1 reunimos as informações gerais das idades dos pacientes, a idade parental no momento da concepção e as características clínicas. Na mesma tabela encontram também informações moleculares sobre a CNV: tamanho, número de marcadores, os genes mórbidos envolvidos pela CNV e a origem parental da CNV. Neste estudo com 12 trios, três casais possuem a mesma idade (casos 023, 036 e 042). Nos casos 013, 026 e 033, os

genitores responsáveis pela transmissão da CNV são mais velhos que suas parceiras, que curiosamente, nos três casos a origem parental é paterna.

Tabela 1. Características clínicas e moleculares dos pacientes que apresentaram CNV *de novo* patogênica e análise da origem da CNV

Tipo de CNV	Caso	Idade (anos)	Sexo	Características Clínicas*	Idade M**	Idade P**	Nomenclatura microarranjo***	Tamanho (Mb)	Nº Marcadores	Genes mórbidos selecionados pelo OMIM	Origem da CNV
	006	11	F	DI	25	28	arr[GRCh37] Xq27.3q28(144.580.614-148.757.072)x1 dn	4,18	10150	CXorf1/TMEM257, FMRI, TMEM185A, IDS	Materna
	011	8	F	DI,ME,DC	25	29	arr[GRCh37] Ip31.3p31.1(68.693.129-79.580.916)x1 dn	10,89	10080	AK5	Materna
	023	1	F	ADG,BE,ME,A CM	31	31	arr[GRCh37] 4p16.3(68.345-3.926.333)x1 dn	3,86	3848	ZNF141, IDUA, LETM1, WHSC1, NELFA/WHSC2	Paterna
	026	6	F	ADG, ME, ACM	32	38	arr[GRCh37] 4p16.3p16.2(4.177.795-5.724.404)x1 dn	1,55	1884	MSX1, STK32B, C4orf6, EVC2, EVC	Paterna
	034	4	M	ADG	29	47	arr[GRCh37] 1p32.3(53.894.316-55.487.208)x1 dn	1,59	974	DHCR24	Materna
	036	15	M	DI	23	23	arr[GRCh37] 22q11.21(18.916.842-21.800.797)x1 dn	2,88	3788	PRODH, DGCR5, DGCR9, DGCR10, DGCR2, DGCR11, DGCR14, GSC2, SLC25A1, CLTCL1, HIRA, MRPL40, UFD1L, CLDN5, SEPT5, GP1BB, TBX1, GNB1L, COMT, ARVCF, DGCR8, DGCR6L, ZNF74, SCARF2, MED15, SNAP29, CRKL, LZTR1, SLC7A4, HIC2	Materna
	042	6	F	DI,M	28	28	arr[GRCh37] 6q26q27(162.708.065-170.919.482)x1 dn	8,2	7709	BRP44L, RNASET2, TBP	Materna
	043	18	F	DI	22	26	arr[GRCh37] 17q21.31(43.648.662-44.212.416)x1 dn	0,56	924	KANSL1	Materna
	002	11	M	DI	37	25	arr[GRCh37] 17p11.2(16.769.800-20.446.820)x3 dn	3,68	4151	COPS3, SMCR9, RAI1, SMCR5, TOM1L2, LRRc48, ATPAF2, DRG2, MYO15A, ALKBH5, FLII, SMCR8, SHMT1, USP32P2, CCDC144B, B9DI, MFAP4, RNF112	Paterna
	007	4	F	ADG, ME, ACM	33	47	arr[GRCh37] 18q11.1q23(18.608.373-78.014.123)x2-3 dn	59,41	53197	Trissomia parcial 18q	Paterna
	013	16	M	DI	15	19	arr[GRCh37] Xp22.3p21.3(168.546-25.887.307)x3 dn	25,72	31456	NLGN4X, API52, NHS, CDKL5, RPS6KA3, MBTPS2, SMS, ARX	Materna
	023	1	F	ADG, BE, ME, ACM	31	31	arr[GRCh37] 7q11.23(72.718.277-74.147.166)x3 dn	1,43	1540	TRIM50, FKBP6, FZD9, BAZ1B, BCL7B, TBL2, MLXIPL, VPS37D, WBSCR27, WBSCR28, ELN, LIMK1, EIF4H, LAT2, CLIP2, GTF2IRD1, GTF2I	Paterna
	023	1	F	ADG, BE, ME, ACM	31	31	arr[GRCh37] 9p24.3p24.2(203.861-4.416.073)x3 dn	4,21	5548	DOCK8, KANK1, SMARCA2, VLDLR	Materna

* DI=Deficiência Intelectual; ADG= Atraso no Desenvolvimento Global; M=Microcefalia; ME=Múltiplos estigmas; AE=Alteração no Eletroencefalograma; DC=Desordem do comportamento; ACM=Anomalias Congênicas Múltiplas; BE=Baixa Estatura.

