



UNIVERSIDADE FEDERAL DE GOIÁS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E BIODIVERSIDADE
DOCTORADO EM BIOTECNOLOGIA E BIODIVERSIDADE

**AUMENTO NA TAXA DE DIAGNÓSTICO GENÉTICO DOS
PACIENTES A PARTIR DA IDENTIFICAÇÃO DE CNVS, POR
CMA, ENVOLVENDO GENES IMPLICADOS COM A
MANIFESTAÇÃO CLÍNICA DA DEFICIÊNCIA INTELECTUAL**

IRENE PLAZA PINTO

Goiânia
Goiás - Brasil
©2019

**TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR
VERSÕES ELETRÔNICAS DE TESES E DISSERTAÇÕES
NA BIBLIOTECA DIGITAL DA UFG**

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: Dissertação Tese

2. Identificação da Tese ou Dissertação:

Nome completo do autor: Irene Plaza Pinto

Título do trabalho: Aumento na taxa de diagnóstico genético dos pacientes a partir da identificação de CNVs, por CMA, envolvendo genes implicados com a manifestação clínica da deficiência intelectual

3. Informações de acesso ao documento:

Concorda com a liberação total do documento SIM NÃO¹

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.



Assinatura do(a) autor(a)²

Ciente e de acordo:


Prof. Aparecido D. de Cruz, PhD
Coordenador - RE 3989
Núcleo de Pesquisas Replicon

Assinatura do(a) orientador(a)²

Data: 05/06/2019

¹ Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

² A assinatura deve ser escaneada.

IRENE PLAZA PINTO

**AUMENTO NA TAXA DE DIAGNÓSTICO GENÉTICO DOS
PACIENTES A PARTIR DA IDENTIFICAÇÃO DE CNVS, POR
CMA, ENVOLVENDO GENES IMPLICADOS COM A
MANIFESTAÇÃO CLÍNICA DA DEFICIÊNCIA INTELECTUAL**

Orientador: Prof. Dr. Aparecido Divino da Cruz, *PhD*

Co-Orientador: Prof. Dr. Robert Pogue, *PhD*

Tese apresentada ao Programa de Pós-Graduação em
Biotecnologia e Biodiversidade da Rede Pró-Centro-
Oeste como requisito para obtenção do título de *Doctor
Scientiae*.

Goiânia

Goiás – Brasil

©2019

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Plaza Pinto, Irene

Aumento na taxa de diagnóstico genético dos pacientes a partir da identificação de CNVs, por CMA, envolvendo genes implicados com a manifestação clínica da deficiência intelectual [manuscrito] / Irene Plaza Pinto. - 2019.

115 f.

Orientador: Prof. Dr. Aparecido Divino da Cruz; co-orientador Dr. Robert Edward Pogue.

Tese (Doutorado) - Universidade Federal de Goiás, Instituto de Patologia Tropical e Saúde Pública (IPTSP), Programa de Pós graduação em Biotecnologia e Biodiversidade, Cidade de Goiás, 2019.

Anexos. Apêndice.

Inclui gráfico, tabelas, lista de figuras, lista de tabelas.

1. Análise Cromossômica por Microarranjo. 2. CNTNAP2. 3. SHANK3. 4. IL1RAPL1. I. Divino da Cruz, Aparecido, orient. II. Título.

CDU 60



UNIVERSIDADE FEDERAL DE GOIÁS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E
BIODIVERSIDADE
Rua 235, S/N - Setor Universitário - Goiânia/GO CEP 74605-050
Fone (62) 3209.6362
email: pgbh.goias@gmail.com

ATA DA REUNIÃO DA BANCA EXAMINADORA DA DEFESA DE TESE DE IRENE PLAZA PINTO - Aos vinte e sete dias do mês de maio do ano de 2019 (27/05/2019), às 14:00 horas, reuniram-se os componentes da Banca Examinadora: Profs. Drs. APARECIDO DIVINO DA CRUZ, CLÁUDIO CARLOS DA SILVA, LYSA BERNARDES MINASI, KÁTIA KARINA VEROLLI DE OLIVEIRA MOURA e MARIA DAS GRAÇAS NUNES BRASIL, para, sob a presidência da primeira, e em sessão pública realizada na ÁREA IV DA PUC-GOÍÁS SALA DO MESTRADO DE GENÉTICA, procederem à avaliação da defesa de tese intitulada: "AUMENTO NA TAXA DE DIAGNÓSTICO GENÉTICO DOS PACIENTES A PARTIR DA IDENTIFICAÇÃO DE CNVS, POR CMA, ENVOLVENDO GENES IMPLICADOS COM A MANIFESTAÇÃO CLÍNICA DA DEFICIÊNCIA INTELECTUAL", em nível de DOUTORADO, área de concentração em BIOTECNOLOGIA, de autora de IRENE PLAZA PINTO, discente do PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E BIODIVERSIDADE, da Universidade Federal de Goiás. A sessão foi aberta pelo orientador da discente, Prof. Dr. APARECIDO DIVINO DA CRUZ, que fez a apresentação formal dos membros da Banca e orientou a Candidata sobre como utilizar o tempo durante a apresentação de seu trabalho. A palavra a seguir, foi concedida a autora da tese que, em 30 minutos procedeu à apresentação de seu trabalho. Terminada a apresentação, cada membro da Banca arguiu a Candidata, tendo-se adotado o sistema de diálogo sequencial. Terminada a fase de arguição, procedeu-se à avaliação da defesa. Tendo-se em vista o que consta na Resolução nº. 1181/2013 do Conselho de Ensino, Pesquisa, Extensão e Cultura (CEPEC), que regulamenta o Programa de Pós-Graduação em Biotecnologia e Biodiversidade a Banca, em sessão secreta, expressou seu julgamento, considerando a candidata **Aprovada ou Reprovada**:

Banca Examinadora	Aprovada / Reprovada
Prof. Dr. Aparecido Divino da Cruz	<u>APROVADA</u>
Prof. Dr. Cláudio Carlos da Silva	<u>Aprovada</u>
Profa. Dra. Lysa Bernardes Minasi	<u>APROVADA</u>
Profa. Dra. Kátia Karina Verolli de Oliveira Moura	<u>APROVADA</u>
Profa. Dra. Maria das Graças Nunes Brasil	<u>Aprovada com louv.</u>

Em face do resultado obtido, a Banca Examinadora considerou a candidata APROVADA/HABILITADA (Habilitada ou não Habilitada), cumprindo todos os requisitos para fins de obtenção do título de DOUTORA EM BIOTECNOLOGIA E BIODIVERSIDADE, na área de concentração em BIOTECNOLOGIA, pela Universidade Federal de Goiás. Cumpridas as formalidades de pauta, às 18 h 15 min, a presidência da mesa encerrou esta sessão de defesa de tese e para constar eu, POLLYANA REZENDE AQUINO, secretária do Programa de Pós-Graduação em Biotecnologia e Biodiversidade lavrei a presente Ata que depois de lida e aprovada, será assinada pelos membros da Banca Examinadora e por mim em duas vias de igual teor.
A Banca Examinadora aprova a seguinte alteração no título da Tese:

→ sem alteração

Prof. Dr. Aparecido Divino da Cruz (PUC/GO) Junc
Prof. Dr. Cláudio Carlos da Silva (PUC/GO) C. Pinheiro
Profa. Dra. Lysa Bernardes Minasi (PUC/GO) Lysa Bernardes Minasi
Profa. Dra. Kátia Karina Verolli de Oliveira Moura (PUC/GO) Katya Verolli
Profa. Dra. Maria das Graças Nunes Brasil (FM/UFG) Procell
Secretário da Pós-Graduação: Pollyana Rezende Aquino



UNIVERSIDADE FEDERAL DE GOIÁS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOTECNOLOGIA E
BIODIVERSIDADE
Rua 235, S/N - Setor Universitário - Goiânia/GO CEP 74605-050
Fone (62) 3209.6362
email: pgbb.goias@gmail.com

FOLHA DE MODIFICAÇÃO DE DISSERTAÇÃO/TESE

Modificações exigidas na Ata da reunião da banca examinadora da defesa de tese de **IRENE PLAZA PINTO**, realizada em 27/05/2019, e que a mesma segue a Norma 006/2015 do PPGBB-Goiás (Modelo de elaboração de Teses/Dissertação):

Nenhuma exigência.

O prazo para a entrega da versão final corrigida é de 40 (quarenta) dias após a defesa, sendo responsável pelo recebimento o orientador da aluna Prof. Dr. Aparecido Divino da Cruz, Presidente da Banca.

Candidata: Irene Plaza Pinto
Irene Plaza Pinto

Atesto que as alterações exigidas foram cumpridas, dentro do prazo estipulado.

Aparecido Divino da Cruz
Prof. Dr. **Aparecido Divino da Cruz**

INSTITUIÇÕES PARTICIPANTES

O desenvolvimento das atividades experimentais do presente estudo contou com a colaboração direta das instituições: Núcleo de Pesquisas Replicon da Escola de Ciências Agrárias e Biológicas da Pontifícia Universidade Católica de Goiás (PUC-Goiás), Laboratório de Citogenética Humana e Genética Molecular do Estado de Goiás (LaGene/LACEN/SES-GO) e Laboratório de Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília. Além disso, o desenvolvimento da tese foi apoiado pelas Universidade Federal de Goiás (UFG) e Universidade de Brasília (UnB) através do Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede Centro-Oeste de Pós-Graduação, Pesquisa e Inovação e Rede ExGens – Rede de Excelência em Genética e Genômica Molecular Aplicada à Saúde Humana.

Ademais, as atividades desenvolvidas tiveram o apoio financeiro das agências de fomento à pesquisa: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) e Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG).

“Dizem que a vida é para quem sabe viver, mas ninguém nasce pronto. A vida é para quem é corajoso o suficiente para se arriscar e humilde o bastante para aprender”.

Clarisse Lispector

DEDICATÓRIAS

DEDICO AO GUSTAVO, minha alma gêmea e meu companheiro de vida e luta. Ter você ao meu lado me deu forças para ser uma pessoa melhor e para lutar pelos meus sonhos. Sou grata pelo apoio, carinho, incentivo, respeito, companheirismo ao longo de nossas vidas e pelo grande amor que compartilhamos.

DEDICO AOS MEUS AMADOS FILHOS GUSTAVO, NATÁLIA E LETÍCIA, meus presentes de vida, que me fizeram evoluir na minha existência. Gratidão por estarem comigo nesta vida e por me apoiarem e estimularem sempre.

DEDICO À MINHA MÃE ALBINEIAR E AO MEU PAI GIL, pelo exemplo de vida e por me mostrarem um mundo melhor, justo e possível, com muita luta e esperança.

AGRADECIMENTOS

A contribuição de várias pessoas com sugestões, críticas, carinho, apoio, ideias, afeto e muita amizade, foi que proporcionou a realização e finalização deste trabalho. Agradeço:

Ao meu marido Gustavo Pinto Silva e aos meus filhos Gustavo Pinto Silva Filho, Natália Plaza Pinto Silva e Letícia Plaza Pinto Silva, por estarem comigo nesta vida e por me apoiarem. Amo muito vocês!

Aos meus pais Gil Santini Pinto e Albineiar Plaza Pinto, por serem um dos meus maiores incentivadores e me ajudarem e estimularem a crescer. Tenho muito orgulho de ser filha de vocês!

Ao Prof. Dr. Aparecido Divino da Cruz, por acreditar em mim e permitir meu crescimento pessoal e profissional. Por sua orientação agradável, inspiradora e acolhedora. Por sua integridade, dignidade e generosidade e principalmente por nossa amizade, carinho e amor construídos durante a nossa caminhada, que me incentivaram a fazer e finalizar este estudo. Minha eterna gratidão por plantar a sementinha da pesquisadora, conselheira e cuidadora.

Ao Prof. Dr. Robert Edward Pogue, pela prontidão em aceitar me coorientar, por confiar em mim e por colaborar com o meu desenvolvimento como pesquisadora. Por todos os momentos de imensa ajuda e disposição, meu profundo agradecimento.

Aos Prof. Dr. Cláudio Carlos da Silva, Prof. Dr. Alex Silva da Cruz e Prof. Rinaldo Welerson Pereira, pela colaboração, apoio e discussões positivas e construtivas referentes ao estudo, que me fizeram crescer profissionalmente e pessoalmente.

À Dra. Nicole de Leeuw, pela presteza e auxílio nas análises e discussões pertinentes ao estudo. Grata por compartilhar seu imenso conhecimento.

À amiga-irmã Profa. Dra. Lysa Bernardes Minasi, que eu aprendi a amar e a respeitar como profissional e pesquisadora, mas principalmente como pessoa. Trabalhar com você é muito gratificante, mas conviver com você no dia a dia, nas nossas viagens e também nas nossas crenças me engrandece. Gratidão por partilhar essa e outras vidas com você.

Aos amigos Danilo Conrado Silva e Emília Oliveira Alves Costa, meus companheiros de espiritualidade, por me ajudarem a ser uma pessoa melhor, mais tolerante e por contribuírem na minha evolução.

Aos amigos Damiana Míriam da Cruz e Cunha, Cristiano Luiz Ribeiro e Eduardo Rocha Pedrosa, por partilharem comigo a amizade e conhecimento. Trabalhar com vocês é agradável.

Aos amigos do Laboratório REPLICON-PUC-GO e LAGENE-SES-GO: Lorryayne Guimarães Oliveira, Juliana Ferreira da Silva, Ana Júlia Cunha Leite, Marcos Vinícius Milk, Andreia Pires Amancio, Samara Socorro Silva Pereira, Rafael Carneiro, Jakeline Soares Fortes, Sabrina Sara Moreira Duarte, Lucas Henrique Nascimento Silva Rodrigues, Gustavo Dib, Victor Cortázio do Prado Santos, Larissa Resende Carvalho, Hugo Leite Filho, Thaís Cidália Vieira, Marc Alexandre Gigonzac, João Antônio Xavier Manso, Raphael Silva da Cruz, Fernanda Ribeiro Godoy, Hérica Samaya Gonçalves de Sousa, Caroline Oliveira de Araújo e Melo, Calebe Bertolino Marins de Campos e Aldaires Vieira de Melo, por compartilharem amizade, apoio e aprendizado.

À toda equipa do Laboratório de Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília, representada por Bianca Simonassi Raso de Paiva, meus sinceros agradecimentos por toda acolhida e ajuda.

A todos os professores do Programa de Pós-Graduação em Biotecnologia e Biodiversidade da Rede Centro-Oeste de Pós-Graduação, Pesquisa e Inovação pela troca de conhecimentos e experiências.

Aos meus queridos e amados irmãos e irmãs Cláudio Plaza Pinto, Joana Plaza Pinto, Pedro Plaza Pinto e Míriam Plaza Pinto por serem quem vocês são e por vivenciarmos nossas vidas com muito amor, admiração, carinho e alegria.

Às minhas sobrinhas e sobrinhos Luísa Plaza dos Santos, Beatriz Plaza dos Santos, Teresa Plaza Pedrosa, Clara Motta, Laura Motta, Martim Plaza Ono, Marco Aurélio de Faria, José de Castro Neto, Rodrigo de Castro, Raul Mendes e Pedro Henrique por me fazerem a tia mais feliz desse mundo e por me ajudarem a ser uma pessoa melhor.

À minha sogra Nelma Pinto Silva (*in memorian*), sogro Antônio da Silva (*in memorian*) e cunhada Lílian Soares de Castro por fazerem parte da minha vida e estarem ao meu lado.

À minha querida avó e madrinha Dineyar Valente Plaza, por sempre ter me apoiado nos estudos, por acreditar em mim e por todo o amor.

A todos e todas pacientes e familiares que, com a sua humildade, esperança e luta, fizeram aprimorar o meu olhar para o outro e sua condição e a entender que a empatia é a chave para um mundo mais humano. Agradeço também por aceitarem participar da pesquisa e colaborarem com meu aprendizado. Minha respeitosa e eterna gratidão!

Meu cordial e verdadeiro agradecimento a todos vocês.

SUMÁRIO

LISTA DE FIGURAS	xiv
LISTA DE TABELAS E QUADROS.....	xvi
RESUMO	xvii
ABSTRACT	xviii
1. INTRODUÇÃO.....	1
2. REFERENCIAL TEÓRICO.....	3
2.1 Deficiência Intelectual	3
2.2 Aspectos Epidemiológicos da Deficiência Intelectual.....	5
2.3 Etiologia da Deficiência Intelectual.....	7
2.3.1 Genética e aspectos moleculares da deficiência intelectual	8
2.4 Diagnóstico Genético da Deficiência Intelectual.....	17
3. JUSTIFICATIVA	20
4. OBJETIVOS.....	22
4.1 Objetivo Geral.....	22
4.2 Objetivos Específicos	22
5. MATERIAL E MÉTODOS.....	23
5.1 Delineamento do estudo.....	23
5.2 Grupo Amostral	23
5.3 Fluxograma de atendimento.....	24
5.4 Coleta e processamento das amostras	25
5.5 Cariotipagem por bandeamento GTG	25
5.6 Extração de DNA	25
5.7 Análise Cromossômica por Microarranjos (CMA)	26
5.8 Desenho dos <i>primers</i> , amplificação por PCR das regiões de CNVs de perda selecionadas e purificação dos amplicons.....	27

5.9 Sequenciamento de Sanger	29
6. RESULTADOS	31
7. DISCUSSÃO	37
a) Gene <i>CNTNAP2</i>	39
b) Gene <i>FGF13</i>	41
c) Gene <i>MID1</i>	42
d) Gene <i>MID2</i>	44
e) Gene <i>SHANK3</i>	44
f) Gene <i>IL1RAPL1</i>	46
g) Gene <i>DMD</i>	48
h) Gene <i>PAK3</i>	49
8. CONCLUSÃO.....	53
9. REFERÊNCIAS BIBLIOGRÁFICAS	55
ANEXO 1 – TCLE DO PROJETO	67
ANEXO 2 – TCLE DO DECIPHER.....	70
APÊNDICE 1 – BLAST DOS PRIMEIROS <i>PRIMERS</i> E DOS RESULTADOS DO SEQUENCIAMENTO	72
APÊNDICE 2 – BLAST DOS <i>PRIMERS</i> REDESENHADOS E DA NOVA SEQUÊNCIA DO GENE <i>CNTNAP2</i>	75
APÊNDICE 3 – ARTIGO PUBLICADO	76
APÊNDICE 4 – CAPÍTULO DE LIVRO PUBLICADO	80

LISTA DE FIGURAS

Figura 1. Imagem de uma página do <i>Papyrus Ebers</i> , uma compilação de textos médicos, formulações farmacêuticas e remédios caseiros para tratar as aflições e os agravos à saúde humana.	3
Figura 2. Proporção de pessoas com DI no território nacional brasileiro, de acordo com o IBGE.	6
Figura 3. Causas Genéticas da DI.	11
Figura 4. Rearranjos genômicos resultantes de recombinação homóloga não-alélica (NAHR) mediada por LCRs.	13
Figura 5. Processamento do pré-mRNA.....	16
Figura 6. Fluxograma indicando o encaminhamento, seleção, triagem e metodologias utilizadas na tentativa de elucidação diagnóstica de casos de DI.	24
Figura 7. Representação esquemática para a seleção dos primers.	27
Figura 8. Representação gráfica do programa de amplificação (PCR).	28
Figura 9. Indicação clínica e resultados do cariótipo dos pacientes com DI atendidos no período de Jan/2013 a Dez/2016 NPR/PUC-Goiás/LaGene/SES/GO.....	31
Figura 10. Tipos de CNVs, classificadas em função da sua patogenicidade, encontradas nos pacientes com DI que fizeram o CMA usando o filtro de diagnóstico recomendando pelo fabricante do GeneChip® CytoScanHD™ array.	32
Figura 11. Fluxograma dos resultados obtidos durante a realização do estudo.	34
Figura 12. Ilustração da distribuição dos exons e dos domínios regulatórios na estrutura do gene <i>CNTNAP2</i>	40

Figura 13. Representação esquemática das isoformas da família <i>FGF11</i> não secretória intracelular.....	41
Figura 14. Representação esquemática do gene <i>MID1</i> , sua localização e proteína <i>MID1</i>	43
Figura 15. Desenho esquemático do gene <i>SHANK3</i> com representação dos exons e domínios e das ilhas CpG (barras azuis).	45
Figura 16. Representação esquemática do gene <i>ILIRAPL1</i> e da proteína transcrita.	47
Figura 17. Organização genômica do gene <i>DMD</i> e representação das isoformas.	49
Figura 18. Representação esquemática da proteína <i>PAK3</i>	50

LISTA DE TABELAS E QUADROS

Tabela 1. Relação do tamanho do produto de PCR com o volume da amostra para sequenciar.	29
Tabela 2. Resultado da análise com filtro de tamanho < 100 kb dos casos negativos do CMA.	33
Tabela 3. CNVs de perda usadas nos experimentos de identificação dos prontos de quebra por Sequenciamento de Sanger, seus respectivos genes, a localização da última sonda <i>upstream</i> e primeira sonda <i>downstream</i> , sequência dos <i>primers forward</i> (F) e <i>reverse</i> (R) utilizados para cada região e o intervalo de tamanho esperado do <i>amplicon</i>	35
Tabela 4. Sequência de oligonucleotídeos (<i>primers</i>) redesenhados para cada região dos genes selecionados e o intervalo de tamanho mínimo e máximo esperado de cada <i>amplicon</i>	36
Quadro 1. Concentração reagentes no preparo do master mix para cada PCR.....	28

RESUMO

Deficiência intelectual (DI) é reconhecida como um transtorno do desenvolvimento neurológico, caracterizado por diminuição significativa em ambas funções cognitivas e adaptativas, originada antes dos 18 anos de idade. Estima-se que em torno de 1–3% das crianças ao redor do mundo são afetadas pela DI e que no Brasil acomete aproximadamente 0,8% da população. As variações no número de cópias (CNVs) são responsáveis por cerca de 15–20% dos casos de DI, que podem comprometer o funcionamento de diversos genes. O objetivo do estudo foi avaliar a ocorrência de CNVs, identificadas pelo CMA com filtro de tamanho < 100 kb, que abrigassem genes funcionalmente associados à DI em pacientes do SUS com diagnóstico clínico de DI. Durante o período de janeiro 2013 a dezembro de 2016, foi realizado o cariótipo por bandeamento GTG em 325 pacientes com indicação clínica de DI, chegando ao diagnóstico genético em 57,2% dos pacientes, demonstrando ser uma importante metodologia de triagem para os pacientes com DI. No entanto, 42,8% dos pacientes apresentaram o cariótipo sem alterações estruturais ou numéricas. Foi realizado nestes pacientes o CMA com filtro de tamanho \geq 100 kb, onde foi possível elucidar o diagnóstico genético em 29,8% dos casos, apresentando um incremento na taxa de diagnóstico de 7,1%. Os casos que permaneceram sem elucidação, foram submetidos ao CMA com filtro de tamanho < 100 kb, no qual foram identificadas CNVs de perda em regiões que continham os genes *CNTNAP2*, *FGF13*, *MID1*, *MID2*, *SHANK3*, *IL1RAPL1*, *DMD* e *PAK3*. A redução do filtro de tamanho demonstrou um incremento na taxa de diagnóstico em 12%, ampliando o espectro de identificação de CNVs presentes em regiões que abrigam genes implicados à manifestação clínica da DI. A aplicação em conjunto das metodologias cariótipo por bandeamento GTG e CMA com filtro de tamanho \geq 100 kb e posteriormente filtro de tamanho < 100 kb, permitiu um aumento no diagnóstico genético da DI e outras comorbidades, promovendo uma compreensão mais ampla dos aspectos genéticos relacionados à essas afecções e possibilitando um manejo adequado das famílias. Finalmente, o acesso ao aconselhamento genético proporciona uma melhor compreensão das causas genéticas da DI, as implicações familiares da contribuição genética e a chance de recorrência.

Palavras-chave: Análise Cromossômica por Microarranjo; *CNTNAP2*; *SHANK3*; *IL1RAPL1*

ABSTRACT

Intellectual disability (ID) is characterized by significant impairment in both cognitive and adaptive functions, originating before the age of 18 years. In addition, it is a common phenotype sign in a cluster of heterogeneous syndromic or non-syndromic disorders, associated with some comorbidities such as autism and congenital malformations. In the worldwide, ID affects around 1–3% of the general population and in Brazil ID affects approximately 0.8% of the population. The Copy number variations account for about 15–20% of children with unexplained ID, compromising the functioning of several genes, with more than 1,416 genes described as causative of this phenotype sign. The aim of the study was to evaluate the occurrence of CNVs, identified by CMA with the size filter of < 100 kb, harboring genes functionally associated with ID in patients from SUS with a clinical diagnosis of ID referred for the genetic diagnosis. During January 2013 to December 2016, GTG banding karyotype was performed in 325 patients with ID, achieve the genetic diagnostic in 57.2%, demonstrating to be an important screening approach for patients with DI. However, 42,8% of the patients showed the karyotype with no visible numerical or structural alterations. The CMA analysis with the size filter of ≥ 100 kb was performed in these patients, where it was possible to elucidate the genetic diagnose in 29.8% of the patients, demonstrating 7,1 % of the increment on the diagnostic. All the cases remained without a diagnosis were submitted to the CMA analysis with size filter of < 100 kb, where it was identified loss CNVs in regions harboring *CNTNAP2*, *FGF13*, *MID1*, *MID2*, *SHANK3*, *IL1RAPL1*, *DMD*, and *PAK3* genes. The reduction of the size filter demonstrated an increase of 12% in the ratio of diagnosis, expanding the spectrum of CNVs identification in regions which harboring genes related to the clinical manifestation of ID. The application of both GTG banding and CMA with the size filter of ≥ 100 kb and later the size filter of < 100 kb allowed an increase in the genetic diagnosis of ID and comorbidities, giving a broad understanding of the genetic aspects related to these conditions and allowing the adequate management of families. Finally, the genetic counseling provides a better understanding of the genetic causes of ID, the familial implications of the genetic contribution and the chance of recurrence.

Keywords: Chromosomal Microarray Analysis; *CNTNAP2*; *SHANK3*; *IL1RAPL1*

1. INTRODUÇÃO

A deficiência intelectual (DI) é caracterizada por uma expressiva redução nas funções cognitivas e adaptativas, manifestada antes dos 18 anos de idade, se tornando evidente durante a primeira infância como Atraso Global no Desenvolvimento (AGD). No entanto, o diagnóstico adequado e preciso ocorre durante os anos escolares (BARTNIK *et al.*, 2014). A DI é mundialmente um dos principais problemas socioeconômicos, causando amplo impacto na vida dos pacientes e suas famílias (WESTERINEN *et al.*, 2017).

As causas genéticas são responsáveis por cerca de 25–50% dos casos de DI, que incluem desequilíbrios cromossômicos microscopicamente visíveis, variação do número de cópias, bem como mutações de ponto. Porém, a etiologia genética permanece desconhecida em pelo menos 50% de todos os casos de DI/AGD (GIJSBERS *et al.*, 2009; AGHA *et al.*, 2014).

O diagnóstico da DI/AGD é desafiador devido à apresentação de um amplo espectro fenotípico, sendo que pacientes com DI/AGD geralmente apresentam anormalidades congênitas e/ou os sinais comuns aos Transtornos do Espectro Autista (TEA) (BRETT *et al.*, 2014). Recentemente, foi demonstrado que rearranjos subteloméricos ou intersticiais são responsáveis por aproximadamente 10 a 25% dos casos de DI. Devido à heterogeneidade de fatores causais associados à DI, sua investigação diagnóstica é dificultada (PRATTE-SANTOS *et al.*, 2016).

Atualmente, sabe-se que o cariótipo por bandeamento GTG deve ser realizado para todo paciente com DI sem outras causas aparentes, sendo recomendado que se tenha um nível de resolução de no mínimo 550 bandas capaz de detectar rearranjos cromossômicos maiores que 5 Mb. Entretanto, um número considerável de desordens clínicas é causado por rearranjos cromossômicos submicroscópicos, menores que 5 Mb, que acaba resultando em um cariótipo normal em relação ao número e à estrutura visível dos cromossomos (VISSERS *et al.*, 2016).

Os microarranjos genômicos têm sido extensivamente usados no estudo das causas genéticas da DI e em desordens genômicas que são consideradas consequência clínica de Variação no Número de Cópias (CNVs) (do inglês, *Copy Number Variation*) de regiões específicas do genoma (VISSERS *et al.*, 2016). A análise cromossômica usando marcadores de Polimorfismo de Nucleotídeo Único (SNP) (do inglês, *Single*

Nucleotide Polymorphism) e marcadores não polimórficos tem sido comumente aplicada como uma ferramenta robusta de diagnóstico laboratorial em pacientes com DI/AGD, TEA e anormalidades congênitas múltiplas (ACM) (CHONG *et al.*, 2014). A Análise Cromossômica por Microarranjo (CMA) (do inglês, *Chromosomal Microarray Analysis*) avalia todo o genoma humano de interesse médico em um único chip com uma resolução pelo menos 10 vezes maior que o cariótipo por bandeamento GTG, sendo recomendado como teste diagnóstico de primeira escolha para pacientes com DI/AGD, TEA e/ou ACM (MILLER *et al.*, 2010; ŽILINA *et al.*, 2014).

Apesar do CMA ser uma técnica robusta e eficiente, com aumento na capacidade diagnóstica para os pacientes com DI/AGD de 15 a 25% (JANG *et al.*, 2019), alguns casos ainda permanecem sem o diagnóstico molecular. Para estes pacientes, o uso do filtro de tamanho < 100 kb pelo CMA é uma possibilidade para o aumento na capacidade diagnóstica desta metodologia. Porém, mesmo com uma alta cobertura de marcadores pelo genoma todo e devido à chance de obtenção de resultados falso-positivos, é recomendado o emprego do Sequenciamento de Sanger (do inglês, *Sanger Sequencing*) para validação dos pontos de quebra das CNVs observadas com filtro de tamanho < 100 kb pelo CMA (STROM *et al.*, 2014; ZAHIR *et al.*, 2017; SILVA *et al.*, 2019).

No presente estudo, realizado em pacientes encaminhados via SUS (Sistema Único de Saúde) com diagnóstico clínico de deficiência intelectual, foram aplicadas as metodologias de cariótipo por bandeamento GTG e CMA. Após a execução do CMA, nos casos que não apresentaram CNVs numa análise com filtro de tamanho ≥ 100 kb, estes considerados negativos, foi empregue o filtro de tamanho < 100 kb e identificadas CNVs presentes em genes importantes relacionados à DI. Para o mapeamento do ponto de quebra, foi utilizado o Sequenciamento de Sanger. Dessa forma, o estudo envolveu o uso das técnicas de cariótipo por bandeamento GTG, CMA e Sequenciamento de Sanger, para tentar mapear os pontos de quebra de CNVs com tamanhos bem reduzidos, que em conjunto aumentaram o incremento na taxa no diagnóstico laboratorial dos pacientes com deficiência intelectual.

2. REFERENCIAL TEÓRICO

2.1 Deficiência Intelectual

A natureza e a definição de deficiência intelectual têm sido debatidas desde do início da classificação de desordens mentais. A primeira referência à deficiência intelectual surgiu ainda na época da medicina Egípcia antiga, no *Papyrus Ebers* (Figura 1), um dos tratados médicos mais antigos e importantes que se conhece, mostrando o interesse médico em relação à essa desordem (HARRIS; GREENSPAN, 2016).



Figura 1. Imagem de uma página do *Papyrus Ebers*, uma compilação de textos médicos, formulações farmacêuticas e remédios caseiros para tratar as aflições e os agravos à saúde humana. O documento foi datado a cerca de 1.500 dc.

Fonte: <http://www.crystalinks.com/egyptmedicine.html>.

A Organização Mundial de Saúde (OMS) (do inglês, *World Health Organization* – *WHO*) define DI como uma desordem do desenvolvimento mental incompleto e capacidade significativamente reduzida de compreender informações novas e complexas e de aprender e aplicar novas habilidades, resultando num agravamento concreto das funções cognitiva, de linguagem, motoras e de sociabilização, com a adaptação ao ambiente comprometida (WHO, 2017).

De acordo com Associação Americana de Deficiência Intelectual e Desenvolvimento (AAIDD), a DI é caracterizada por limitações significativas em ambos o funcionamento intelectual e o comportamento adaptativo, comprometendo várias habilidades sociais e práticas do dia a dia (AAIDD, 2017).

Atualmente, é amplamente aceito que a DI é reconhecida como transtorno do desenvolvimento neurológico, caracterizada por limitações significativas no

funcionamento intelectual e adaptativo, comprometendo várias habilidades sociais e práticas diárias, originando antes dos 18 anos de idade. Outras características importantes que definem a DI são o *score* de QI (quociente de inteligência) abaixo da média (estabelecido em 70) e o déficit em dois ou mais comportamentos adaptativos (BERTELLI *et al.*, 2016; PRATTE-SANTOS *et al.*, 2016; WESTERINEN *et al.*, 2017). A DI pode se manifestar durante a primeira infância como Atraso Global no Desenvolvimento, no entanto é adequadamente diagnosticada durante os anos escolares (BARTNIK *et al.*, 2014).

O Atraso Global no Desenvolvimento é definido como um atraso significativo em dois ou mais domínios do desenvolvimento, incluindo habilidades motoras amplas, linguagem/fala, cognitivo, pessoal/social, e atividades do dia a dia. O termo AGD é aplicado para crianças menores de 5 anos, lembrando que em crianças acima de 5 anos aplica-se o termo DI (MILANI *et al.*, 2015; VISSERS *et al.*, 2016).

De acordo com a quinta edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-5) APA (2014), há três critérios que devem ser levados em consideração quanto ao diagnóstico clínico da DI: déficit nas habilidades mentais gerais, significativa limitação em uma ou mais áreas do comportamento adaptativo em vários ambientes (comunicação, habilidade de autoajuda, habilidades interpessoais entre outras), e evidência de que essas limitações se tornaram aparentes na infância e adolescência. Além disso, no manual há um especificador da gravidade da DI, conforme o funcionamento adaptativo, em 4 níveis:

1. Leve: indivíduos desempenham habilidades sociais e de comunicação, têm prejuízo mínimo nas áreas senso-motoras. Com apoio apropriado, podem viver com poucas limitações na comunidade, de modo independente ou supervisionado;
2. Moderada: indivíduos, em sua maioria, adquirem atraso nas habilidades de comunicação durante os primeiros anos da infância. Há dificuldades no reconhecimento de convenções sociais, no entanto, com moderada supervisão podem cuidar de si mesmos;
3. Grave: os indivíduos adquirem pouca ou nenhuma fala comunicativa, necessitam de supervisão para os cuidados pessoais e possuem o mínimo de entendimento, conseguindo atender e interagir com outros ou fazer algumas atividades domésticas;
4. Profunda: os indivíduos apresentam prejuízos significativos no funcionamento sensorio-motor, com limitações graves quanto aos cuidados pessoais, continência, comunicação e mobilidade.

Além da classificação dos níveis de gravidade, a DI pode ser classificada de acordo com o *score* de QI, em que faz uma avaliação do conhecimento geral do indivíduo, vocabulário, habilidades de resolução de problemas e capacidade de raciocínio (VAN BOKHOVEN, 2011; MARIS *et al.*, 2013). O QI médio deve ter um escore acima de 85, sendo o QI 70-79 considerado limítrofe nas habilidades intelectuais. Conforme o QI, a DI pode ser classificada em: DI leve, com QI de 55–70; DI moderada, com QI de 40–55; DI severa, com QI de 25–40; e DI profunda com QI < 25 (BARTNIK *et al.*, 2014; HARRIS; GREENSPAN, 2016).

O atraso no desenvolvimento global inicia-se desde a vida intrauterina e primeira infância, permanece como sinal fenotípico da DI ao longo da infância e adolescência até a idade adulta como resultado de doenças agudas ou progressivas que afetam o funcionamento do cérebro, incluindo síndromes genéticas, infecções intrauterinas, complicações ao nascimento, infecções subseqüentes, traumatismos cranianos, afogamento, envenenamento e entre outros (WESTERINEN *et al.*, 2017).

Ao exame físico, a DI pode clinicamente ser agrupada em DI Síndrômica (DIS) e DI Não-Sindrômica (DINS). Na DIS, pacientes apresentam características dismórficas, defeitos metabólicos ou comorbidades adicionais à DI. Enquanto que na DINS, é observada a presença da DI com ou sem anormalidades neurológicas, físicas e/ou metabólicas (AGHA *et al.*, 2014; TAN *et al.*, 2016). Contudo, o limite entre a classificação de DIS e DINS é geralmente arbitrário devido às características síndrômicas serem somente reconhecidas quando exames clínicos comparativos cuidadosos são realizados em vários pacientes com uma etiologia comum (VAN BOKHOVEN, 2011).

2.2 Aspectos Epidemiológicos da Deficiência Intelectual

A prevalência da DI é altamente variável dependendo do país estudado, da idade do paciente e do método de determinação do agravo. Nos países de alta renda, geralmente a prevalência é menor do que em países de baixa e média renda, sendo que pode ocorrer em qualquer classe social e cultura (KARAM *et al.*, 2015).

Na população em geral, a DI afeta aproximadamente 1–3%, com causas genéticas e não genéticas bem heterogêneas, impondo uma carga emocional e financeira acentuada às famílias afetadas, ao sistema de saúde e à sociedade (RAFATI *et al.*, 2012a; RAFATI *et al.*, 2012b; PREIKŠAITIENĖ *et al.*, 2016). Estudos de coortes em meta-

análise mostraram prevalências de DI em crianças e adolescentes entre 0,6–1,4% (WESTERINEN *et al.*, 2017).

Em torno de 75–90% dos casos, a DI é leve (MILANI *et al.*, 2015), sendo 30% mais frequente entre os homens (KAUFMAN *et al.*, 2010). As causas genéticas da DI são estimadas por volta de 25–50% (CHELLY *et al.*, 2006), com cerca de 25–30% por rearranjos cromossômicos e de 10–25% em virtude de pequenos rearranjos subteloméricos ou intersticiais (PRATTE-SANTOS *et al.*, 2016). Além disso, em torno de 30–50% dos casos de DI são idiopáticos ou não-sindrômicos (MILANI *et al.*, 2015).

Segundo a Pesquisa Nacional de Saúde (PNS, 2015) feita pelo IBGE (Instituto Brasileiro de Geografia e Estatística), cerca de 6,2% da população brasileira tem algum tipo de deficiência, sendo estas: auditiva, visual, física e/ou intelectual. Em relação à DI, em torno de 0,8% da população possui esse sinal fenotípico. No Brasil, as mulheres apresentaram proporção inferior à dos homens: 0,7% e 0,9%, respectivamente, não havendo diferença significativa entre as grandes regiões (Figura 2).