** Idade M: Idade materna na concepção; Idade P: Idade Paterna na concepção.

*** ISCN 2016.

Foi realizada análise da presença de LCRs circundantes ($\geq 90\%$ de similaridade) às CNVs *de novo* através do *Genome Browser* (<https://genome.ucsc.edu/>) para inferir possível mecanismo responsável pela formação das CNVs. As LCRs foram encontradas flanqueando 7 CNVs (46,7%), no presente estudo (Tabela 2).

Tabela 2. Análise da presença de LCRs flanqueando as CNVs.

Tipo de CNV	Casos	Tamanho da CNV (Mb)	Coordenada da CNV GRCh37/hg19	Origem da CNV	Similaridade das LCRs (%)	Orientação das LCRs
Ganho	002	3,68	chr17:16.769.800-20.446.820	Paterna	98,8	+
	007	25,72	chrX:168.546-25.887.307	Materna	100,0	+
	013	1,43	chr7:72.718.277-74.147.166	Paterna	99,8	-
Perda	011	10,89	1p31.3p31.1(68.693.129-79.580.916)	Materna	91,3	+
	036	2,88	22q11.21(18.916.842-21.800797)	Materna	98,2	-
	043	0,56	17q21.31(43.648.662-44.212.416)	Materna	98,3	-
	063	0,51	17q21.31(43.703.801-44.212.416)	Materna	98,3	-

No estudo de Ma et al. (2017), a análise da presença de LCRs flanqueando as CNVs estudadas detectou 57/87 CNVs recorrentes (65,5%), uma porcentagem superior ao que encontramos. Em contrapartida, o estudo de Hehir-Kwa et al. (2011) identificou que das 118 CNVs identificadas 25 delas estavam circundadas por LCRs, representando 21,2%, valor inferior ao observado no presente estudo. Por meio da análise da presença de LCRs circundantes às CNVs pode-se afirmar que 46,7% das CNVs deste estudo são CNVs recorrentes e são formadas mediante o processo de NAHR, sendo um dos primeiros mecanismos identificados como responsáveis pela formação de desordens genômicas (LUPSKI, 1998; STANKIEWICZ; LUPSKI, 2002).

As CNVs que não estavam circundadas por LCR com similaridade maior que 90% representaram 53% do total de CNVs estudadas (8/15). É provável que estas CNVs não

recorrentes foram formadas por mecanismos diferentes de NAHR, que não necessitam das sequências de LCRs, como o mecanismo de NHEJ e os mecanismos baseados em replicação (RBMs) destacando-se a replicação induzida por ruptura mediada por microhomologia (MMBIR) e Parada da forquilha de replicação e mudança de molde de DNA (FoSTeS).

Os achados do trabalho de Hehir-Kwa et al. (2011) sugerem que as CNVs de origem paterna geralmente não são recorrentes, sendo geradas por mecanismos como NHEJ e RBMs, e além dessas CNVs serem raras, elas estão espalhadas por todo o genoma. Esse dado corrobora com o que foi observado no presente estudo, onde do total de 6 CNVs paternas observadas 4 delas são CNVs não recorrentes, representando 66,7% das CNVs paternas identificadas. Isto explicaria que a maioria das CNVs de origem parental paterna são formadas no processo de replicação do DNA, que ocorre frequentemente nas espermatogônias que realizam inúmeras divisões mitóticas para aumentar a produção de gametas.

Segundo Carvalho e Lupski (2016), as três principais características de LCR que contribuem positivamente para a instabilidade genômica favorecendo o processo de NAHR é: (1) LCRs maior que 10 Kb, consideradas LCRs grandes, (2) posição a uma distância aproximada de 10 Mb e (3) similaridade de sequência de 97%. Neste atual estudo foi analisada a similaridade da sequência das LCRs circundantes, 6/7 CNVs (85,7%) possuem LCRs circundantes com similaridade das sequências maiores que 98%.

Em geral as doenças do desenvolvimento neurológico, em especial a DI, muitas vezes estão correlacionadas com CNVs recorrentes raras de pontos específicos do cromossomo, sendo que muitas dessas CNVs possuem penetrância incompleta e a expressividade influenciada por outros fatores genéticos e ambientais (TORRES et al. 2016; CARVALHO; LUPSKI, 2016). Diferentemente do que já foi observado em outros estudos, o presente estudo identificou que dentre as 15 CNVs patogênicas *de novo* a maioria é de origem materna e foi observado também um maior conjunto de CNVs não recorrentes. Em contrapartida, a maioria das CNVs não recorrentes foram de origem paterna. Portanto, diante do que foi identificado estudos de origem cromossômica das variantes estruturais e seus mecanismos de formação devem ser ampliados e são importantes para a compreensão do papel dos eventos não relacionados ao NAHR, ou seja, não dependentes da presença de LCRs, na geração de CNVs *de novo* na DI.