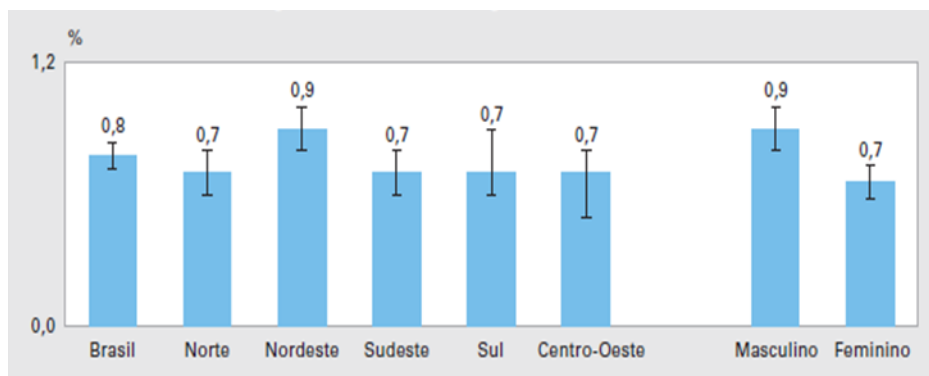


Figura 2. Proporção de pessoas com DI no território nacional brasileiro, de acordo com o IBGE. O gráfico mostra a proporção de pessoas com deficiência intelectual, na população total, com intervalo de confiança de 95%, segundo as grandes regiões geográficas do Brasil e o sexo.

Fonte: IBGE, 2015. Diretoria de Pesquisas, Coordenação de Trabalho e Rendimento. Pesquisa Nacional de Saúde 2013.

No Brasil, 0,5% da população total apresenta DI desde o nascimento, enquanto 0,3% é adquirida devido a doença ou acidente. Essa desordem foi a única, dentre as deficiências acima citadas, em que o percentual de pessoas que nasceu com ela foi superior ao percentual de pessoas que a adquiriu por acidente ou como comorbidade em função de uma doença (PNS/IBGE, 2015).

Dados do IBGE mostraram também que 54,8% da população com DI possuía nível intenso ou muito intenso de limitação, não conseguindo realizar suas atividades

habituais, tais como brincar, ir à escola e trabalhar. Ademais, cerca de 30% dessa população frequentam algum serviço de reabilitação em saúde (PNS/IBGE, 2015).

Karam e colaboradores (2015), em um estudo de coorte Brasileiro de uma população acompanhada prospectivamente desde o nascimento, concluíram que em torno de 20% dos casos de DI foram resultantes de alterações genéticas, sendo a síndrome de Down observada como a mais frequente das desordens cromossômicas que apresentam DI entre os seus sinais clínicos. Além disso, foi observado naquele estudo que as anormalidades cromossômicas foram responsáveis por aproximadamente 40% dos casos severos de DI e 10% dos casos leves de DI.

2.3 Etiologia da Deficiência Intelectual

A DI é um sinal fenotípico comum em um agrupamento de desordens heterogêneas, complexas e multifatoriais, que se apresentam de formas sindrômicas ou não-sindrômicas, que prejudicam as habilidades intelectuais das pessoas acometidas (VAN BOKHOVEN, 2011). De modo geral, a DI aparece como comorbidade acompanhando uma condição primária como o autismo, a síndrome de Down, a síndrome do X-Frágil, a paralisia cerebral, entre outras (GAMSIZ *et al.*, 2013; ASHA, 2017; CIACCIO *et al.*, 2017).

As causas da DI são bastante variáveis, podendo ser consequência de fatores ambientais, genéticos e epigenéticos, sendo que a DI pode aparecer como um traço isolado (DI não sindrômica) ou como uma deficiência associada com alterações físicas e funcionais (DI sindrômica) (MARIS *et al.*, 2013; BERTOK *et al.*, 2015; VISSERS *et al.*, 2016). Os fatores que prejudicam e interferem na maturação e crescimento do cérebro durante os períodos pré-natal, peri-natal e pós-natal são subjacentes à DI (BERTELLI *et al.*, 2016). Entretanto, devido à heterogeneidade clínica da DI ser reflexo da heterogeneidade genética, em torno de 40% dos casos de DI permanecem sem um diagnóstico molecular específico, que é crucial para o estabelecimento de risco recorrente e para a promoção do aconselhamento genético com as famílias dos pacientes (REDIN *et al.*, 2014; ROSELLÓ *et al.*, 2014; MICLEA *et al.*, 2015; VISSERS *et al.*, 2016).

Os fatores ambientais implicados com a etiologia da DI englobam: 1. Toxicidade – abuso materno de álcool, cocaína, anfetamina e outras drogas durante a gestação; 2. Uso de teratógenos; 3. Infecções congênitas; 4. Deficiências nutricionais extremas; 5. Complicações no parto; 6. Alterações metabólicas – hiperbilirrubinemia; e 7. Traumas

(KARAM *et al.*, 2015; MILANI *et al.*, 2015; BERTELLI *et al.*, 2016; VISSERS *et al.*, 2016).

As causas genéticas da DI incluem rearranjos cromossômicos (microscópicos ou submicroscópicos) que resultam principalmente de efeitos de doses de genes, CNVs, desequilíbrio de genes específicos imprintados ou de regiões genômicas, disfunção de um único gene, desordens genômicas e variações ligadas ao cromossomo X. Adicionalmente, alterações epigenéticas devido à variações genômicas que afetam genes envolvidos na regulação epigenômica e dissomia uniparental resultante da presença de longos trechos contíguos em homozigose, ocasionando homozigose cromossômica ou segmentar, contribuem para a etiologia da DI (ROSELLÓ *et al.*, 2014; IOUROV *et al.*, 2015; KARAM *et al.*, 2015; MILANI *et al.*, 2015; BERTELLI *et al.*, 2016; VISSERS *et al.*, 2016).

A presença de Longos Trechos Contíguos em Homozigose (LCSH) (do inglês, *Long Contiguous Stretches of Homozygosity*), pode afetar a expressão de diversos genes envolvidos na regulação epigenômica, por esses genes estarem em homozigose alélica e poderem alterar padrões de metilação de genes imprintados. A presença de LCSH pode ser indicativo de consanguinidade, dissomia uniparental ou mutação recessiva de um único gene (LI *et al.*, 2006; KIRIN *et al.*, 2010; IOUROV *et al.*, 2015).

No banco de dados *online* de Herança Mendeliana em Humanos (OMIM) (do inglês, *Online Mendelian Inheritance in Man*), mais de 1.000 condições genéticas já foram descritas com a DI incluída como um componente fenotípico e mais de 50 síndromes que também cursam com DI já foram descritas em associação somente com o cromossomo X (NOWAKOWSKA, 2017).

Com o desenvolvimento de novas tecnologias, como análise cromossômica por microarranjos e sequenciamento de nova geração, a promessa é o aumento da compreensão da etiologia da DI, com a identificação de genes e mecanismos que contribuem para o desenvolvimento deste sinal fenotípico tão evidente na fenotipagem dos pacientes (ELLISON *et al.*, 2013; REDIN *et al.*, 2014).

2.3.1 Genética e aspectos moleculares da deficiência intelectual

O cérebro é um órgão complexo com uma grande quantidade de tipos celulares interconectados. Durante o desenvolvimento e o funcionamento diário do cérebro, numerosas proteínas necessitam estarem funcionalmente ativas na quantidade, no espaço

e no tempo correto. Mutações e rearranjos que afetam qualquer um dos genes que codificam essas proteínas poderão ter graves consequências para o desenvolvimento do cérebro ou o seu funcionamento cognitivo (VISSERS *et al.*, 2016).

Eventos genéticos únicos são responsáveis pela maioria dos casos de DI, que variam de mutação de ponto em um único gene a variações no número de cópias afetando vários genes ou mesmo um cromossomo inteiro. As formas monogênicas são caracterizadas por uma heterogeneidade genética extrema, com centenas de genes implicados à DI ligada ao X e vários outros associados às formas autossômicas recessivas ou autossômicas dominantes (REDIN *et al.*, 2014).

Rearranjos cromossômicos são considerados a causa mais frequente de DI, AGD e anomalias congênitas múltiplas, explicando mais de 25–30% dos casos e associados com cerca de 25% das mortes perinatais (RAFATI *et al.*, 2012a; CAPALBO *et al.*, 2017). Alterações citogeneticamente visíveis ao microscópio são observadas em torno de 15% dos casos de DI, sendo as mais frequentes a trissomia do cromossomo 21 (aneuploidia cromossômica mais comum associada à DI), a trissomia do 13 e a trissomia do 18 (MICLEA *et al.*, 2015). Outras alterações microscopicamente visíveis incluem deleções, duplicações, translocações e cromossomos marcadores supranumerários. Todavia, rearranjos cromossômicos críticos, particularmente rearranjos subteloméricos e intersticiais, como por exemplo deleções e duplicações cromossômicas menores que 5 Mb, têm se mostrado como uma significativa causa da DI, respondendo por mais de 30% dos casos (CHELLY *et al.*, 2006; ELLISON *et al.*, 2013; MICLEA *et al.*, 2015).

Frequentemente, pacientes com alterações cromossômicas apresentam características clínicas adicionais ao sinal fenotípico da DI, como anormalidades congênitas múltiplas, alterações dismórficas e atraso no crescimento. Além disso, há uma alta incidência de abortamentos e infertilidade em familiares desses pacientes, e nascimentos prematuros com malformações múltiplas, que usualmente morrem logo ao nascerem, elevando a taxa de mortalidade pré- e perinatal (ROSELLÓ *et al.*, 2014). Contudo, as consequências clínicas dos rearranjos cromossômicos estão geralmente relacionadas com a localização do rearranjo, a quantidade de genes envolvidos e sua função (PRATTE-SANTOS *et al.*, 2016).

Alterações no cromossomo X representam 5–10% do total de casos com DI, dos quais 30% mais homens são afetados do que mulheres (MUTHUSAMY *et al.*, 2017), sendo as formas de DI ligadas ao X facilmente identificadas pelo característico padrão de

herança mendeliano (MOYSÉS-OLIVEIRA *et al.*, 2015). Deficiência intelectual ligada ao X é geneticamente heterogênea e fenotipicamente variável, sendo que há mais de 200 formas sindrômicas ou não sindrômicas de DI ligada ao X. Além disso, têm sido identificadas mutações em mais de 100 genes ligados ao X. Contudo, nenhum dos genes atualmente implicados explicam mais de 0,1% das DI, mas, coletivamente, eles explicam mais de 10% da DI em homens (TAN *et al.*, 2016; VISSERS *et al.*, 2016; GIORGIO *et al.*, 2017; MUTHUSAMY *et al.*, 2017).

Causas monogênicas ligadas ao X são responsáveis por mais de 10% de casos de DI. A Síndrome do X-Frágil (FXS) (do inglês, *Fragile X Syndrome*) é o defeito monogênico mais comum (observada em 5% dos pacientes com DI), sendo a segunda causa de DI e a primeira de DI herdada. É a causa de DI mais prevalente em homens, considerada o primeiro tipo de DI que pode ser explicado por um defeito em um gene específico (MICLEA *et al.*, 2015; CIACCIO *et al.*, 2017). O diagnóstico é baseado na detecção de uma alteração na região promotora do gene *FMRI*, localizado em Xq27.3, sendo que 99% dos indivíduos com FXS tem perda de função do gene *FMRI*, causada pela mutação caracterizada pela expansão de uma zona de repetição de um trinucleotídeo (CGG) na região 5' não traduzida (tipicamente > 200 tripletes) do gene (CIACCIO *et al.*, 2017). Outro gene ligado ao X observado como causativo para DI é o *MECP2*, originalmente implicado na Síndrome de Rett (ELLISON *et al.*, 2013).

Em contrapartida, a prevalência de alterações nos cromossomos autossômicos em pacientes com DI varia de 16–55%. Enquanto que DI autossômica recessiva é menos prevalente, a estimativa da contribuição de mutações recessivas na DI é estimada acima de 25%, sendo a maioria em famílias consanguíneas (AGHA *et al.*, 2014; TAN *et al.*, 2016). Mais de 300 genes têm sido identificados em formas autossômicas recessivas de DI, na maioria por mapeamento de homozigose usando polimorfismo de Variante de Nucleotídeo Único (SNV) (do inglês, *Single Nucleotide Variants*) (VISSERS *et al.*, 2016).

Em torno de 50–60% das microdeleções e microduplicações intersticiais são identificadas em pacientes com DI e anormalidades congênicas múltiplas. Metade dessas alterações genômicas está localizada em regiões comuns de síndromes de microdeleção e microduplicação em autossomos, como a Síndrome de DiGeorge (deleção em 22q11), Síndrome de Prader-Willi (deleção de origem paterna em 15q11.2), Síndrome de Angelman (deleção de origem materna em 15q11.2), Síndrome de William-Beuren

(deleção em 7q11.23), Síndrome de Smith-Magenis (deleção em 17p11.2), Síndrome de Miller-Dieker (deleção em 17p13.3), Síndrome de Deleção 1p36 e Síndrome de Sotos (deleção em 5q35) (RAFATI *et al.*, 2012a; CLASSEN *et al.*, 2013; MILANI *et al.*, 2015; VISSERS *et al.*, 2016) (Figura 3).

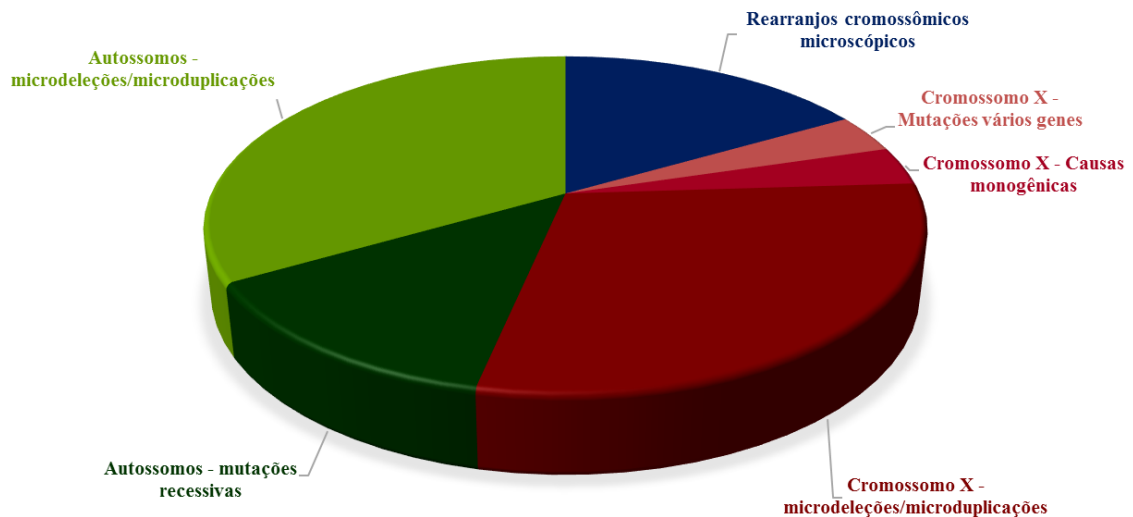


Figura 3. Causas Genéticas da DI. Gráfico representado pelos 30% dos rearranjos cromossômicos considerados a causa mais frequente de DI, TEA e ACM. Destes, 15% são por rearranjos cromossômicos microscópicos (azul escuro) e 85% são por rearranjos cromossômicos submicroscópicos (evidenciados pelo restante das cores).

Outro grupo de desordens do desenvolvimento neurológico que tem sido observado altamente associado com DI é o TEA. Estimativas populacionais apontam que até 75% dos pacientes com diagnóstico de autismo apresentam concomitantemente a DI. Mutações em um grande número de genes têm sido identificadas em pacientes com TEA e DI, incluindo mutações em *IL1RAP1*, *SHANK2*, *SHANK3*, *NLGN3*, *NLGN4*, *GRIN2B*, *TCF4*, *AUTS2*, *CNTN4*, *CNTNAP2*, que codificam proteínas sinápticas. Alterações *de novo* em SNV são encontradas em 8–20% dos pacientes com autismo e DI (VAN BOKHOVEN, 2011; HARRIS; GREENSPAN, 2016; GRIESI-OLIVEIRA; SERTIÉ, 2017).

As CNVs, definidas como um intervalo genômico que desvia do estado diploide normal, apresentam segmentos de DNA > 50 pb (pares de base), presentes em torno de 12 a 16% do genoma humano, podendo estar associadas com doenças ou representar a diversidade fenotípica de cada indivíduo. O termo desordem genômica é considerada um grupo de doenças genéticas causadas por rearranjos no DNA resultante de CNVs de

ganho, perda ou por ruptura de genes no qual sua dosagem se torna crítica (LEE *et al.*, 2007; CAPALBO *et al.*, 2017; HAREL; LUPSKI, 2018).

As CNVs não estão distribuídas no genoma humano randomicamente. Porém, elas tendem a se agrupar em regiões de arquitetura genômica complexa ricas em repetições de poucas cópias (LCR) (do inglês, *Low Copy Repeats*) ou duplicações segmentais (SD) (do inglês, *Segmental Duplication*). Alguns destes agrupamentos podem estar relacionados à ausência de genes sensíveis à dosagem, mas há ampla evidência de que a arquitetura cromossômica específica também esteja envolvida. As CNVs benignas geralmente são pequenas, intergênicas ou englobando genes que toleram uma mudança no número de cópias, enquanto que as CNVs patogênicas são consideravelmente enriquecidas por genes envolvidos no desenvolvimento e genes com restritos padrões evolutivos de ganhos e perdas (HASTINGS *et al.*, 2009; HAREL; LUPSKI, 2018).

Além das CNVs com consequências clínicas já bem estabelecidas e conhecidas, CNVs recorrentes raras presentes em regiões cromossômicas específicas têm sido implicadas à fatores de risco para doenças do desenvolvimento neurológico. Ademais, muitas dessas CNVs possuem penetrância incompleta e expressividade variável que podem ser influenciadas por outros fatores genéticos, epigenéticos e ambientais (HAREL; LUPSKI, 2018).

As CNVs são classificadas, de acordo com seu mecanismo de formação, em dois grupos: CNVs recorrentes e CNVs não-recorrentes. As CNVs recorrentes, compartilham pontos de quebra, conteúdo genômico e tamanho em indivíduos não relacionados, frequentemente, surgem por um mecanismo de recombinação conhecido como Recombinação Homóloga Não-Alélica (NAHR) (do inglês, *Nonallelic Homologous Recombination*), mediada por blocos duplicados que flanqueiam o evento de CNV. Em contrapartida, as CNVs não recorrentes têm pontos de quebra que geralmente não resultam de uma arquitetura genômica predisponente, possuem tamanhos únicos e conteúdo genômico distintos em indivíduos não relacionados, sendo que surgem por diversos mecanismos de recombinação, incluindo Junção de Extremidades Não Homólogas (NHEJ) (do inglês, *Nonhomologous End Joining*) (CAPALBO *et al.*, 2017; HAREL; LUPSKI, 2018). Desordens genômicas, causadas por rearranjos de perda ou de ganho genômicos, usualmente possuem tamanho < 5Mb, abrangem vários genes e podem ser herdadas ou esporádicas, dependendo se o rearranjo foi transmitido através da linhagem germinativa ou se ocorreu *de novo* (CAPALBO *et al.*, 2017).

O principal mecanismo responsável pela formação de CNVs é a NAHR, tanto durante a meiose quanto a mitose, mediado por dois segmentos parálogos de DNA repetitivo com similaridade $\geq 90\%$, conhecidos como repetições de poucas cópias (LCR) ou duplicações segmentais (SD), que flanqueiam o segmento genômico rearranjado (DITTWALD *et al.*, 2013; HAREL; LUPSKI, 2018). A NAHR pode ocorrer tanto entre cromossomos distintos, sejam eles homólogos ou não, quanto entre diferentes loci do mesmo cromossomo. Quando acontece NAHR mediada por LCRs que estão no mesmo sentido circundando o local da recombinação, há a formação de CNVs de perda ou de ganho. Por outro lado, o rearranjo do tipo inversão pode acontecer por NAHR mediada por LCRs que estão em sentido invertido flanqueando o local da recombinação (Figura 4) (INOUE; LUPSKI, 2002; LEE *et al.*, 2007; GU *et al.*, 2008; KIM *et al.*, 2008).

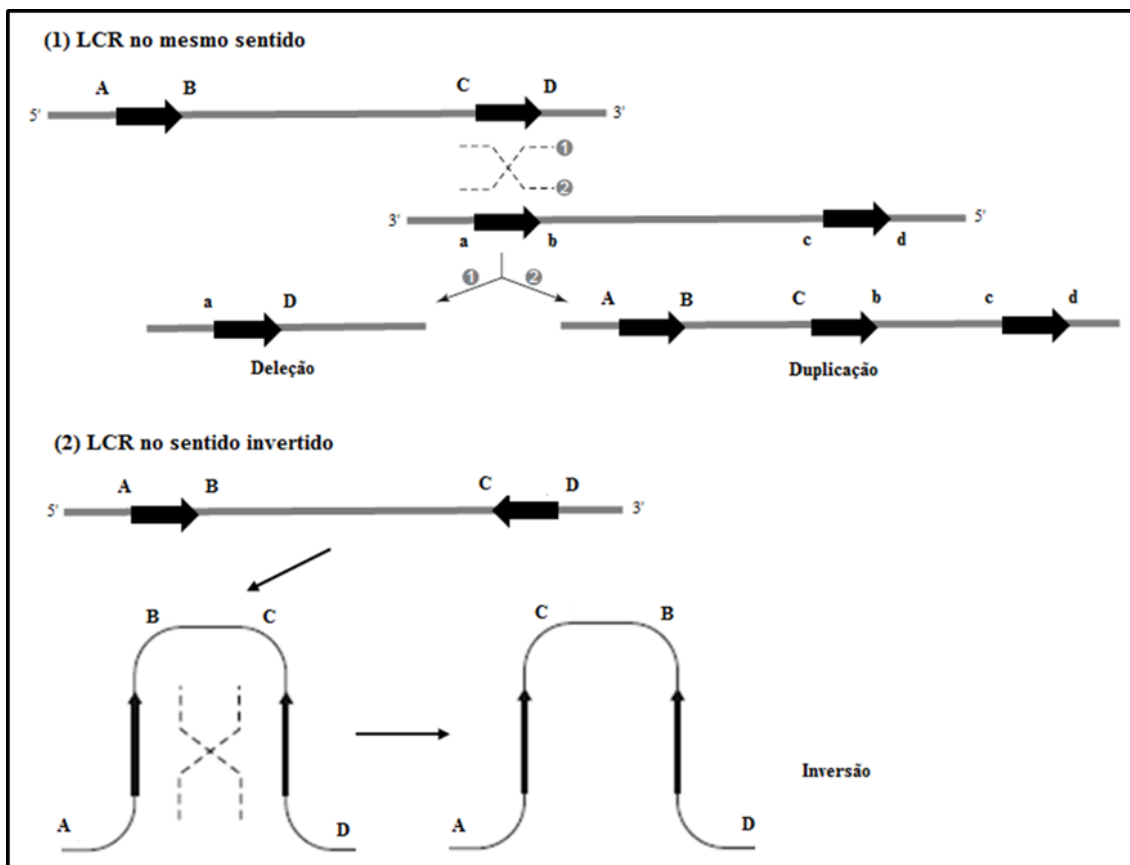


Figura 4. Rearranjos genômicos resultantes de recombinação homóloga não-alélica (NAHR) mediada por LCRs. As LCRs estão representadas pelas flechas pretas com a orientação indicada pelo sentido da flecha. As linhas tracejadas referem ao evento de recombinação com os resultados desse evento mostrados pelos números 1 e 2. (1) NAHR entre LCRs que estão no mesmo sentido resultando em perdas e/ou ganhos. (2) NAHR entre LCRs no sentido inverso resultando em inversão.

Fonte: Adaptado de LUPSKI, 1998; GU *et al.*, 2008; HAREL; LUPSKI, 2018.

A presença de CNVs no genoma é uma importante fonte de variantes normais e patogênicas. Em algumas situações, as CNVs podem ser neutras, em outra podem ser cooptadas ao longo de nossa evolução e em outras elas podem ser CNVs deletérias, sendo um grande desafio estimar se essa variação é benigna ou afeta o funcionamento biológico vital resultando em doença. Há um esforço constante na identificação de CNVs que não produzem consequências fenotípicas, mas que apenas fazem parte da diversidade do genoma humano e que podem ajudar na interpretação de achados de CNVs tanto para o diagnóstico clínico quanto para as aplicações em pesquisa (NOWAKOWSKA, 2017). Da mesma maneira, CNVs podem impactar na expressão gênica indiretamente através do efeito de posição, revelando mutação recessiva ou alterando a comunicação entre os alelos por deleção de elementos regulatórios, levando a inativação ou perda de função do gene (NOWAKOWSKA, 2017; HAREL; LUPSKI, 2018).

As CNVs clinicamente relevantes são encontradas em cerca de 15–25% de crianças com inexplicável DI, AGD, TEA e ACM (WECKSELBLATT; RUDD, 2015; NOWAKOWSKA, 2017). CNVs microscópicas e submicroscópicas estão entre as causas genéticas mais comuns da DI e AGD (BERTOK *et al.*, 2015). A maioria das CNVs afetam um grande número de genes, podendo alterar os níveis de expressão dos genes incluídos nestas regiões, permitindo níveis de transcrição mais altos ou inferiores aos que podem ser alcançados pelo controle de transcrição de cópias únicas por genoma (HASTINGS *et al.*, 2009). Ganhos e perdas recorrentes são responsáveis por aproximadamente 20% dos rearranjos patogênicos intracromossômicos (WECKSELBLATT; RUDD, 2015). Além disso, interrupção direta de um gene por um ponto de quebra de CNV e deleções presentes em regiões intrônicas que podem interferir no processo de *splicing*, levando à inativação ou perda de função deste gene (HAREL; LUPSKI, 2018).

O processo de *splicing*, reconhecido como um eficiente e preciso processo de reconhecimento e remoção dos íntrons e ligação dos exons adjacentes, ocorre na maioria dos genes dos eucariotos, podendo ser do tipo constitutivo ou alternativo (CHABOT; SHKRETA, 2016; NIK; BOWMAN, 2019). Os eventos de *splicing* ocorrem através de duas reações bioquímicas de transesterificação, que são catalisadas por uma sofisticada maquinaria ribonucleoprotéica, conhecida como *spliceossomo*, composta por 5 pequenas ribonucleoproteínas (snRNP) (do inglês, *Small Nuclear Ribonucleoproteins*) (U1, U2, U4, U5, U6) e mais de 300 proteínas diferentes (FREDERICKS *et al.*, 2015; GRODECKÁ *et al.*, 2017; ANNA; MONIKA, 2018; GALEJ, 2018).

O processamento do pré-mRNA acontece através de sítios de *splice* (ss), que são sequências conservadas ao redor das junções entre exons e íntrons. Os limites dessas junções são delimitados por pequenas sequências consenso no sítio de *splice* 5' (GU) e sítio de *splice* 3' (AG), que mediam o reconhecimento pelo *spliceossomo*. Ademais, as ribonucleoproteínas *spliceossomais* também interagem com duas regiões nos íntrons que são primordiais, a sequência de ponto de ramificação (BPS) (do inglês, *Branch Point Sequence*), que contém uma Adenina (A) catalítica que promove formação do laço, e uma região rica em pirimidinas (Y) que está próxima ao 3'ss do íntron (KATAOKA, 2017; VAZ-DRAGO *et al.*, 2017; NIK; BOWMAN, 2019; YAN *et al.*, 2019).

A primeira reação de transesterificação, com objetivo de reconhecimento do sítio de *splicing* entre a junção íntron/exon, a 2'-OH da adenina (A) presente no BPS “ataca” o 5'-fosfato da guanina presente no 5' ss, resultando na formação de um laço intrônico intermediário com ligação 2',5'-fosfodiéster. Em seguida, uma segunda reação de transesterificação ocorre onde o 3'ss é “atacado” pelo grupo 3'-OH do nucleotídeo do 5'ss-A, promovendo a remoção do laço intrônico e a ligação de exons. O reconhecimento pelo *spliceossomo* das sequências consensus presentes em 5'ss, 3'ss e BPS é uma etapa crucial para os eventos de *splicing* que são modulados por potenciadores de *splicing* intrônicos (ISE) (do inglês, *Intronic Splicing Enhancer*) e potenciadores de *splicing* exônicos (ESE) (do inglês, *Exonic Splicing Enhancer*), além de silenciadores de *splicing* intrônicos (ISS) (do inglês, *Intronic Splicing Silencers*) e silenciadores de *splicing* exônicos (ESS) (do inglês, *Exonic Splicing Silencers*), que são reconhecidos por auxiliarem os fatores de *splicing*, incluindo proteínas ricas em serina/arginina (SR) e nRNPs (Figura 5) (SCOTTI; SWANSON, 2016; SHI 2017; GALEJ, 2018).

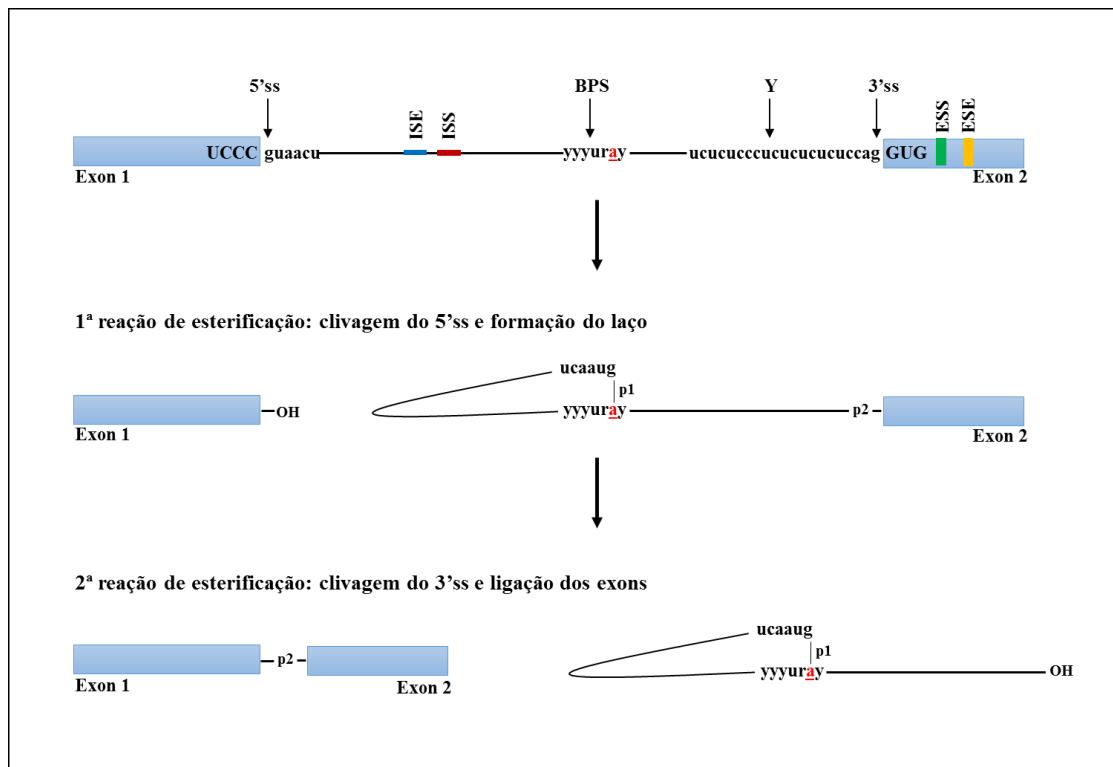


Figura 5. Processamento do pré-mRNA. Representação esquemática dos sítios de *splicing*, moduladores dos *spliceossomos* e reação de transesterificação com remoção dos íntrons e ligação dos exons.

Fonte: Adaptado de WILL; LÜHRMANN, 2011; VAZ-DRAGO *et al.*, 2017.

Genes responsáveis pelas alterações no desenvolvimento têm sido identificados presentes em regiões crípticas de CNVs patogênicas (ELLISON *et al.*, 2013). Em alguns casos, porém, um único gene sensível à dosagem pode ser identificado como a causa do sinal fenotípico da DI, como o gene *RAI1* na Síndrome de Potocki-Lupski, o *CHD7* na Síndrome CHARGE, o *MECP2* na Síndrome Duplicação em Xq28, o *NSD1* na Síndrome de Sotos, o *SHANK3* na Síndrome de Phelan-McDermid, o *UBE3A* na Síndrome de Angelman dentre outros (VAN BOKHOVEN, 2011; ELLISON *et al.*, 2013; VISSERS *et al.*, 2016; NOWAKOWSKA, 2017).

Em torno de 1.416 genes responsáveis pelo sinal fenotípico da DI isolado ou em desordens associadas à DI já foram identificados ligados ao X, autossômicos dominantes e autossômicos recessivos (REUTER *et al.*, 2017). Mutações *de novo* em genes associados à DI ocorrem entre 13–35% dos casos de DI severa (VISSERS *et al.*, 2016).

Uma quantidade altamente variável de informações está disponível em diferentes genes, podendo elucidar a probabilidade e possibilidade de causar doenças genéticas quando mutados. Nos últimos anos, mutações em vários genes que podem resultar em DI sindrômica ou não sindrômica têm sido identificadas e caracterizadas.

Essas mutações podem ser altamente impactantes na função dos genes e tendem a afetar importantes domínios de genes do desenvolvimento, sendo que a maioria dos genes afetados está localizada no cromossomo X (VELTMAN; BRUNNER, 2012; VERPELLI *et al.*, 2014).

Ademais, mais de 50% das proteínas relacionadas à DI que não são fatores de transcrição ou de remodelamento de cromatina, estão claramente presentes nos compartimentos pré ou pós-sinápticos e estão implicadas na função sináptica devido à regulação do rearranjo do citoesqueleto de actina, plasticidade sináptica ou formação da sinapse. E também, há outros genes que codificam proteínas que regulam a integração do circuito neuronal e a correção da rede cerebral (VERPELLI *et al.*, 2014).

2.4 Diagnóstico Genético da Deficiência Intelectual

Atualmente, existem várias técnicas capazes de elucidar a alteração cromossômica estrutural ou numérica e alterações genômicas que podem estar presentes no paciente que apresenta o sinal fenotípico da DI e AGD. Os testes usados para a investigação das causas genéticas da DI e AGD incluem o Cariótipo por bandeamento GTG, a Análise Cromossômica por Microarranjo, Pesquisa para a Síndrome do X-Frágil, Sequenciamento de um único gene ou painéis de genes. Além disso, o Sequenciamento do Genoma Total ou de Exoma podem fornecer uma identificação ampla e de alta resolução das variantes genéticas, sendo uma promessa como ensaios diagnósticos eficientes, apesar da limitação pelos altos preços destes exames e da produção de uma grande quantidade de informação que ainda não pode ser adequadamente interpretada do ponto de vista clínico (BOWLING *et al.*, 2017).

O cariótipo por bandeamento GTG, uma técnica simples e barata, tem sido há mais de 35 anos o teste padrão para a busca de desequilíbrios cromossômicos em pacientes com DI e AGD (MILANI *et al.*, 2015), com uma capacidade de evidenciar alterações cromossômicas em 3–5% dos casos de pacientes com DI e AGD (BARTNIK *et al.*, 2014). Para a realização desta técnica, na presença de um meio de cultura específico, os cromossomos são estimulados ao processo de divisão celular, posteriormente, corados e então avaliados com o auxílio de um microscópio. O cariótipo é capaz de detectar alterações estruturais, como deleções e duplicações > 5 Mb, translocações, inversões, além de aneuploidias. Contudo, um número relevante de afecções clínicas, incluindo a DI e AGD, ocorrem por rearranjos cromossômicos, que

envolvem ganhos e perdas genômicos com tamanho < 5 Mb, não sendo detectados por esta metodologia (WECKSELBLATT; RUDD, 2015; VISSERS *et al.*, 2016).

A expansão trinucleotídica CGG, com > 200 repetições, observada na região promotora do gene *FMRI* é capaz de alterar a expressão deste gene e causar a Síndrome do X-Frágil, que apresenta o sinal fenotípico da DI além de outras alterações fenotípicas. Atualmente, para a determinação da expansão trinucleotídica CGG, a amplificação por PCR da região promotora do gene *FMRI* e posteriormente a determinação da quantidade de repetições CGG por eletroforese capilar tem sido um método eficaz e acessível para os pacientes que apresentam DI (GIGONZAC *et al.*, 2016).

Nos últimos anos, novas tecnologias têm sido desenvolvidas, criando e ampliando um novo campo de investigação que transformou a prática da genética clínica. O CMA é uma técnica que está sendo cada vez mais utilizada como teste genético para pacientes com DI/AGD idiopáticos, TEA e ACM. Esta técnica utiliza marcadores não polimórficos e marcadores polimórficos (SNP), que são capazes de detectar, com alta resolução, perdas ou ganhos no genoma com tamanhos de 1kb a milhões de pares de bases, utilizando um único chip (MILLER *et al.*, 2010; ŽILINA *et al.*, 2014; WECKSELBLATT; RUDD, 2015; NOWAKOWSKA, 2017). Além disso, os marcadores de SNP possibilitam a detecção de longos trechos contínuos em homozigose (LCSH), que pode ser indicativo de identidade por descendência se estiverem amplamente distribuídos por todo o genoma, ou indicativo de dissomia uniparental (UPD) se estiverem presentes em um único cromossomo (HAREL; LUPSKI, 2018).

Para fazer a classificação das CNVs visualizadas pelo CMA, utiliza-se as recomendações dos consensos propostos por Miller e colaboradores (2010) e Battaglia e colaboradores (2013), e também das diretrizes do Colégio Americano de Genética Médica (ACMG) (do inglês, *American College of Medical Genetics*) (KEARNEY *et al.*, 2011) e das diretrizes Europeias para análise citogenômica constitucional (SILVA *et al.*, 2019).

Além disso, utiliza-se como referência os bancos de dados públicos que contém informações sobre indivíduos com diferentes fenótipos clínicos. Há dois tipos de bancos de dados: (1) os que contém informações de dados populacionais de indivíduos saudáveis, como por exemplo o Bancos de Dados de Variantes Genômicas (DGV) (do inglês *Database of Genomic Variants*); e (2) os bancos de dados que apresentam detalhes genéticos e fenotípicos de indivíduos, como exemplos temos o OMIM, o Banco de Dados de Imbalanços Cromossômicos e Fenótipos em Humanos usando Recursos Ensemble

(DECIPHER) (do inglês *Database of Chromosomal Imbalance and Phenotype in Humans using Ensemble Resources*), o Banco de Dados da Associação Europeia de Registros Citogenéticos de Aberrações Cromossômicas Não Balanceadas (ECARUCA) (do inglês *European Cytogeneticists Association Register of Unbalanced Chromosome Aberrations*) e o Banco de Dados de Padrões Internacionais para Arranjos Citogenômicos (ISCA) (do inglês *International Standards for Cytogenomic Arrays*) (CAPALBO *et al.*, 2017).