7 – CONCLUSÃO

Nosso estudo investigou a origem parental das CNVs *de novo* patogênicas para compreender a contribuição de cada genitor para formação de novas CNVs de interesse clínico. Ao término do estudo podemos concluir que:

- No chort de pacientes para as quais foi feito a CMA, 17,1% apresentaram CNVs *de novo* patogênicas. Entre as CNVs *de novo* patogênicas detectadas, 66,7% foram de perda e 33,3% são de ganho, demonstrando que as CNVs de perdas tendem a ser mais patogênicas e mais presentes em estudos com pacientes com DI.
- O incremento do diagnóstico genético da DI com o uso da metodologia de CMA observado no presente estudo, foi de 4,5%.
- No presente estudo, a origem parental de 60% das CNVs foi materna e 40% e paterna. Ao analisar as CNVs de perda, 70% delas eram de origem materna e 30% de origem paterna. E as CNVs de ganho, 60% são de origem paterna e 40% são de origem materna.
- Com análise da presença de LCRs circundando as CNVs, detectamos que 46,7% das CNVs *de novo* patogênicas são recorrentes e o mecanismo de formação foi NAHR. Sendo que 71,4% das CNVs recorrentes são de origem materna.

8. REFERÊNCIAS BIBLIOGRÁFICAS

1000 GENOMES PROJECT CONSORTIUM. 2010. **A map of human genome variation from population-scale sequencing.** *Nature*, v. 467, n. 7319, p. 1061.

ADAMS, D.; OLIVER, C. 2011. **The expression and assessment of emotions and internal states in individuals with severe or profound intellectual disabilities.** *Clinical psychology review*, v. 31, n. 3, p. 293-306.

ARLT, M. F.; OZDEMIR, A. C.; BIRKELAND, S. R.; et al. 2011. **Comparison of constitutional and replication stress-induced genome structural variation by SNP array and mate-pair sequencing.** *Genetics*, v. 187, n. 3, p. 675-683.

ARLT, M. F.; RAJENDRAN, S.; BIRKELAND, S. R.; ET al. 2014. **Copy number variants are produced in response to low-dose ionizing radiation in cultured cells.** *Environmental and molecular mutagenesis*, v. 55, n. 2, p. 103-113.

BALLIF, B. C.; SULPIZIO, S. G.; LLOYD, R. M.; et al. 2007. **The clinical utility of enhanced sub-telomeric coverage in array CGH.** *Am J Med Genet.* v. 143, n. 16, p. 1850-1857.

BAUMAN, J. G.; WIEGANT, J.; BORST, P.; et al.; 1980. **A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA.** *Exp. Cell Res.* v. 128, n. 2, p. 485-490.

BINSBERGEN, V. E. 2011. **Origins and breakpoint analyses of copy number variations: up close and personal.** *Cytogenetic and genome research*, v. 135, n. 3-4, p. 271-276.

BLOOM, R. J.; KÄHLER, A. K.; COLLINS, A. L.; et al. 2013. **Comprehensive analysis of copy number variation in monozygotic twins discordant for bipolar disorder or schizophrenia.** *Schizophrenia research*, v. 146, n. 1, p. 289-290.

BOGGULA, V. R.; AGARWAL, M.; KUMAR, K.; et al. 2015. **Recurrent benign copy number variants & issues in interpretation of variants of unknown significance identified by cytogenetic microarray in Indian patients with intellectual disability.** *The Indian journal of medical research*, v. 142, n. 6, p. 699.

BORLOT, F. REGAN, B. M., BASSETT, A. S.; ET al. 2017. **Prevalence of Pathogenic Copy Number Variation in Adults With Pediatric-Onset Epilepsy and Intellectual Disability.** *JAMA neurology*, v. 74, n. 11, p. 1301-1311.

BUNTING, S. F.; NUSSENZWEIG, A. 2013. **End-joining, translocations and cancer.** *Nature reviews Cancer*, v. 13, n. 7, p. 443-454.

CAMPBELL, C. D.; SAMPAS, N.; TSALENKO, A.; et al. 2011. **Population-genetic properties of differentiated human copy-number polymorphisms.** *The American Journal of Human Genetics*, v. 88, n. 3, p. 317-332,.

CARVALHO, C. M. B.; LUPSKI, J. R. 2016. **Mechanisms underlying structural variant formation in genomic disorders.** *Nature Reviews Genetics*, v. 17, n. 4, p. 224-238.

- CHAMBERLIN, J.; MAGENIS, R. E. 1980. **Parental origin of de novo chromosome rearrangements.** *Human genetics*, v. 53, n. 3, p. 343-347.
- CONRAD, D. F.; BIRD, C.; BLACKBURNE, B.; et al. 2010. **Mutation spectrum revealed by breakpoint sequencing of human germline CNVs.** *Nature genetics*, v. 42, n. 5, p. 385-391.
- COSTANTINO, L.; SOTIRIOU, S. K.; RANTALA, J. K.; et al. 2014. **Break-induced replication repair of damaged forks induces genomic duplications in human cells.** *Science*. p. 1243211.
- DAVIS, A. J.; CHEN, D. J. 2013. **DNA double strand break repair via non-homologous end-joining.** *Translational cancer research*, v. 2, n. 3, p. 130.
- DELIO, M.; GUO, T.; MCDONALD-MCGINN, D.; et al. 2013. **Enhanced maternal origin of the 22q11. 2 deletion in velocardiofacial and DiGeorge syndromes.** *The American Journal of Human Genetics*, v. 92, n. 3, p. 439-447.
- ESCARAMÍ, G.; DOCAMPO, E.; RABIONET, R. 2015. **A decade of structural variants: description, history and methods to detect structural variation.** *Briefings in functional genomics*, v. 14, n. 5, p. 305-314.
- GIRIRAJAN, S.; CAMPEBELL, C. D.; EICHLER, E. E.; 2011. **human copy number variation and complex genetic disease.** *Annu Rev Genet.* , v. 45, p. 203-226.
- GU W.; ZHANG, F.; LUPSKI, J. R. 2008. **Mechanisms for human genomic rearrangements.** *Pathogenetics*, v. 1, n. 1, p. 4.
- HAREL, T.; LUPSKI, J. R. 2017. **Genomic disorders 20 years on-mechanisms for clinical manifestations.** *Clinical genetics*. DOI: 10.1111/cge.13146
- HEHIR-KWA, J. Y.; RODRÍGUEZ-SANTIAGO, B.; VISSERS, L. E.; et al. 2011. **De novo copy number variants associated with intellectual disability have a paternal origin and age bias.** *Journal of medical genetics*, v. 48, n. 11, p. 776-778.
- HOANG, M. L.; TAN, F. J.; LAI, D. C.; et al. 2010. **Competitive repair by naturally dispersed repetitive DNA during non-allelic homologous recombination.** *PLoS genetics*, v. 6, n. 12, p. e1001228.
- HOCHSTENBACH, R.; BINSBERGEN, E. V.; ENGELEN J.; et al. 2009. **Array analysis and karyotyping: workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands.** *European journal of medical genetics*, v. 52, n. 4, p. 161-169.
- HOOK, E. B. 1977. **Exclusion of chromosomal mosaicism: tables of 90%, 95% and 99% confidence limits and comments on use.** *Am J Hum Genet.* v. 29, n. 1, p. 94.
- HU, L.; YAO, X.; HUANG, HU.; et al. 2018. **Clinical significance of germline copy number variation in susceptibility of human diseases.** *Journal of Genetics and Genomics*.
- KALLIONIEMI, A.; KALLIONIEMI, O. P.; SUDAR, D.; et al. 1992. **Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors.** *Science*, v. 258, n. 5083, p. 818-821,