Uma significativa proporção de indivíduos com DI, porém, permanece sem o diagnóstico genético, provavelmente por apresentarem CNVs patogênicas bem menores que a capacidade de resolução de todos os arrays, que podem ocorrer e afetar regiões ou genes que estão relacionados com o sinal fenotípico da DI (WECKSELBLATT; RUDD, 2015; NOWAKOWSKA, 2017; ZAHIR *et al.*, 2017). Portanto, o uso do filtro de tamanho < 100 kb no CMA pode ser uma alternativa para identificação de CNVs muito pequenas que abrigam genes importantes relacionados à DI. Adicionalmente, o uso do sequenciamento de Sanger poderá auxiliar no mapeamento dos pontos de quebras e no entendimento dos mecanismos que levaram a formação destas CNVs. Além disso, o Sequenciamento de Sanger é, apesar do advento das técnicas de Sequenciamento de Nova Geração (NGS) (do inglês, *Next Generation Sequencing*), considerado o padrão-ouro para o diagnóstico molecular de doenças genéticas (STROM *et al.*, 2014; ZAHIR *et al.*, 2017).

3. JUSTIFICATIVA

A deficiência intelectual pode ocorrer de forma isolada ou combinada com outras malformações congênitas, características neurológicas, e problemas comportamentais, sendo caracterizada por uma condição fenotipicamente e geneticamente complexa, com diagnóstico diferencial bastante desafiante, podendo em muitos casos ainda não ser alcançado (VASUDEVAN; SURI, 2017).

A aplicação do CMA na prática clínica tem aumentado em 15–25 % o diagnóstico das condições em que há DI, AGD e TEA e epilepsia, tornando-se o espectro de diagnóstico mais completo que abrange grande variedade de fenótipos e suas gravidades para diversas e diferentes patologias (HOWELL *et al.*, 2013; BARTNIK *et al.*, 2014; JANG *et al.*, 2019). Entretanto, dentro do universo dos pacientes com DI, nem todos terão o diagnóstico concluído, mesmo após a combinação das metodologias da citogenética e da citogenômica, como o bandeamento GTG e o CMA, e da genética molecular, como a PCR para a FRAX (BRETT *et al.*, 2014).

Após a análise com filtro de tamanho ≥ 100 kb recomendado pelo fabricante e identificação e classificação das CNVs de acordo com os consensos de Miller e colaboradores (2010), Battaglia e colaboradores (2013) e as diretrizes do Colégio Americano de Genética Médica (KEARNEY *et al.*, 2011), das diretrizes Europeias para análise citogenômica constitucional (SILVA *et al.*, 2019), e também com os diferentes bancos de dados, incluindo DGV, DECIPHER, CytoScan™ HD Array Database, uma significativa proporção de indivíduos com DI, porém, permanece sem o diagnóstico citogenômico, provavelmente por apresentarem CNVs bem menores, que podem ocorrer e afetar regiões ou genes que estão relacionados ao sinal fenotípico da DI, sendo geralmente menores que a capacidade de resolução de todos os arrays (WECKSELBLATT; RUDD, 2015; NOWAKOWSKA, 2017; ZAHIR *et al.*, 2017).

A utilização do filtro de tamanho < 100 kb na técnica de CMA pode aumentar a capacidade diagnóstica, aumentando o espectro de detecção de CNVs muito pequenas que estão presentes próximos ou dentro de genes importantes que são responsáveis pela manifestação da DI, contribuindo para elucidação da causa genética envolvida no fenótipo do paciente. Porém, para identificação dos pontos de quebra destas CNVs é recomendado o uso do sequenciamento de Sanger, possibilitando a obtenção de uma maior sensibilidade e especificidade ao ensaio (STROM *et al.*, 2014; ZAHIR *et al.*, 2017)

O mapeamento dos pontos de quebra das CNVs muito pequenas, através do uso da metodologia de sequenciamento de Sanger, é uma forma de contribuir para uma melhor compreensão do mecanismo responsável pela formação destes rearranjos e também da genética da DI, possibilitando a elucidação diagnóstica e o estabelecimento de caracterizações e prognósticos diferenciais desse transtorno altamente complexo e heterogêneo, além de contribuir para o fortalecimento do desenvolvimento biotecnológico regional e nacional e a atenção aos pacientes atendidos pelo SUS.

Por fim, o conhecimento dos genes envolvidos e as informações geradas a partir de uma avaliação completa, utilizando diferentes metodologias como cariótipo, CMA e sequenciamento, auxiliam na elucidação da causa da DI, podendo ser usada para compreender a biologia e os mecanismos moleculares subjacentes ao desenvolvimento da DI. Isto por sua vez, ajudará na conduta do melhor tratamento e contribuirá na evolução e no delineamento de um plano de reabilitação adequado. Além disso, ajudará as famílias a entenderem a condição que suas crianças apresentam, permitindo que tenham acesso a educação especial e ao aconselhamento genético preciso com informações sobre chances de recorrência.

4. OBJETIVOS

4.1 Objetivo Geral

Avaliar a ocorrência de CNVs, identificadas pelo CMA com filtro de tamanho < 100 kb, que abrigassem genes funcionalmente associados à deficiência intelectual em pacientes do SUS com diagnóstico clínico de DI encaminhados para o diagnóstico genético laboratorial.

4.2 Objetivos Específicos

- Realizar o cariótipo por bandeamento GTG, em indivíduos com sinal fenotípico de DI, para identificar alterações cromossômicas numéricas ou estruturais;
- Realizar o CMA, após o cariótipo, em pacientes com DI que apresentarem cariótipo sem alterações numéricas e estruturais;
- Identificar CNVs menores que o tamanho de 100 kb, que abrigam genes relacionados ao desenvolvimento da DI, a partir de uma seleção feita após a realização do CMA com filtro de tamanho reduzido;
- Identificar pontos de quebra, por Sequenciamento de Sanger, das regiões de CNVs de perda onde os genes selecionados estivessem presentes.

5. MATERIAL E MÉTODOS

5.1 Delineamento do estudo

Foi realizado um estudo transversal de população afetada de jan/2013 a dez/2016, conduzido no Núcleo de Pesquisas Replicon da Pontifícia Universidade Católica de Goiás (NPR/PUC-Goiás) e no LaGene/LACEN/SES-GO, em parceria com Laboratório Centro de Genômica de Alto Desempenho do Distrito Federal da Universidade Católica de Brasília (UCB), em pacientes com indicação clínica de DI/AGD para diagnóstico genético laboratorial.

5.2 Grupo Amostral

Foram analisados pacientes de ambos os sexos, com indicação clínica de DI/AGD, encaminhados por médicos assistentes das redes públicas de saúde de Goiás. Após o esclarecimento sobre o estudo, os pacientes e seus pais ou responsáveis legais foram convidados a participar e aqueles que concordaram, assinaram o termo de consentimento livre e esclarecido (TCLE) do estudo (Anexo 1) e também o TCLE do banco de dados DECIPHER (Anexo 2). Além disso, foi aplicado um questionário para coleta dos dados clínicos.

Após a realização do cariótipo e do CMA, foram selecionados aqueles pacientes que apresentaram o cariótipo sem alteração e com CMA não demonstrando CNVs informativas no filtro de tamanho ≥ 100 kb, para proceder a análise dos dados do CMA com filtro de tamanho < 100 kb e posteriormente sequenciamento de Sanger para identificar os pontos de quebras das regiões de CNVs de perda encontradas que continham genes associados à DI (Figura 6).

A inclusão dos participantes seguiu os seguintes critérios: pacientes com indicação clínica de DI encaminhados via SUS, cujos pais e/ou responsáveis concordaram voluntariamente em assinar os TCLEs. Foram excluídos pacientes com diagnóstico já elucidado ou cujos pais e/ou responsáveis não concordaram em assinar os TCLEs.

A presente proposta foi aprovada pelo Comitê de Ética em Pesquisa (CEP) da PUC-Goiás (1721/2011). Toda a documentação e as amostras foram arquivadas no NPR/PUC-Goiás e pode ser disponibilizada sob pedido formal. Todos os protocolos foram conduzidos de acordo com critérios internacionais de qualidade e ética em pesquisa com seres humanos.

5.3 Fluxograma de atendimento

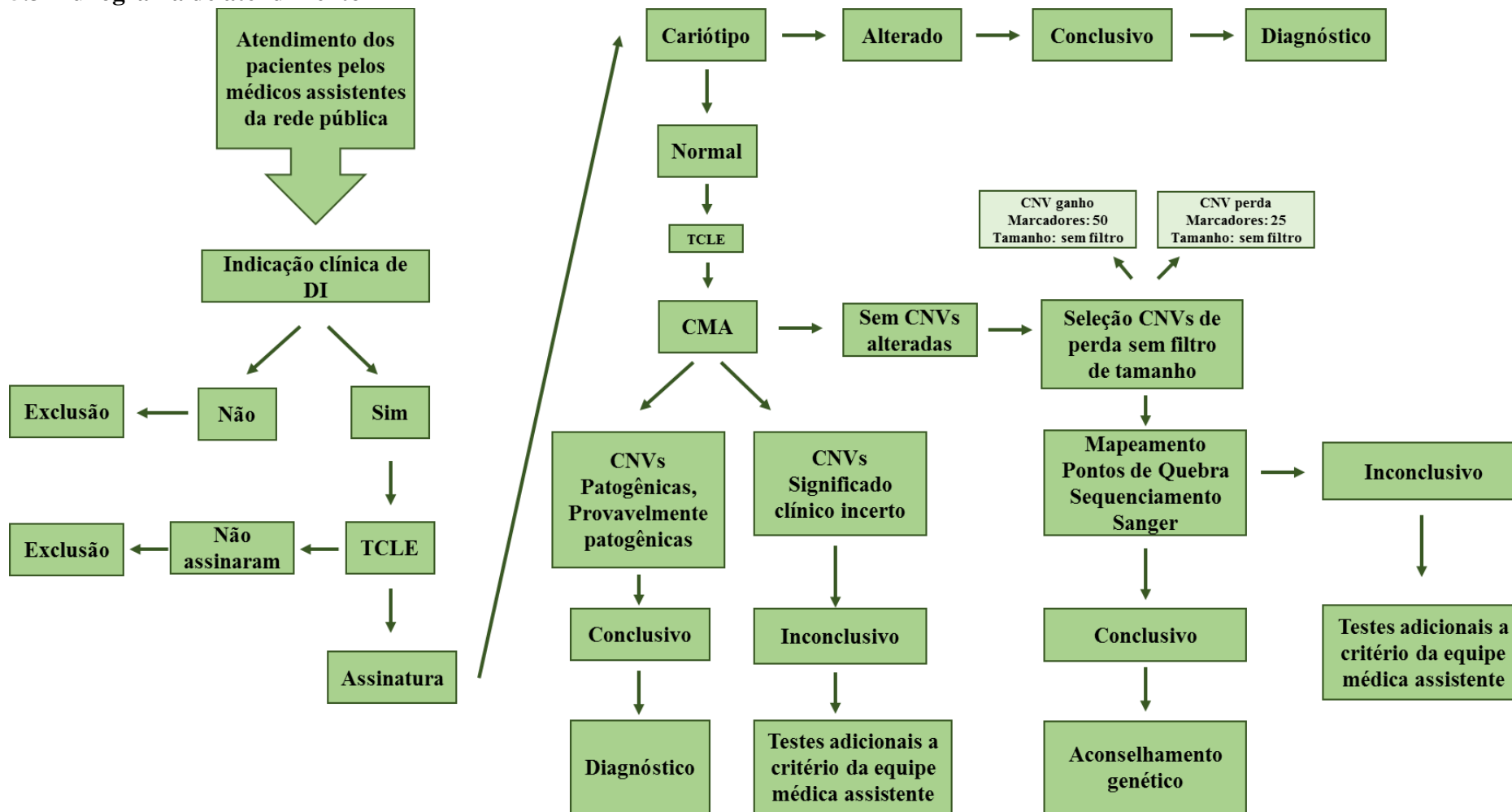


Figura 6. Fluxograma indicando o encaminhamento, seleção, triagem e metodologias utilizadas na tentativa de elucidação diagnóstica de casos de DI.

5.4 Coleta e processamento das amostras

Após a assinatura dos TCLEs, foram obtidas as amostras biológicas dos pacientes para realização do cariótipo por bandeamento GTG e CMA, e dos seus pais para realização do CMA. Foram coletados 5 mL de sangue periférico em tubos a vácuo heparinizados e 5 mL de sangue periférico em tubos a vácuo com Ácido Etilonodiamino Tetra-acético (EDTA) (do inglês, *Ethylenediamine tetraacetic acid*), graduados e identificados. Posteriormente, as amostras heparinizadas foram utilizadas para realização do cariótipo e as amostras com EDTA foram processadas para a extração de DNA e realização da técnica de CMA e sequenciamento, segundo informações descritas abaixo.

5.5 Cariotipagem por bandeamento GTG

As amostras de sangue periférico heparinizadas dos pacientes foram submetidas à cultura de linfócitos T, a curto prazo, em meio RPMI 1640[®] (do inglês, *Roswell Park Memorial Institute*) enriquecido com fitohemaglutinina, L-glutamina e soro fetal bovino, conforme protocolo modificado para a obtenção de metáfases, adotado no NPR/PUC-Goiás/LaGene/LACEN/SES/GO (VERMA; BABU, 1995). Após 47h, foi adicionado ao meio de cultura colchicina, reagente utilizado para impedir a formação do fuso acromático, parando as divisões celulares na fase de metáfase, fase em que os cromossomos estão mais individualizados e condensados. Após esse processo, foi feita a hipotonização e fixação das células com solução hipotônica e fixador Carnoy, respectivamente. Posteriormente, a suspensão de células foi gotejada sobre vapor de banho-maria a 60°C, em lâminas limpas e desengorduradas, para o espalhamento das metáfases. Para a análise do cariótipo por bandeamento GTG (método Giemsa-Tripsina-Giemsa), as lâminas foram envelhecidas por sete dias e posteriormente tratadas em solução de tripsina (Gibco[®], Thermo Fisher Scientific, Massachusetts, EUA), diluídas em tampão PBS (Gibco[®], Thermo Fisher Scientific, Massachusetts, EUA) a 0,25% e coradas em solução Giemsa (Gibco[®], Thermo Fisher Scientific, Massachusetts, EUA) a 4%. Para a análise cromossômica, as metáfases foram capturadas com auxílio de microscopia óptica motorizada (Carl Zeiss[®], Alemanha) e analisadas com o auxílio *software* IKAROS[®] (Metasystems Corporation, EUA).

5.6 Extração de DNA

Para a extração de DNA genômico foi utilizado o Kit para Purificação de DNA *Illustra Blood GenomicPrep Mini Spin*[®] (GE Healthcare Life Sciences, EUA), de acordo com protocolo

sugerido pelo fabricante. A quantificação da concentração de DNA genômico foi realizada em um espectrofotômetro NanoVue™ Plus (GE Healthcare, Life Sciences, EUA).

5.7 Análise Cromossômica por Microarranjos (CMA)

O CMA, dos pacientes e de seus respectivos pais, foi conduzido utilizando o GeneChip® CytoScanHD™ array (Affymetrix, Santa Clara, EUA), de acordo com protocolo sugerido pelo fabricante sem modificação. A matriz de genotipagem do CMA oferece um alto poder de robustez e cobertura para investigação de variações estruturais no genoma. O GeneChip® CytoScanHD™ contem aproximadamente 2,6 milhões de marcadores, sendo ~ 1,9 milhões de marcadores não polimórficos e cerca de 750 mil marcadores polimórficos (SNPs), com capacidade para analisar CNVs, determinar LCSH e também baixos níveis de mosaïcismo.

Foi realizada digestão do DNA genômico amostral com a enzima de restrição (NspI) fornecida pelo fabricante. Este passo requer aproximadamente 250 ng de DNA total. Em seguida, o DNA digerido foi ligado aos adaptadores específicos e posteriormente amplificados por PCR anelados a *primers* universais. O amplicon foi fragmentado por digestão enzimática, marcado com biotina e hibridizado no chip por 18 horas. Após a hibridização, os chips foram lavados numa estação fluídica 450 e digitalizados no *GeneChip® Scanner 3.000 7G* (Affymetrix, Santa Clara, EUA). A captura e digitalização dos sinais obtidos pelo escaneamento foi feita pelo *software GeneChip® Command Console®* (AGCC®, Affymetrix, USA). Os arquivos CEL obtidos pelo escaneamento foram analisados usando o *software Chromosome Analysis Suite 3.0 (ChAS®)*. Os arrays que estavam dentro dos parâmetros de controle de qualidade determinados pelo fabricante, apresentando o MAPD ≤ 0.25 e SNP-QC ≥ 15 , foram analisados. Usando o ChAS®, foi possível investigar alterações estruturais ao longo do genoma, permitindo identificar CNVs por ganhos e perdas de segmentos de DNA, LCSH e conteúdo gênico presentes nas amostras estudadas de modo comparativo, facilitando as interpretações dos resultados.

A primeira análise de ganhos e perdas, presentes no genoma dos pacientes, foi feita utilizando o filtro de diagnóstico recomendado pelo fabricante, onde para ganho usou o filtro com 50 marcadores e para perda o filtro com 25 marcadores, ambos com o tamanho ≥ 100 kb. Todas as CNVs visualizadas foram classificadas de acordo com os consensos e diretrizes, apontados por Miller e colaboradores (2010), Kearney e colaboradores (2011), Battaglia e colaboradores (2013) e Silva e colaboradores (2019). Posteriormente, os pacientes que não apresentaram CNVs informativas no filtro de tamanho ≥ 100 kb, foram submetidos a um filtro

A amplificação de região de CNV de perda por PCR foi realizada utilizando a Taq DNA Polimerase Recombinante (Thermo Fisher Scientific, Massachusetts, EUA), juntamente com outros reagentes apresentados no Quadro 1, com um volume final de reação de 10 μ L, utilizando o termociclador Veriti™ Thermal Cycler (Applied Biosystems, Thermo Fisher Scientific, EUA). O ciclo utilizado está ilustrado na Figura 8, sendo que as temperaturas de anelamento foram ajustadas de acordo com cada par de *primers*.

Quadro 1. Concentração dos reagentes usados para o preparo do master mix para cada PCR dos genes selecionados para o estudo em pacientes com DI.

Reagentes	Concentração Inicial	Concentração final
<i>Primer F</i>	10 μ M	0,25 μ M
<i>Primer R</i>	10 μ M	0,25 μ M
Tampão	10x	1x
dNTP	2,5 mM	0,2 mM
MgCl ₂	50 mM	* mM
Taq	5 U/ μ L	0,5 U/reação

* A concentração de MgCl₂ variou para cada gene.