- KE. X.; LIU, J. 2015. **Deficiência Intelectual**. In Rey JM (ed), **IACAPAP e-Textbook of Child and Adolescent Mental Health**. (edição em Português; Dias Silva F, ed). *Genebra: International Association for Child and Adolescent Psychiatry and Allied Professions*
- KEARNEY, H. M.; THORLAND, E. C.; BROWN K. K.; et al. 2011. **American College of Medical Genetics standards and guidelines for interpretation and reporting of postnatal constitutional copy number variants**. *Genetics in Medicine*, v. 13, n. 7, p. 680.
- LANDER, E. S. 2011. **Initial impact of the sequencing of the human genome**. *Nature*. v.470, n.7333, p.187-97, Feb 10.
- LEE, C.; IAFRATE, A. J.; BROTHMAN, A. R.; 2007. **Copy number variations and clinical cytogenetic diagnosis of constitutional disorders**. *Nature genetics*, v. 39, p. S48-S54.
- LEJEUNE, J.; GAUTIER, M.; TURPIN, R. 1959. **Etude des chromosomes somatiques de neuf enfants mongoliens**. [Study of somatic chromosomes from 9 mongoloid children]. *C R Hebd Seances Acad Sci*, v.248, n.11, p.1721-2, Mar 16.
- LEJEUNE, J.; LAFOURCADE, J.; BERGER, R.; et al. 1963. **Trois cas de délétion partielle du bras court d'un chromosome 5**. [3 Cases of Partial Deletion of the Short Arm of a 5 Chromosome]. *C R Hebd Seances Acad Sci*, v.257, p.3098-102, Nov 18.
- LIU, P.; LACARIA, M.; ZHANG, F.; et al. 2011. **Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over**. *The American Journal of Human Genetics*, v. 89, n. 4, p. 580-588.
- LUCKASSON, R.; BORTHWICK-DUFFY, S.; BUNTINX, W. H.; et al. 2002. **Mental retardation: Definition, classification, and systems of supports**. *American Association on Mental Retardation*. 10TH ED. Washington.
- LUPSKI J. R. 1998. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet*. v. 14, n. 10, p. 417-422.
- LUPSKI J. R.; BELMONT J. W.; BOERWINKLE E.; et al. 2011. **Clan genomics and the complex architecture of human disease**. *Cell*; v. 147, n. 1, p. 32-43.
- MA, R.; DENG, L.; XIA, Y.; et al. 2017. **A clear bias in parental origin of de novo pathogenic CNVs related to intellectual disability, developmental delay and multiple congenital anomalies**. *Scientific Reports*, v. 7. p. 44446.
- MALUF, S.W.; RIEGEL, M.; 2011. **Citogenética humana** Artmed, São Paulo.
- MATTEVI, M.S.; MIRANDA, J.A. 2011. **História da citogenética clínica**. In: MALUF, S.W. et al. (Org.). *Citogenética Humana*. Porto Alegre: Artmed,. p. 11-16.
- MAULIK, P. K.; MASCARENHAS, M. N.; MATHERS, C. D.; et al. 2011. **Prevalence of intellectual disability: a meta-analysis of population-based studies**. *Res. Dev. Disabil*. v. 32, n. 2, p. 419-436.
- MEFFORD, H. C.; BATSHAW, M. L.; HOFFMAN, E. P.; 2012. **Genomics, intellectual disability, and autism**. *N. Engl. J. Med*. v. 366, n. 8, p. 733-743.

- MILLER, D. T.; ADAM, M. P.; ARADHYA S.; et al. 2010. **Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies.** *Am J Hum Genet.* v. 86, n. 5, p. 749-764.
- MILLS, R. E.; WALTER, K.; STEWART, C.; et al. 2011. **Mapping copy number variation by population-scale genome sequencing.** *Nature,* v. 470, n. 7332, p. 59.
- NG, S.B.; BUCKINGHAM, K. J.; LEE, C.; et al. 2010. **Exome sequencing identifies the cause of a mendelian disorder.** *Nature Genetics.* v.42, p.30 – 35.
- NOWAKOWSKA, B. 2017. **Clinical interpretation of copy number variants in the human genome.** *Journal of applied genetics,* v. 58, n. 4, p. 449-457.
- PANNUNZIO, N. R.; WATANABE, G.; LIEBER, M. R.; 2017. **Nonhomologous DNA End Joining for Repair of DNA Double-Strand Breaks.** *Journal of Biological Chemistry,* p. jbc.TM117. 000374.
- PASQUIER, L.; FRADIN, M.; CHÉROT, E.; et al. 2016. **Karyotype is not dead (yet)!. European journal of medical genetics,** v. 59, n. 1, p. 11-15.
- PINKEL D.; SEGRAVES, R.; SUDAR, D.; et al. 1998. **High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays.** *Nat Genet.* v. 20, n. 2, p. 207.
- POLLACK J. R.; PEROU, C. M.; ALIZADEH, A. A.; et al. 1999. **Genome-wide analysis of DNA copy-number changes using cDNA microarrays.** *Nat Genet.* v. 23, n. 1, p. 41.
- PRATTE-SANTOS, R.; RIBEIRO, K. H.; SANTOS, T. A.; et al. 2016. **Análise de anomalias cromossômicas por CGH-array em pacientes com dismorfias e deficiência intelectual com cariótipo normal.** *Einstein* (16794508), v. 14, n. 1 , p. 30-34
- RADHAKRISHNAN, S. K.; JETTE, N.; LEES-MILLER, S. P. 2014. **Non-homologous end joining: emerging themes and unanswered questions.** *DNA repair,* v. 17, p. 2-8.
- RAVNAN J. B.; TEPPERBERG J. H.; PAPHENHAUSEN P.; 2006. et al. **Subtelomere FISH analysis of 11 688 cases: an evaluation of the frequency and pattern of subtelomere rearrangements in individuals with developmental disabilities.** *J Med Genet.* v. 43, n. 6, p. 478-489.
- ROCHA, N. B. 2014. **Busca de microrrearranjos no cromossomo X em meninos com deficiência intelectual.** *Dissertação (Mestrado em Ciências da Saúde).* Universidade de Brasília, Brasília, 21-MAI-2014.
- ROBERS, H. H. 2008. **Genetics of intellectual disability.** *Curr Opin Genet Dev.* v. 18, n. 3, p. 241-250.
- SCHAAF, C. P.; WISZNIEWSKA, J.; BEAUDET, A. L.; 2011. **Copy number and SNP arrays in clinical diagnostics.** *Annual review of genomics and human genetics,* v. 12, p. 25-51.
- SCHALOCK, R. L.; BORTHWICK-DUFFY, S. A.; BRADLEY, V. J.; ET AL. 2010. **Intellectual disability: Definition, classification, and systems of supports.** *American*

Association on Intellectual and Developmental Disabilities. 444 North Capitol Street NW Suite 846, Washington, DC 20001.

SEBAT J.; LAKSHMI B.; TROGE J.; et al. 2004. **Large-scale copy number polymorphism in the human genome**. *Science*. v. 305, n. 5683, p. 525-528.

SEKIGUCHI, J. M.; FERGUSON, D. O. 2006. **DNA double-strand break repair: a relentless hunt uncovers new prey**. *Cell*, v. 124, n. 2, p. 260-262.

SHAW, C. J.; LUPSKI, J. R.; 2004. **Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease**. *Human molecular genetics*, v. 13, n. suppl_1, p. R57-R64.

SOLINAS-TOLDO S.; LAMPEL, S.; STILGENBAUER, S.; et al. 1997. **Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances**. *Genes Chromosomes Cancer*. v. 20, n. 4, p. 399-407.

SOUTH S. T.; LEE C., LAMB A. N.; et al. 2013. **ACMG standards and guidelines for constitutional cytogenomic microarray analysis, including postnatal and prenatal applications: revision 2013**. *Genet Med*. v. 15, n. 11, p. 901.

STAMOULI, S.; ANDERLID, B. M.; WILLFORS, C.; et al. 2018. **Copy Number Variation Analysis of 100 Twin Pairs Enriched for Neurodevelopmental Disorders**. *Twin Research and Human Genetics*, p. 1-11.

STANKIEWICZ, P.; LUPSKI, J. R. 2002. **Genome architecture, rearrangements and genomic disorders**. *TRENDS in Genetics*, v. 18, n. 2, p. 74-82.

STANKIEWICZ, P.; LUPSKI, J. R. 2010. **Structural variation in the human genome and its role in disease**. *Annual review of medicine*, v. 61, p. 437-455.

SUDMANT P. H.; HUDDLESTON, J.; CATAACCHIO, C. R.; et al. 2013. **Evolution and diversity of copy number variation in the great ape lineage**. *Genome Res*. v. 23, n. 9, p. 1373-1382.

SUN, M.; ZHANG, H.; LI, G.; et al. 2017. **Molecular characterization of 20 small supernumerary marker chromosome cases using array comparative genomic hybridization and fluorescence in situ hybridization**. *Scientific reports*, v. 7, n. 1, p. 10395.

TAO, V. Q.; CHAN, K. Y. K.; CHU, Y. W. Y.; et al. 2014. **The clinical impact of chromosomal microarray on paediatric care in Hong Kong**. *PLoS One*, v. 9, n. 10, p. e109629.

THEISEN, A. 2008. **Microarray-based comparative genomic hybridization (aCGH)**. *Nature Education*, v. 1, n. 1, p. 45.

THOMAS, N. S.; DURKIE, M.; VAN Z. B.; et al. 2006. **Parental and chromosomal origin of unbalanced de novo structural chromosome abnormalities in man**. *Human genetics*, v. 119, n. 4, p. 444.

TORRES, F.; BARBOSA, M.; MACIEL, P. 2016. **Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications**. *J Med Genet*. p. jmedgenet-2015-103366.

VERMA, R. S.; BABU, A. 1995. **Human chromosomes: principles and techniques.** *McGraw-Hill.*

VISSERS, L. E. L .M.; GILISSEN, C.; VELTMAN, J. A.; 2016. **Genetic studies in intellectual disability and related disorders.** *Nature reviews. Genetics*, v. 17, n. 1, p. 9.