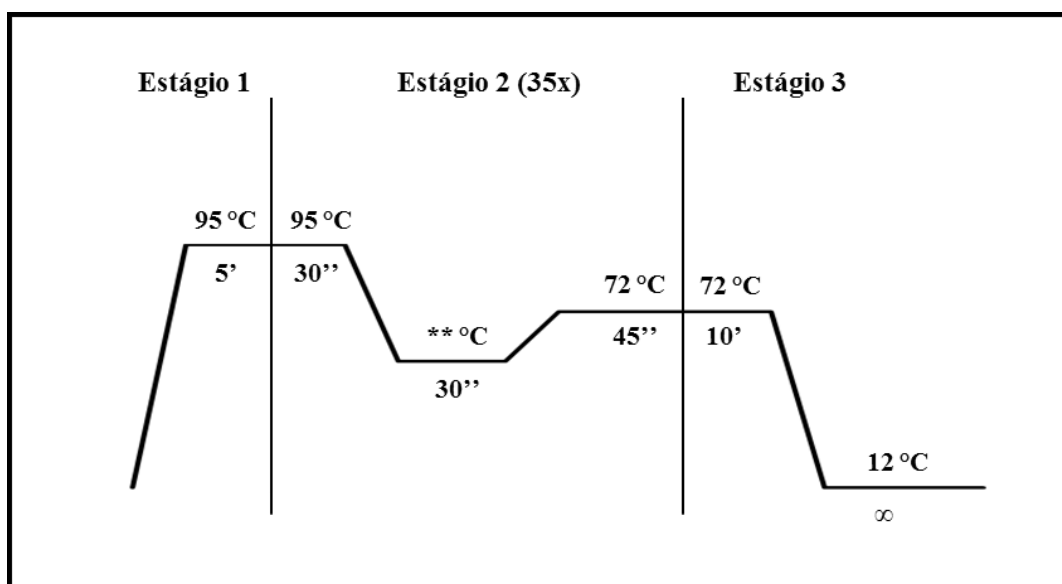


Figura 8. Representação gráfica do programa de amplificação (PCR).

** A temperatura de anelamento variou de acordo com os *primers*.

Após a amplificação dos produtos de PCR, foi feita a purificação destes produtos de PCR utilizando enzima ExoProStar 1 Step (GE Healthcare Life Sciences, EUA) de acordo com protocolo do fabricante modificado, para o qual foi usado 5,0 μ L do produto de PCR juntamente

com 2,0 μL da ExoStar e colocado no termociclador por 20 min à 37 °C para atividade da enzima e 15 min à 85 °C para inativação da enzima.

Após a purificação do material amplificado, os amplicons foram enviados para o Laboratório de Biotecnologia, Pós-Graduação em Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília (UCB-DF) para realizarem o serviço de sequenciamento de Sanger.

5.9 Sequenciamento de Sanger

O sequenciamento de Sanger das regiões de CNVs de perdas, que continham genes relacionados à DI, foi realizado em parceria com o Centro de Genômica de Alto Desempenho do Distrito Federal, Laboratório de Biotecnologia, Pós-Graduação em Ciências Genômicas e Biotecnologia da Universidade Católica de Brasília (UCB).

Após a reação da PCR e purificação do amplicon, as amostras foram enviadas para o Centro de Genômica de Alto Desempenho, onde a concentração do amplicon purificado, que é em torno de 10–20 ng, depende do tamanho do amplicon. As quantidades de amostra sugeridas pelo fabricante do kit de sequenciamento são demonstradas na Tabela 1.

Para a realização do sequenciamento, foram utilizados 1 μL dos *primers Forward* e *Reverse* concentração de 3,2 pmol, “X” μL de amostra (dependendo do tamanho do produto, vide tabela 1) e quantidade suficiente de H₂O Ultrapura para completar o volume final de 7,0 μL .

Tabela 1. Relação do tamanho do produto de PCR com o volume da amostra para sequenciar.

Tamanho do produto da PCR (pb)	Volume da amostra (ng)
100–200	1–3
200–500	3–10
500–1000	5–20
1000–2000	10–40
>2000	20–50

O sequenciamento dos fragmentos purificados foi feito pelo método de Sanger com eletroforese capilar, utilizando o sequenciador automático de DNA ABI 3130xl (Applied Biosystems, California, EUA).

Após a obtenção das sequências, as análises dos resultados foram feitas através do *software* ChromasPro – B11 – Versão 2.1.3, do *software* BioEdit Sequence Alignment Editor

e do *software* CLC Sequence Viewer 7. Além disso, as sequências analisadas foram comparadas com o banco de dados de sequências Nucleotide BLAST (*Nucleotide Basic Local Alignment Search Tool*) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) do NCBI (*National Center for Biotechnology Information*), com o objetivo de achar regiões de similaridade entre as sequências obtidas.

6. RESULTADOS

No período de janeiro de 2013 a dezembro de 2016 foram encaminhados 752 pacientes aos laboratórios NPR/PUC-Goiás/LaGene/LACEN/SES/GO, sendo 325/752 (43,2%) destes com indicação clínica de DI e o restante 427/752 (56,8%) com outras indicações.

Foi feita coleta de amostras de sangue periférico heparinizado dos pacientes para a realização do cariótipo por bandejamento GTG. Dos 325 pacientes com DI que fizeram o cariótipo, 186/325 (57,2%) obtiveram resultados diversos com o cariótipo, dos quais 78/186 (41,9%) foram diagnosticados com Síndrome de Down e 108/186 (58,1%) foram outros. O restante, 139/325 (42,8%) pacientes, apresentou o cariótipo sem alterações estruturais ou numéricas visíveis (Figura 9).

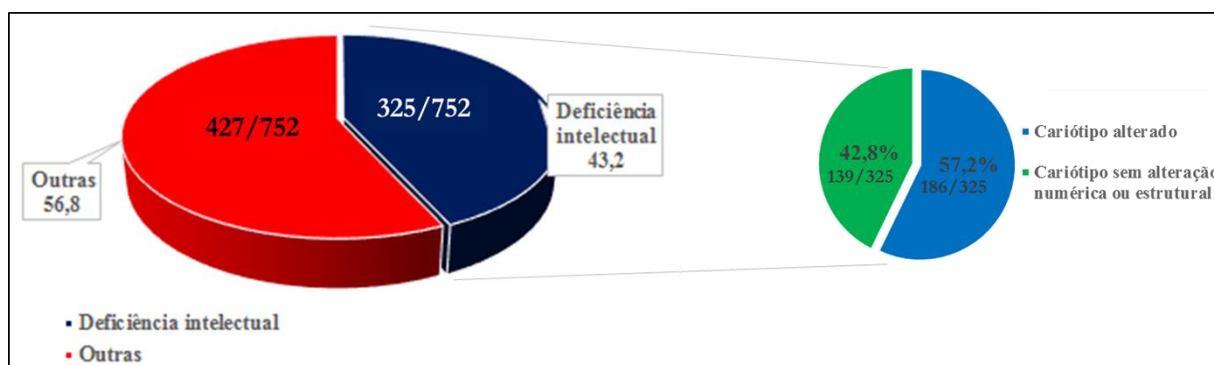


Figura 9. Indicação clínica e resultados do cariótipo dos pacientes com DI atendidos no período de Jan/2013 a Dez/2016 NPR/PUC-Goiás/LaGene/SES/GO.

Após esclarecimentos sobre o projeto, o grupo de 139 pacientes e seus respectivos pais biológicos foram convidados a participar do estudo e 77/139 (55,4%) trios aceitaram e assinaram os TCLEs. Os outros 62/139 (44,6%) pacientes não aceitaram participar do estudo por diversos motivos, sendo as três razões mais frequentes foram por se contentarem com o diagnóstico sem alteração do cariótipo, por não quererem fazer investigações adicionais ou por não terem comparecido à entrevista.

Dos 77 pacientes com diagnóstico de DI participantes do estudo, 40/77 (51,9%) eram do sexo masculino e 37/77 (48,1%) do sexo feminino. Todos os participantes e seus pais biológicos fizeram coleta de sangue com EDTA para a realização do CMA. Na análise dos resultados do CMA foi aplicado o filtro de diagnóstico recomendado pelo fabricante, cujos parâmetros foram 50 marcadores para ganho, 25 marcadores para perda e para ambos o tamanho

≥ 100kb. Após a realização do CMA, 14/77 (18,1%) pacientes apresentaram CNVs patogênicas, 9/77 casos (11,7%) apresentaram CNVs provavelmente patogênicas, 19/77 (24,7%) apresentaram CNVs de significado clínico incerto e 35/77 (45,5%) dos casos foram considerados negativos, pois o CMA não mostrou CNVs alteradas (Figura 10). A análise dos pais foi feita em função de seus respectivos filhos com DI para identificar a origem das CNVs observadas, se estas eram *de novo* ou herdadas.

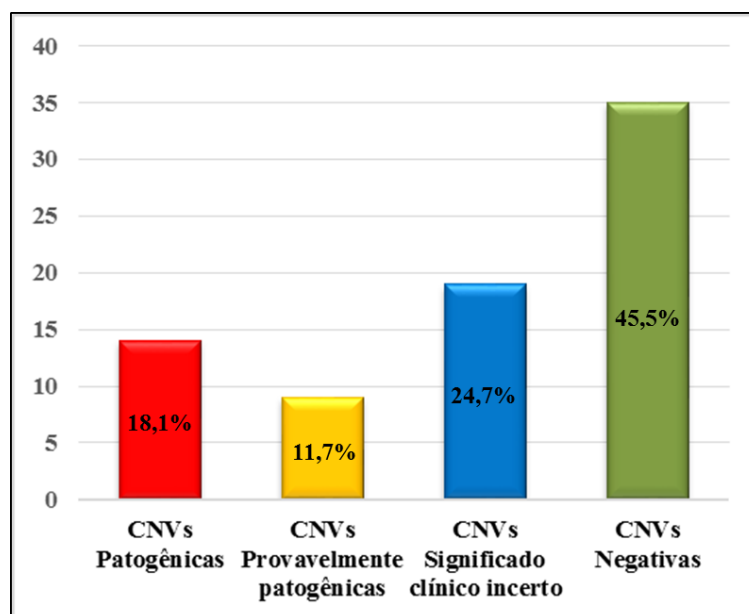


Figura 10. Tipos de CNVs, classificadas em função da sua patogenicidade, encontradas nos pacientes com DI que fizeram o CMA usando o filtro de diagnóstico recomendando pelo fabricante do GeneChip® CytoScanHD™ array.

Nos 35 casos negativos, foi feita uma análise adicional utilizando-se os parâmetros reduzidos do filtro de tamanho, mantendo-se os parâmetros de quantidade de marcadores no segmento genômico analisado (50 marcadores para ganho e 25 marcadores para perda, ambos com tamanho < 100 kb). Destes casos, 16/35 (45,7%) apresentaram pequenas CNVs presentes em genes potencialmente relacionados à DI e 19/35 casos (54,3%) mantiveram-se negativos. Nos 16 casos positivos, foram observadas 30 CNVs *de novo* (com tamanhos variando de 0,3–30 kb), sendo 15 CNVs por ganho e 15 CNVs por perda genômica (Tabela 2). Destas, foram selecionadas as CNVs de perda para a realização do sequenciamento de Sanger para a identificação dos pontos de quebra. No entanto, as CNVs de ganho e os genes nelas contidos não foram selecionadas para o sequenciamento de Sanger porque em segmentos de ganhos esta metodologia não é eficiente devido ao comprimento do segmento (Figura 11).

Dos 12/16 casos observados com CNVs de perda, 3 casos apresentaram 2 CNVs e os 9 casos restantes apresentaram apenas uma CNV. As 15 CNVs de perda observadas estavam presentes nos cromossomos 7 (1), 22 (1) e X (13), sendo os loci observados: 7q35, 22q13.33, Xp21.1, Xp21.2, Xp21.3, Xp22.2, Xq22.3, Xq23, e Xq27.1. Os 8 genes relacionados à DI presentes nas CNVs de perda foram: *CNTNAP2*, *SHANK3*, *DMD*, *ILIRAPL1*, *MID1*, *MID2*, *PAK3*, *FGF13*. Fazendo análise das 9 regiões cromossômicas que apresentaram CNVs de perda, 4 regiões foram observadas em mais de um paciente e as outras 5 regiões foram observadas em apenas 1 paciente.

Tabela 2. Resultado da análise com filtro de tamanho < 100 kb dos casos negativos do CMA.

Casos	Idade	Gênero	CNVs	Localização	Tamanho (kb)	Genes	Nomenclatura do microarranjo
001	9a	F	Ganho	Xp22.13	15	<i>NHS</i>	arr[GRCh37] Xp22.13(17.483.771-17.498.971)x3 dn
			Perda	Xq23	0,3	<i>PAK3</i>	arr[GRCh37] Xq23(110.356.320-110.356.593)x1 dn
			Perda	Xq27.1	20	<i>FGF13</i>	arr[GRCh37] Xq27.1(138.192.061-138.211.584)x1 dn
005	9a	F	Perda	Xp22.2	2	<i>MID1</i>	arr[GRCh37] Xp22.2(10.630.475-10.632.537)x0 dn
009	10a	F	Perda	Xq23	0,47	<i>PAK3</i>	arr[GRCh37] Xq23(110.356.320-110.356.790)x1 dn
010	9a	F	Ganho	Xp21.1	22	<i>DMD</i>	arr[GRCh37] Xp21.1(31.872.841-31.894.898)x3 dn
			Ganho	Xq22.1	6	<i>SRPX2</i>	arr[GRCh37] Xq22.1(99.900.084-99.906.042)x3 dn
			Ganho	Xq26.2	7	<i>GPC4</i>	arr[GRCh37] Xq26.2(132.434.060-132.440.801)x4 dn
			Ganho	Xq26.2	11	<i>GPC3</i>	arr[GRCh37] Xq26.2(133.015.257-133.026.071)x3 dn
			Ganho	Xq28	10	<i>MECP2</i>	arr[GRCh37] Xq28(153.325.445-153.335.837)x3 dn
016	5a	F	Ganho	Xq26.3	11	<i>ARHGEF6</i>	arr[GRCh37] Xq26.3(135.771.766-135.782.790)x3 dn
			Ganho	Xq27.3	6	<i>FMR1</i>	arr[GRCh37] Xq27.3(147.014.625-147.020.741)x3 dn
			Ganho	Xq28	3	<i>AFF2</i>	arr[GRCh37] Xq28(147.893.513-147.896.666)x3 dn
017	8a	F	Ganho	Xq12	30	<i>OPHN1</i>	arr[GRCh37] Xq12(67.551.934-67.582.422)x3 dn
			Ganho	Xq27.3	8	<i>FMR1</i>	arr[GRCh37] Xq27.3(147.012.658-147.020.538)x3 dn
031	1a	F	Perda	7q35	21	<i>CNTNAP2</i>	arr[GRCh37] 7q35(146.552.052-146.573.427)x1 dn
			Perda	Xp21.2	3	<i>ILIRAPL1</i>	arr[GRCh37] Xp21.2(29.316.508-29.319.883)x1 dn
			Ganho	Xq28	10	<i>AFF2</i>	arr[GRCh37] Xq28(147.765.026-147.775.292)x3 dn
032	1a	F	Perda	22q13.33	9	<i>SHANK3</i>	arr[GRCh37] 22q13.33(51.106.757-51.115.526)x1 dn
035	16a	F	Perda	Xp21.1	1	<i>DMD</i>	arr[GRCh37] Xp21.1(32.742.108-32.743.344)x1 dn
037	1m	F	Perda	Xp21.1	11	<i>DMD</i>	arr[GRCh37] Xp21.1(32.919.479-32.930.885)x1 dn
057	7a	F	Perda	Xp21.2	24	<i>ILIRAPL1</i>	arr[GRCh37] Xp21.2(29.318.639-29.342.203)x1 dn
			Ganho	Xq26.2	7	<i>GPC4</i>	arr[GRCh37] Xq26.2(132.434.060-132.440.801)x3 dn
065	9m	M	Perda	Xp21.3	14	<i>ILIRAPL1</i>	arr[GRCh37] Xp21.3(29.160.655-29.174.711)x0 dn
066	26a	F	Perda	Xq22.3	2	<i>MID2</i>	arr[GRCh37] Xq22.3(107.110.275-107.112.208)x1 dn
			Perda	Xq23	1	<i>PAK3</i>	arr[GRCh37] Xq23(110.356.320-110.357.527)x1 dn
067	2a	M	Perda	Xp21.1	18	<i>DMD</i>	arr[GRCh37] Xp21.1(33.099.140-33.117.042)x0 dn
073	6a	M	Ganho	Xq13.3	24	<i>ZDHC15</i>	arr[GRCh37] Xq13.3(74.697.085-74.720.997)x2 dn
			Ganho	Xq26.2	12	<i>GPC3</i>	arr[GRCh37] Xq26.2(132.683.434-132.695.673)x2 dn
077	12a	F	Perda	Xp21.3	14	<i>ILIRAPL1</i>	arr[GRCh37] Xp21.3(29.160.655-29.174.711)x0 dn

A sequência dos *primers* utilizados para mapeamento ponto de quebra das CNVs de perda selecionadas que abrigavam genes relacionados com a DI e o intervalo de tamanho mínimo e máximo esperado de cada amplicon estão representados na Tabela 3.

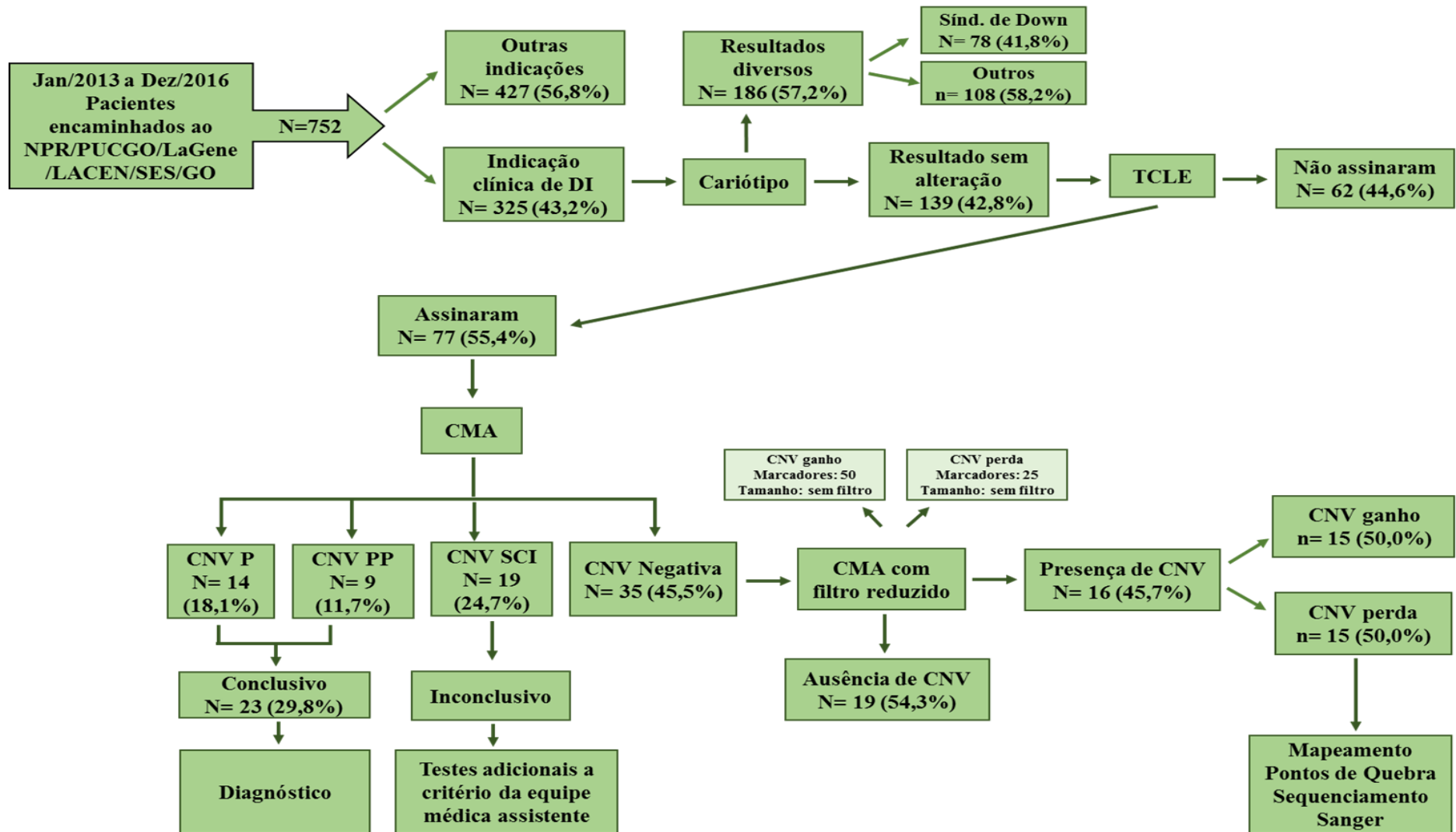


Figura 11. Fluxograma dos resultados obtidos durante a realização do estudo.

CNV P=CNV Patogênica; CNV PP= CNV Provavelmente Patogênica; CNV SCI= CNV Significado Clínico Incerto; N=Pacientes; n=Número de CNVs encontradas.

Tabela 3. CNVs de perda usadas nos experimentos de identificação dos prontos de quebra por Sequenciamento de Sanger, seus respectivos genes, a localização da última sonda *upstream* e primeira sonda *downstream*, sequência dos *primers forward* (F) e *reverse* (R) utilizados para cada região e o intervalo de tamanho esperado do amplicon.

Casos	Genes	Nomenclatura do microarranjo (GRCh37)	Última sonda <i>upstream</i>	Primeira sonda <i>downstream</i>	Conversão para GRCh38	Posição para construção <i>primer</i>	<i>Primer</i>	Intervalo de tamanho do amplicon (pb)
001	<i>PAK3</i>	chrX:110.356.320-110.356.593	110.356.220	110.356.723	chrX:111.112.992-111.113.495	chrX:111.111.992-111.114.495	F - TCC AGG CAC CAT TAG CTA CC R - GCA CTT CAG AAA AGG GCA AG	1.611-1.842
	<i>FGF13</i>	chrX:138.192.061-138.211.584	138.191.934	138.211.820	chrX:139.109.772-139.129.658	chrX:139.108.772-139.130.658	F - CAG ACA CGA TTT GGC ATT TG R - GAT CTG CCT CCT AGC CCT CT	314-677
005	<i>MID1</i>	chrX:10.630.475-10.632.537	10.630.449	10.632.651	chrX:10.662.409-10.664.611	chrX:10.661.409-10.665.611	F - CAA AAT TAG CTG GGC ATG GT R - GGA AGA TCA AAT GCC AAG GA	733-873
009	<i>PAK3</i>	chrX:110.356.320-110.356.790	110.356.220	110.356.723	chrX:111.112.992-111.113.495	chrX:111.111.992-111.114.495	F - TCC AGG CAC CAT TAG CTA CC R - GCA CTT CAG AAA AGG GCA AG	808-1.645
031	<i>CNTNAP2</i>	chr7:146.552.052-146.573.427	146.550.842	146.573.485	chr7:146.853.750-146.876.393	chr7:146.852.750-146.877.393	F - CAG GAA TGC CAA TGA TGT TG R - TGC ATG GGT ATG ACA GAG GA	965-2233
	<i>ILIRAPL1</i>	chrX:29.316.508-29.319.883	29.316.464	29.319.967	chrX:29.298.347-29.301.850	chrX:29.297.347-29.302.850	F - TGC CTT AAG ATC AGG CAA CA R - GCC ATC CTG GAT CAC TGT CT	320-448
032	<i>SHANK3</i>	chr22:51.106.757-51.115.526	51.104.143	51.120.080	chr22:50.665.715-50.681.652	chr22:50.664.715-50.683.652	F - GAT CTG CTC ACC TTG GCT TC R - GTG CAC GCA TGC AAT CAC T	594-7762
035	<i>DMD</i>	chrX:32.742.108-32.743.344	32.741.911	32.743.390	chrX:32.723.794-32.725.273	chrX: 32.722.794-32.726.273	F - CCC CTT AAA TCA GGG GTA TG R - TTG CTG AAT TGA CCC CTT TC	333-576
037	<i>DMD</i>	chrX:32.919.479-32.930.885	32.919.163	32.931.540	chrX:32.901.046-32.913.423	chrX:32.900.046-32.914.423	F - CTC AGC CTC CCA AAA CTC TG R - TTA CAG AGA AGC AGT GGG TGC T	744-1.825
057	<i>ILIRAPL1</i>	chrX:29.318.639-29.342.203	29.318.252	29.347.243	chrX:29.300.135-29.329.126	chrX:29.299.135-29.330.126	F - GAG TAA TGC GGG TGT GTC CT R - GTG GAA TTG CTG GGT CAA AT	476-5.901
065	<i>ILIRAPL1</i>	chrX:29.160.655-29.174.711	29.160.572	29.174.927	chrX:29.142.455-29.156.810	chrX:29.141.455-29.157.810	F - TGC TAC AAA ATG GCT GAC CA R - CAA AAC CCT TCG ACC TGT GT	197-486
066	<i>MID2</i>	chrX:107.110.275-107.112.208	107.110.203	107.116.180	chrX:107.866.973-107.872.950	chrX:107.865.973-107.873.950	F - CCC CAT TTT TCA GAT GAG GA R - GAG GCT ATC CAG GGG AAG AC	403-4447
	<i>PAK3</i>	chrX:110.356.320-110.357.527	110.356.220	110.357.670	chrX:111.112.992-111.114.442	chrX:111.111.992-111.115.442	F - TCC AGG CAC CAT TAG CTA CC R - CAG TTG GTT TTG TGG CCT TT	664-907
067	<i>DMD</i>	chrX:33.099.140-33.117.042	33.091.271	33.117.384	chrX:33.073.154-33.099.267	chrX:33.072.154-33.100.267	F - CTG GGC AGC ATA ATT TCC AT R - GCC TTT AAT GGG CAA GAT GA	798-9.010
077	<i>ILIRAPL1</i>	chrX:29.160.655-29.174.711	29.160.572	29.174.927	chrX:29.142.455-29.156.810	chrX:29.141.455-29.157.810	F - TGC TAC AAA ATG GCT GAC CA R - CAA AAC CCT TCG ACC TGT GT	197-486

Após a amplificação dos segmentos genômicos que continham as CNVs de perda, os produtos amplificados foram purificados e em seguida encaminhados para a realização do sequenciamento de Sanger. Dentre as 15 regiões de CNVs de perda encaminhadas para o sequenciamento, nenhuma sequência obtida alinhou com a região esperada (APÊNDICE 1).

Para tentar solucionar este problema e identificar a região de quebra, novos *primers* foram redesenhados para anelar em uma região mais conservada do gene, aumentando a especificidade dos primers durante a amplificação por PCR. Para isso, foram escolhidas 3 regiões de CNVs de perda que continham os genes *CNTNAP2*, *FGF13* e *ILIRAPLI* para serem amplificadas por esses novos *primers* (Tabela 4).

Tabela 4. Sequência de oligonucleotídeos (*primers*) redesenhados para cada região dos genes selecionados e o intervalo de tamanho mínimo e máximo esperado de cada amplicon.

Gene	Primer	Intervalo do amplicon (pb)
<i>CNTNAP2</i>	F - TGA GAG GTA AGA GAC CAG GCC T R - ATG CAT GGG TAT GAC AGA GGA A	619 -1887
<i>ILIRAPLI</i>	F - TCT GAC AGA GTA TGG ATG AAT AGC A R - TGT TTG GAT ATT TTA CAG CAC TTC A	935 - 1234
<i>FGF13</i>	F - AAA ACT AGG GCT TTC TCG GC R - TGG GCT TGA GAC GGA CAA G	533 - 899

Após a PCR com os novos *primers*, os produtos amplificados foram purificados e em seguida encaminhados para a realização do sequenciamento de Sanger. Dos três novos *primers* desenhados, apenas o da região que continha o gene *CNTNAP2* apresentou um amplicon dentro do intervalo de tamanho esperado e foi encaminhado para o sequenciamento. Para as outras duas regiões, que continham os genes *FGF13* e *ILIRAPLI*, não foram obtidos amplicons dentro do intervalo de tamanho esperado e, portanto, os produtos espúrios não foram sequenciados.

A nova sequência obtida da amplificação do gene *CNTNAP2* não alinhou novamente com a região esperada (APÊNDICE 2). Diante deste resultado, tomou-se a decisão de clonar o gene *CNTNAP2* usando a clonagem TA, para posteriormente amplificar os clones utilizando-se os *primers* específicos para região do gene e *primers* específicos para região do plasmídeo. Os produtos serão sequenciados, porém, esta etapa não foi concluída e os resultados não farão parte do produto final desta tese.

7. DISCUSSÃO

A deficiência intelectual é uma desordem neurológica, caracterizada por uma diminuição significativa nas funções cognitivas e adaptativas, comprometendo várias habilidades sociais e práticas diárias. Por ser um grande problema sócio-econômico e de saúde pública, o diagnóstico específico e preciso é fundamental para o gerenciamento eficiente de comorbidades associadas e o aconselhamento genético adequado para a família, com o pressuposto de melhorar a qualidade de vida dos pacientes acometidos pela DI e suas famílias.

Embora os aspectos sócio-econômicos, doenças infecciosas e agentes tóxicos contribuam para a prevalência da DI, os fatores genéticos são considerados diversos e responsáveis pela maioria dos casos. Em populações não relacionadas, as aberrações cromossômicas e mutações de ponto são as causas mais frequentes de DI. Entretanto, em famílias consanguíneas, a herança autossômica recessiva de genes mutados importantes para o desenvolvimento cerebral aparece como a causa mais comum da DI (REUTER *et al.*, 2017).

Testes genéticos laboratoriais, incluindo o cariótipo por bandeamento GTG, CMA, sequenciamento de um único gene e/ou painel de genes, e sequenciamento de exoma estão entre as metodologias rotineiramente usadas para tentar elucidar as causas genéticas da deficiência intelectual em todo o mundo (BOWLING *et al.*, 2017).

A análise cromossômica por microarranjo, com maior resolução, alta densidade de marcadores e robustez de análise tem sido eficiente na identificação de perdas e ganhos genômicos associados à DI apresentando um incremento no diagnóstico de até 25% (RODRÍGUEZ-REVENGA *et al.*, 2013; PEREIRA *et al.*, 2014; NOWAKOWSKA, 2017). Contudo, ainda há uma grande quantidade de pacientes com DI que permanecem sem o diagnóstico laboratorial, mostrando que a aplicação do filtro de tamanho < 100 kb no CMA e o mapeamento dos pontos de quebra, pelo sequenciamento, destas regiões de CNV menores que a capacidade de resolução de diagnóstico do CMA, se torna uma ferramenta interessante para auxiliar na identificação de CNVs presentes em genes responsáveis pelo desenvolvimento do sinal fenotípico da DI e que podem estar interferindo na expressão destes genes.

O presente estudo teve como objetivo geral avaliar a ocorrência de CNVs, identificadas pelo CMA com filtro de tamanho < 100 kb, que abrigassem genes funcionalmente associados à deficiência intelectual, aumentando assim a taxa de diagnóstico etiológico da DI. As amostras foram constituídas por pacientes do SUS provenientes de Goiânia (Goiás, Brasil),

com diagnóstico clínico de DI, encaminhados para o diagnóstico genético laboratorial. O estudo foi desenvolvido em pacientes com deficiência intelectual, encaminhados aos laboratórios NPR/PUCGO e LaGene/LACEN/SES/GO para o diagnóstico genético laboratorial, que apresentaram o diagnóstico laboratorial sem alteração após a utilização dos testes de cariótipo por bandeamento GTG e de CMA.

O teste de cariótipo por bandeamento GTG, realizado com os 325 pacientes que apresentavam DI, conseguiu identificar, em 57,2% dos casos, alterações numéricas e estruturais diversas, dos quais 24,0% foram referentes à Síndrome de Down, incluindo trissomia livre do cromossomo 21, trissomia parcial do braço longo do cromossomo 21, translocação Robertsoniana (14;21) e trissomia 21 em mosaico. Os nossos resultados estão em consonância com estudos de Bertelli e colaboradores (2016) e Kruszka e colaboradores (2017), que demonstraram que uma grande parcela dos casos de DI observados na população se deve a Síndrome de Down, sendo conhecidamente a principal causa deste sinal fenotípico.

Novas tecnologias têm aumentado a capacidade de detecção de pequenos rearranjos em todo o genoma, nos ajudando a entender melhor a função destas variações genômicas nas doenças, principalmente na DI (CAPALBO *et al.*, 2017). O CMA, realizado nos 77 pacientes com DI que aceitaram participar do estudo e que apresentavam o resultado do cariótipo sem alteração numérica ou estrutural, possibilitou uma taxa de diagnóstico de 29,8%, estando esse achado de acordo com os achados da literatura, que mostraram um aumento no diagnóstico genético, utilizando técnica de CMA, em média de 20% (BATTAGLIA *et al.*, 2013; BERTOK *et al.*, 2015; CAPALBO *et al.*, 2017).

A realização, nos nossos pacientes encaminhados com DI, do CMA com filtro de diagnóstico após o cariótipo por bandeamento GTG, proporcionou um incremento na taxa de diagnóstico de 7,1%, demonstrando que a utilização dessas duas metodologias permitiu, em 64,5% dos pacientes, a identificação das causas genéticas da DI. Aproximadamente 40% dos casos requerem outras metodologias, no entanto, quando a causa a ser investigada é DI, TEA e ACM, o consenso recomenda a utilização do CMA (MILLER *et al.*, 2010; KEARNEY *et al.*, 2011; BATTAGLIA *et al.*, 2013).

A aplicação da análise cromossômica por microarranjo neste estudo proporcionou um aumento na taxa de diagnóstico, mas ainda assim, 45,5% dos 77 pacientes encaminhados permaneceram sem a identificação de CNVs alteradas. Com o uso do filtro de tamanho reduzido, para tentar identificar CNVs menores e que pudessem estar presentes em regiões que apresentavam genes relacionados à DI, foi possível a identificação de CNVs *de novo* de perda

e de ganho, com tamanhos variando de 0,3–30 kb, em 45,7% destes casos negativos. Com a redução do filtro de tamanho, a taxa de incremento na identificação de CNVs de perda presentes em genes relacionados à DI foi de 12%, demonstrando que a redução do filtro de diagnóstico pode ser uma alternativa para a elucidação do diagnóstico genético para os pacientes com DI. Principalmente devido as CNVs de perda ocasionarem interrupção do gene ou da junção dos pontos de quebra, interferindo na sensibilidade à dosagem deste gene.

No presente estudo, foram identificados genes reconhecidamente envolvidos no processo de desenvolvimento do sistema nervoso central, implicados na função sináptica e plasticidade cerebral. De acordo com Verpelli e colaboradores (2014), mais da metade dos genes relacionado à DI localizados no cromossomo X codificam proteínas presentes nos compartimentos pré- ou pós-sinápticos e estão envolvidos na formação da sinapse e plasticidade cerebral.

Dos 8 genes detectados, detalhadamente descritos abaixo, os genes *CNTNAP2*, *FGF13*, *MID1* e *MID2*, estão envolvidos no processo de desenvolvimento do sistema nervoso central, incluindo morfogênese, crescimento, diferenciação e adesão celular, progressão do ciclo celular, além da regulação das estruturas do citoesqueleto (ZHANG *et al.*, 2012; POOT, 2015; PURANAM *et al.*, 2015; GHOLKAR *et al.*, 2016; SMOGAVEC *et al.*, 2016; POOT, 2017). Já os genes *SHANK3*, *IL1RAPL1*, *DMD* e *PAK3* estão envolvidos na sinaptogênese, função sináptica, plasticidade, bem como na regulação da dinâmica do citoesqueleto (GAO; MCNALLY, 2015; LEE *et al.*, 2017; MONTANI *et al.*, 2017; RENKILARAJ *et al.*, 2017).

a) Gene *CNTNAP2*

O gene *CNTNAP2*, denominado gene da Proteína Associada à Conectina Tipo 2 (OMIM604569), localizado em 7q35, é um dos maiores genes do genoma humano com tamanho aproximado de 2,3 Mb, apresentando uma moldura aberta de leitura (ORF) (do inglês, *Open Reading Frame*) de tamanho total de 3.996 nucleotídeos correspondendo a um transcrito de 24 exons. Este gene codifica uma proteína (CASPR2) pertencente à família das neurexinas, que no sistema nervoso funcionam como moléculas e receptores de adesão celular, regulando o contato entre neurônios e células da glia (SMOGAVEC *et al.*, 2016; POOT, 2017).

O transcrito completo de *CNTNAP2* codifica uma grande região extracelular da proteína CASPR2. A proteína CASPR2 é uma proteína transmembrana composta por vários domínios: 8 extracelulares, 1 transmembrana e 2 C-terminal intracelular. Os domínios extracelulares são compostos por 4 domínios laminina G, 2 domínios associados ao fator de

crescimento epidérmico (FCE), 1 domínio associado à discoidina e 1 domínio associado ao fibrinogênio. Estes domínios extracelulares mediam interações célula-célula e ligações com receptores, ligantes e com a matriz extracelular. A pequena parte C-terminal citoplasmática contém um sítio de ligação domínio PDZ, que supostamente media interações proteínas-proteínas que contém PDZ (Figura 12) (POOT, 2015; SMOGAVEC *et al.*, 2016).

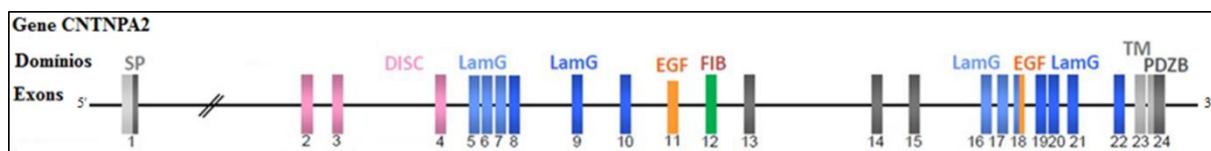


Figura 12. Ilustração da distribuição dos exons e dos domínios regulatórios na estrutura do gene *CNTNAP2*. Cada cor diferente mostra os domínios codificantes/exons. SP=peptídeo sinal; DISC=domínio associado à discoidina; LamG= domínio laminina-G; EGF=domínio associado ao fator de crescimento epidérmico; FIB=domínio associado ao fibrinogênio; TM=região transmembrana; PDZB=sítio ligação domínio PDZ.

Fonte: Adaptado de SMOGAVEC *et al.*, 2016.

No cérebro do feto humano, o gene *CNTNAP2* é predominantemente expresso em regiões anteriores perisilvianas frontal e temporal, no estriado e no tálamo. Da 18ª a 20ª semana de gestação, o cérebro fetal mostra um enriquecimento dos sinais de *CNTNAP2* na substância cinzenta frontal, aproximadamente entre o giro orbital e superior frontal cortical, na área de broca, parte do cérebro humano responsável pela expressão da linguagem, e outras regiões perisilvianas do cérebro. Portanto, o *CNTNAP2* é altamente expresso na medula espinhal e várias regiões do cérebro, particularmente no circuito córtico-estriado-talâmico, envolvido em diversos processos cognitivos de ordem superior, incluindo linguagem, recompensa e função executiva frontal. Além disso, a *CASPR2* possui uma função organizadora no desenvolvimento de neurônios, sendo essencial para a montagem de circuitos neurais no sistema nervoso central (RODENAS-CUADRADO *et al.*, 2016; SMOGAVEC *et al.*, 2016; POOT, 2017).