WATSON, J. D.; et al. 2015. **Biologia molecular do gene.** *Artmed Editora.*

WORLD HEALTH ORGANIZATION et al. 2011. **World report on disabilities.** *Geneva: World Health Organization.*

WOU, K.; LEVY, B.; WAPNER, R. J. 2016. **Chromosomal microarrays for the prenatal detection of microdeletions and microduplications.** *Clinics in laboratory medicine*, v. 36, n. 2, p. 261-276.

XU Y.; PENG B.; FU Y.; et al. 2011. **Genome-wide algorithm for detecting CNV associations with disease.** *BMC Bioinformatics.* v. 12, n. 1, p. 331.

9 – ANEXOS

9.1 Anexo 1

Termo de Consentimento Livre e Esclarecido

Dados da Identificação

Título do Projeto: Investigação das Causas Genéticas e Genômicas do Retardo Mental Autossômico

Pesquisador Responsável: Dr. Aparecido Divino da Cruz

Nome (sujeito da pesquisa): _____,

RG nº _____, Nacionalidade _____, Idade _____,

Endereço _____,

Neste ato representado por mim:

Nome dos representantes legais:

Pai _____ Nacionalidade

_____, Idade _____, Estado Civil _____,

Profissão _____, Endereço _____

Mãe _____ Nacionalidade

_____, Idade _____, Estado Civil

_____, Profissão _____, Endereço _____

Está sendo convidado(a) a participar de um estudo intitulado: Investigação das Causas Genéticas e Genômicas do Retardo Mental Autossômico cujos objetivos e justificativas são: propor a investigação genética do paciente com indicação clínica de retardo mental, visto que o diagnóstico de um paciente com essa síndrome estabelece uma oportunidade valiosa de fazer estudos na família, identificando outros afetados e portadores, e de realizar um aconselhamento genético eficiente permitindo a prevenção de novos casos, possibilitando

assim, um tratamento específico, diminuindo o impacto econômico, social e pessoal que esta síndrome acarreta.

Os critérios de inclusão são: pacientes encaminhados ao Núcleo de Pesquisas Replicon e LaGene com indicação clínica de retardo mental, acompanhados dos pais biológicos que concordarem em assinar o termo de consentimento livre e esclarecido (TCLE).

Os critérios de exclusão são: indivíduos maiores de 18 anos, com indicação clínica de retardo mental, que não concordarem em assinar o termo de consentimento livre e esclarecido, ou menores, cujos pais biológicos não aceitarem a participação do(a) filho(a) neste estudo, ou que não assinarem o TCLE, ou ainda, pacientes que apresentarem outras doenças genéticas já diagnosticadas.

A minha participação no referido estudo será no sentido de permitir a doação voluntária de uma amostra de sangue de meu(minha) filho(a) biológico(a) para colaborar com o diagnóstico do retardo mental, juntamente com uma amostra, doada voluntariamente, de meu sangue para comparação com a finalidade de determinar causas genéticas e sua hereditariedade.

Fui alertado de que, da pesquisa a se realizar, posso esperar alguns benefícios, tais como: a confirmação da causa do retardo mental (quando tiver uma etiologia genética). Estou ciente ainda, que a técnica proposta já foi realizada em outros estudos e é reprodutiva, evitando assim possíveis riscos.

Recebi, por outro lado, os esclarecimentos necessários sobre os possíveis desconfortos e riscos decorrentes do estudo, levando-se em conta que é uma pesquisa, e os resultados positivos ou negativos somente serão obtidos após a sua realização. Durante a coleta do sangue, você poderá sentir uma dor leve a moderado, em decorrência da aplicação da agulha. Podem, também, ocorrer a formação de hematomas que não são comuns, e caso isso ocorra, você será imediatamente encaminhado(a) ao Serviço Médico da Pontifícia Universidade Católica de Goiás (PUC-GO).

Estou ciente de que a minha privacidade será respeitada, ou seja, meu nome ou qualquer outro dado, ou elemento, que possa, de qualquer forma, me identificar, será mantido em sigilo. Também fui informado de que posso me recusar a participar do estudo, ou retirar

meu consentimento a qualquer momento, sem precisar justificar, e de, por desejar sair da pesquisa, não sofrerei qualquer prejuízo à assistência que venho recebendo.

O pesquisador responsável, envolvido com o referido projeto é: Aparecido Divino da Cruz, e com ele poderei manter contato pelos telefones: (62)3946-1443/3946-1086.

É assegurada a minha assistência durante toda pesquisa, bem como me é garantido o livre acesso a todas as informações e esclarecimentos adicionais sobre o estudo e suas consequências, enfim, tudo o que eu queira saber antes, durante e depois da minha participação.

Tenho sido orientado ao teor de todo o aqui mencionado e compreendido a natureza e o objetivo do já referido estudo, autorizo a participação de (nome do sujeito da pesquisa) _____ na referida pesquisa, e manifesto meu livre consentimento em participar, estando totalmente ciente de que não há nenhum valor econômico, a receber ou pagar.

No entanto, caso eu tenha qualquer despesa decorrente da participação na pesquisa, haverá ressarcimento na forma de dinheiro em espécie. De igual maneira, caso ocorra algum dano decorrente da minha participação no estudo, serei devidamente indenizado, conforme determina a lei.

Goiânia, ____ de _____ de 201__.

(nome e assinatura do Pai)

(nome e assinatura da Mãe)

Pesquisador Dr. Aparecido Divino da Cruz

9.2 Anexo 2

DECIPHER

DatabasE of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources

Banco de Dados de Desequilíbrio Cromossômico e Fenótipo em seres Humanos utilizando o Ensembl Uma iniciativa de colaboração para atribuir fenótipos a locos no mapa genômico humano

O banco de dados DECIPHER de desequilíbrio cromossômico submicroscópico recolhe informações clínicas sobre microdeleções cromossômicas/ duplicações e inversões e exibe esta informação no mapa do genoma humano, com os seguintes objetivos:

- Aumentar o conhecimento médico e científico sobre microdeleções cromossômicas/ duplicações e inversões.
- Melhorar a assistência médica e aconselhamento genético para os indivíduos/famílias com desequilíbrio cromossômico submicroscópico.
- Facilitar a investigação sobre o estudo de genes que afetam a saúde e o desenvolvimento humano.

Informações adicionais sobre o banco de dados DECIPHER de microdeleções/duplicações/ inversões cromossômicas estão disponíveis na Internet em: <http://decipher.sanger.ac.uk>.

Termo de Consentimento

Eu dou permissão no meu nome/ no nome do meu filho(a)

.....

nascido(a) em para o Dr. Prof. APARECIDO DIVINO DA CRUZ do NÚCLEO DE PESQUISAS REPLICON E LABORATÓRIO DE CITOGENÉTICA HUMANA E GENÉTICA MOLECULAR – LAGENE, RUA 235, nº40, Área 4, Bloco L, Setor Universitário, Goiânia – GO, para enviar:

- Dados clínicos anônimos sobre mim/ meu filho(a) para o banco de dados DECIPHER:
SIM/ NÃO
(por favor, circule sua opção) Fotografias minhas/ do meu filho(a) para o banco de dados DECIPHER (opcional): SIM/ NÃO (por favor, circule sua opção)
- Eu li o folheto explicativo “Bem-vindo ao banco de dados DECIPHER - Informação para as famílias”:

SIM/ NÃO

(por favor, circule sua opção)

Eu entendo que:

- Meu nome/ o nome do meu filho(a) não será publicado. ○ O material será colocado no site DECIPHER em <http://decipher.sanger.ac.uk>. O site será visualizado e usado por médicos (principalmente os geneticistas clínicos) e cientistas (principalmente citogeneticistas e biólogos moleculares).
- O site DECIPHER pode ser acessado por médicos e cientistas que trabalham em países fora da União Europeia, por exemplo EUA, Canadá e Austrália. A legislação de proteção de dados é uniforme em toda a União Europeia, mas pode variar em países não pertencentes à UE.
- Dados resumidos e totalmente anônimos depositados no DECIPHER poderão ser visualizados através de navegadores genômicos disponíveis publicamente na internet e serão acessíveis a outros países.
- Os organizadores do DECIPHER não podem aceitar qualquer responsabilidade legal pelo uso indevido de informações.
- Qualquer fotografia será protegida por senha e ficará disponível somente para os médicos /clínicos citogeneticistas membros registrados do Consórcio DECIPHER.

Assinatura do(a) paciente/ pai/ mãe/ responsável:.....

Por favor, indique a sua relação com a pessoa, ex: eu mesmo/ mãe/ pai/ responsável:.....

Data da assinatura:

Assinatura do(a) biomédico(a) geneticista/ conselheiro geneticista solicitante do TCLE:.....

Data da assinatura do(a) biomédico(a) geneticista/ conselheiro geneticista:.....

The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambs CB10 1SA, UK
 Tel: +44 (0)1223 834244. Fax: +44 (0)1223 494919 www.ensembl.org/helpdesk/
 Cambridge University Dept of Medical Genetics, Box 134 Addenbrookes Hospital, Cambridge CB2 2QQ, UK Tel:
 +44 (0)1223 216446 Fax; +44 (0)1223 217054