Devido ao *CNTNAP2* ser um dos maiores genes do genoma humano, ele é alvo de rearranjos estruturais, incluindo inversões e translocações, CNVs, mutações nos exons e sítios de ligações dos fatores de transcrição, e por modificações epigenéticas. Por perturbação em mecanismos complexos de penetrância, expressividade e efeitos nas interações com outros alelos e loci, *CNTNAP2* pode estar relacionado ao autismo, DI, epilepsia e problemas na linguagem (ZWEIER, 2011; RODENAS-CUADRADO *et al.*, 2014; POOT, 2015).

Adicionalmente, mutações com perda de função do gene *CNTNAP2*, rearranjos genômicos heterozigóticos e CNVs na região de 7q35, englobando o gene *CNTNAP2*, têm sido

implicados com uma variedade de desordens neurocomportamentais, como a Síndrome de Pitt-Hopkins (OMIM610042), Susceptibilidade ao Autismo (OMIM612100), Comprometimento Específico da Linguagem (OMIM612514), com sinal fenotípico da DI, epilepsia e esquizofrenia (ZWEIER, 2011; RODENAS-CUADRADO *et al.*, 2016; SMOGAVEC *et al.*, 2016; LISKA *et al.*, 2017; POOT, 2017).

b) Gene *FGF13*

O gene *FGF13*, conhecido como o gene do Fator de Crescimento de Fibroblastos tipo 13 (OMIM300070), localizado em Xq27.1, codifica uma proteína que pertence à família de proteínas de sinalização, fatores de crescimento de fibroblastos (FGF), que são responsáveis pelo crescimento celular, diferenciação, morfogênese e processo de desenvolvimento (ZHANG *et al.*, 2012).

Nos mamíferos, os FGFs são divididos em 7 famílias, entre elas FGF4, FGF5, FGF8, FGF9, FGF10, FGF11 e FGF15/19. O gene *FGF13* faz parte da família FGF11, também conhecida como FGFs intracelulares (iFGFs), que não possuem receptor peptídeos sinal, atuam dentro da célula e expressam predominantemente no sistema nervoso. Todas os membros das famílias FGFs possuem uma região homóloga conhecida como domínio core. Em particular, os genes da família *FGF11* (*FGF11*, *FGF12*, *FGF13* e *FGF14*) têm duas isoformas produzidas por diferentes iniciadores de transcrição, sendo que a isoforma A dos membros desta família contém um sinal de localização nuclear N-terminal (NLSs) (Figura 13). Cada um desses quatro genes possui dois ou mais sítios de iniciação da transcrição, gerando múltiplos mRNAs com primeiros exons distintos unidos com os exons comuns II-V, que produzem proteínas com diferentes sequências amino-terminais (GOLDFARB, 2005; ZHANG *et al.*, 2012).

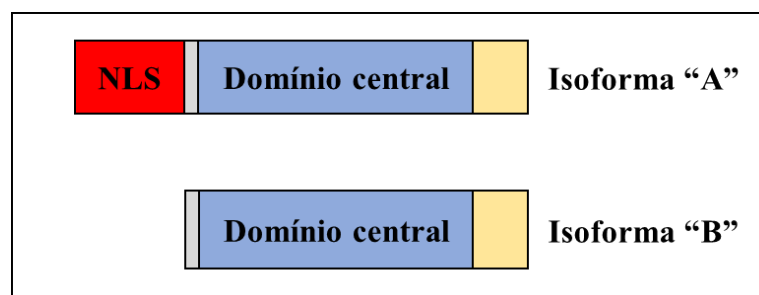


Figura 13. Representação esquemática das isoformas da família FGF11 não secretória intracelular. Fazem parte desta família os genes *FGF11*, *FGF12*, *FGF13* e *FGF14*.
Fonte: Adaptado de ZHANG *et al.*, 2012.

As proteínas FGF intracelulares, pertencentes à família FGF11, não são segregadas e não ativam os receptores de FGF, e podem interagir com os domínios intracelulares de canais de Na⁺ de tensão controlada, que são necessários para o funcionamento motor normal. Além disso, durante o desenvolvimento, membros da família *FGF11* são expressos em neurônios cortical cerebral e o *FGF13* é amplamente distribuído com alta expressão no cérebro durante o desenvolvimento e tem função essencial no estabelecimento dos circuitos neurais no córtex cerebral e possibilitam funções cognitivas (GUILLEMOT; ZIMMER, 2011; WU *et al.*, 2012; PURANAM *et al.*, 2015).

A proteína FGF13 age dentro da célula como uma proteína estabilizadora de microtúbulos necessária para o axônio, levando ao processo de desenvolvimento e migração neuronal no córtex cerebral. FGF13 está enriquecida nos cones de crescimento axonal e interage diretamente com os microtúbulos. Além disso, FGF13 polimeriza tubulinas e estabiliza os microtúbulos (WU *et al.*, 2012; PURANAM *et al.*, 2015; YANG *et al.*, 2016).

A perda de FGF13 prejudica a polarização neuronal e aumenta a ramificação dos axônios e o processo condutor. A função cerebral e a plasticidade dependem da arquitetura complexa dos circuitos neurais que são estabelecidos durante o desenvolvimento. Defeitos na morfogênese neuronal, migração celular e sinaptogênese do córtex cerebral podem levar à DI, transtornos do espectro autista e epilepsia. Mutações no gene *FGF13* podem causar DI ligada ao X síndrome ou não síndrome, sendo que o gene alterado pode afetar no aprendizado e na memória, além de causar DI de leve a moderada (WU *et al.*, 2012; ZHANG *et al.*, 2012; PURANAM *et al.*, 2015).

c) Gene *MIDI*

O gene *MIDI* (OMIM300552), localizado na região Xp22.2, tem aproximadamente 400 kb de sequência genômica e possui 9 exons (SCHWEIGER; SCHNEIDER, 2003; JI *et al.*, 2014). Foi identificado pela primeira vez por Quaderi e colaboradores em 1997, que descobriram que mutações neste gene são responsáveis por uma doença monogênica rara chamada de Síndrome Opitz BBB/G (OMIM300000) (QUADERI *et al.*, 1997; WINTER *et al.*, 2016).

MIDI codifica uma ubiquitina ligase E3, proteína associada à microtúbulos que pertence à subfamília de proteínas motivo tripartite (TRIM) de proteínas de “dedo *RING*” (do inglês, *RING-finger proteins*). A região N-terminal desta proteína é caracterizada pelo domínio RING, dois domínios B-box (B-box1 e B-box2) e um domínio de “bobina enrolada” (do inglês,

coiled coil domain) (HÜNING *et al.*, 2013; MIGLIORE *et al.*, 2013). Há 9 subfamílias de proteínas TRIM, cada uma se diferencia da outra pela variação no seu domínio C-terminal. A proteína MID1 pertence à subfamília C–I, consiste do subgrupo C-terminal constituído pelos domínios fibronectina III e B30.2. Esta proteína forma homodímeros que se associam a microtúbulos no citoplasma. A função da região C-terminal de MID1 não está completamente entendida, mas sabe-se que contribui para à interação proteína-microtúbulos e proteína-proteína (Figura 14) (JI *et al.*, 2014; WRIGHT *et al.*, 2017).

A proteína MID1 é encontrada apenas no citoplasma e tem atividade ubiquitina ligase que regula a degradação catalítica da subunidade da proteína fosfatase 2A (PP2A) através da interação com a subunidade 4 α da PP2A. A MID1 liga-se à subunidade 4 α através do domínio B-box1, sendo que a associação com os microtúbulos é conseguida através da domínio C-terminal B30.2, no qual está ativo durante o ciclo celular (MNAYER *et al.*, 2006; GRIESCHE *et al.*, 2016; NAKAMURA *et al.*, 2017). A interação MID1- α 4 é um ponto focal importante para o controle da quantidade imensa de processos biológicos essenciais à níveis celular e organizacional (FONTANELLA *et al.*, 2008; NICHOLSON, 2017).

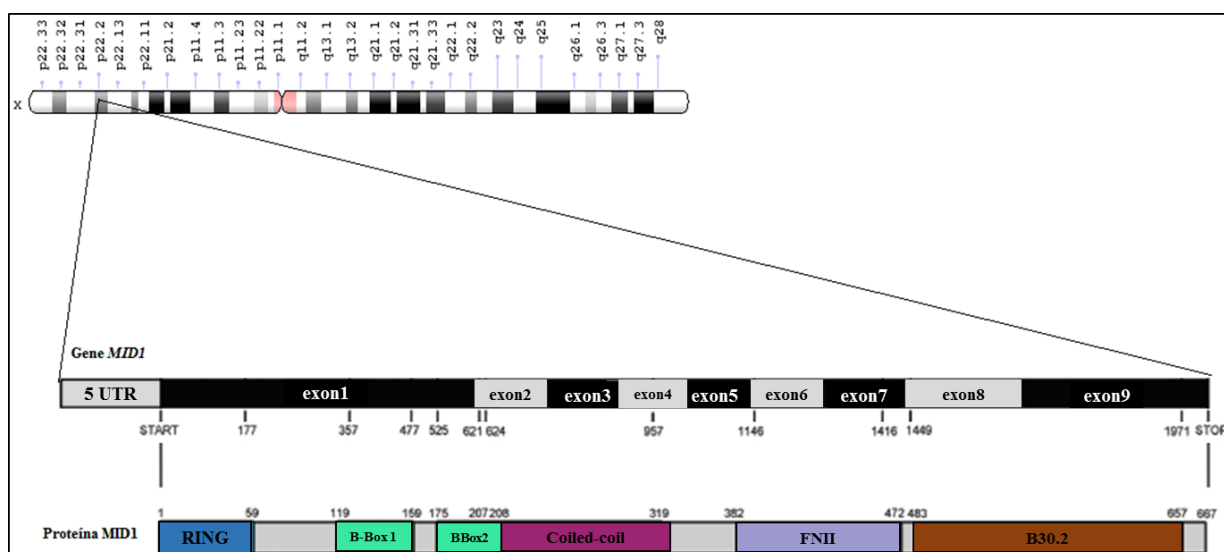


Figura 14. Representação esquemática do gene *MID1*, sua localização e proteína MID1.
Fonte: Adaptado de SCHWEIGER; SCHNEIDER, 2003 e WINTER *et al.*, 2016.

Mutações germinativas no gene *MID1* causam a síndrome Optiz BBB/G, uma desordem monogênica rara, que envolve malformações da linha média ventral incluindo hipertelorismo, hipospádia, fenda labial e palatina, defeitos cardíacos, ânus imperfurado, atraso no desenvolvimento e DI. O espectro de mutações neste gene engloba mutações *missense*,

mutações sem sentido, pequenas inserções e deleções, alterações sítio de splice, bem como deleções de exons (HÜNING *et al.*, 2013; MIGLIORE *et al.*, 2013; UNTERBRUNER *et al.*, 2018). Além disso, o gene *MID1* está associado com desenvolvimento e progressão de várias outras doenças incluindo o câncer e doenças neurodegenerativas (WINTER *et al.*, 2016; UNTERBRUNER *et al.*, 2018).

d) Gene *MID2*

O gene *MID2* (OMIM300204), codifica uma proteína ubiquitina ligase com 83% de similaridade com a proteína *MID1*. Semelhante ao gene *MID1*, o gene *MID2* está localizado no cromossomo X, mas na região Xq22.3. Ambos os genes codificam proteínas conservadas associadas com microtúbulos pertencentes à superfamília de RBCC/TRIM (motivo tripartite de proteína de “dedo *RING*”, com domínio B-box e *coiled coil*), no quais são caracterizadas pela habilidade de agir como suporte com múltiplas interações proteína-proteína. *MID1* e *MID2* podem justapor funções: ambas estão distribuídas entre os microtúbulos de todos os tipos celulares, ambas interagem com subunidade 4 α da PP2A no qual regula o volume nos microtúbulos e ambas têm atividade ubiquitina ligase E3, com papéis variáveis na homeostase celular, incluindo sinalização, regulação das estruturas do citoesqueleto e progressão do ciclo celular (SUZUKI *et al.*, 2010; GHOLKAR *et al.*, 2016; LI *et al.*, 2016).

Quantidade de dose dos genes *MID1/MID2* pode ser crítica na regulação durante o desenvolvimento inicial, sendo que a expressão de ambos os genes foi observada em estruturas craniofaciais embriônicas. Além disso, a expressão de *MID2* é particularmente forte dentro da porção anterior do cérebro posterior que dará origem ao cerebelo em um estágio posterior. Mutações no gene *MID2* tem sido relacionada à DI ligada ao X (OMIM300928), mas também demonstra uma ampla variedade nas características fenotípica (SHORT *et al.*, 2002; SUZUKI *et al.*, 2010; GEETHA *et al.*, 2014; GHOLKAR *et al.*, 2016; LI *et al.*, 2016).

e) Gene *SHANK3*

O gene *SHANK3*, conhecido como Múltiplos Domínios Repetidos de Anquirina e SH3 (OMIM606230), localizado em 22q13.3, é um dos três membros da família *SHANK* que codificam proteínas de suporte que permitem a interação com uma ampla variedade de proteínas diferentes, conectam receptores de neurotransmissores, canais iônicos e outras proteínas de membrana ao citoesqueleto de actina e as vias de sinalização acopladas à proteína G (COSTALES; KOLEVZON, 2015; GÖRKER *et al.*, 2017; LEE *et al.*, 2017).

As proteínas da família SHANK são uma das proteínas mais abundantes nas sinapses excitatórias neuronais, desempenhando um papel importante na formação de sinapses e na maturação da coluna dendrítica. A família dessas proteínas é altamente enriquecida em um compartimento específico da membrana pós-sináptica de sinapses excitatórias, conhecido como densidade pós-sináptica (PSD) (COSTALES; KOLEVZON, 2015; HARONY-NICOLAS *et al.*, 2015; LEE *et al.*, 2017).

A isoforma completa da SHANK abriga cinco domínios de interação proteína-proteína, o domínio de repetição de anquirina (ANK), um domínio Src homólogo 3 (SH3), um domínio PDZ, uma região rica em prolina e um domínio α -motivo estéril (SAM) (Figura 15). Através do domínio SAM, as proteínas SHANK podem se auto montar para formar uma plataforma SHANK no PSD, que desempenha um papel vital na sinaptogênese, manutenção sináptica e plasticidade. Através de outros domínios, as proteínas SHANK também podem formar interação direta e indireta com proteínas de suporte adicionais dentro do PSD, com os receptores glutamatérgicos (receptores AMPA e NMDA), com moléculas de adesão celular na membrana e com componentes do citoesqueleto dentro da pós-sinapse (LEBLOND *et al.*, 2014; WANG *et al.*, 2014; HARONY-NICOLAS *et al.*, 2015; JARAMILLO *et al.*, 2017).

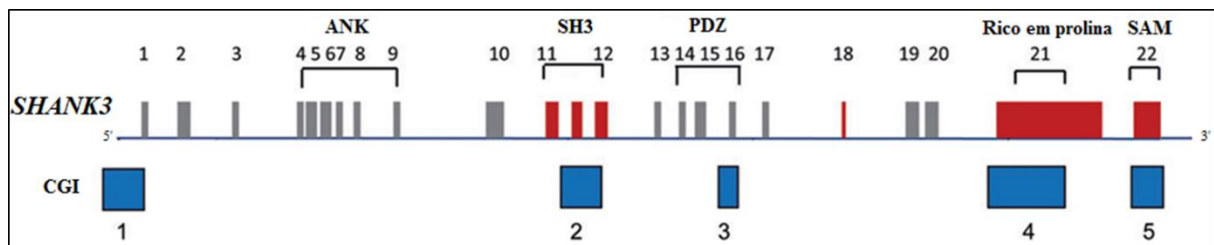


Figura 15. Desenho esquemático do gene *SHANK3* com representação dos exons e domínios e das ilhas CpG (barras azuis). ANK=domínio de repetição de anquirina; SH3=domínio de homologia Src 3; PDZ=discos de zonas longas; Rico em prolina=região rica em prolina; SAM=domínio α -motivo estéril.

Fonte: Adaptado de LEBLOND *et al.*, 2014 e COSTALES; KOLEVZON, 2015.

As isoformas de *SHANK3* são altamente expressas em todas as regiões do cérebro durante o início do desenvolvimento pós-natal, sendo que picos de expressão se correlacionam com um aumento significativo na sinaptogênese e maturação sináptica, processo crítico no aprendizado e desenvolvimento da memória. Além disso, há uma concentrada expressão de *SHANK3* no córtex cerebral e cerebelo no humano adulto (COSTALES; KOLEVZON, 2015; PHILIPPE *et al.*, 2015; GÖRKER *et al.*, 2017).

O *SHANK3* é um dos genes mais ricos em GC no genoma, apresentando 5 ilhas CpG (CGIs), na região 5' do promotor e dentro do corpo do gene (CGIs intragênicas), que exibem um padrão de metilação de DNA específico numa região do cérebro (Figura 13). As CGIs intragênicas estão implicadas na expressão específica das isoformas e expressão tecido específica do *SHANK3* (ZHU *et al.*, 2013; HARA *et al.*, 2015). Zhu e colaboradores (2013), observaram que a metilação do DNA nas ilhas CpG intragênicas de *SHANK3* estava significativamente aumentada em tecidos cerebrais pós-morte de indivíduos com autismo, mostrando o papel causal da metilação do DNA na alteração da expressão específica de isoformas de *SHANK3*.

O gene *SHANK3* está localizado na região crítica da Síndrome de Deleção 22q13.3, também conhecida como Síndrome Phelan-McDermid (OMIM606232), caracterizada clinicamente por AGD, DI moderada à profunda, atraso ou ausência da fala, TEA, hipotonia e características faciais dismórficas distintas. (WANG *et al.*, 2014; LEE *et al.*, 2017; TERRONE *et al.*, 2017).

Deleções e mutações de ponto no gene *SHANK3* envolvendo mais exons proximais têm sido observadas em indivíduos com AGD, DI e autismo, sendo a haploinsuficiência do *SHANK3* e alterações epigenéticas no gene também responsáveis por manifestações neuropsiquiátricas (OBERMAN *et al.*, 2015; REIM *et al.*, 2017; TERRONE *et al.*, 2017). Por outro lado, duplicações estão associadas ao transtorno de déficit de atenção e hiperatividade e desordem bipolar (CHOI *et al.*, 2015; LEE *et al.*, 2017).

f) Gene *ILIRAPL1*

O gene *ILIRAPL1*, conhecido como gene da Proteína Acessória Receptora de Interleucina tipo 1 (OMIM300206), foi descrito e mapeado na região Xp22.1-p21.3, por Carrié e colaboradores (1999). É um gene sinaptogênico e está localizado distal ao gene *DMD* (distrofina), *GK* (glicerol quinase) e *NROBI* (receptor nuclear subfamília 0, grupo B, membro 1) e proximal ao gene *ARX* (homeobox relacionado com *Aristaless*) (CARRIÉ *et al.*, 1999; BERTELLI *et al.*, 2016).

É um gene que contém 11 exons, dos quais 10 exons codificam uma proteína que é um membro da família de receptores de interleucina 1, chamada Proteína Acessória do Receptor de Interleucina 1 tipo 1 (*ILIRAPL1*), que pertence à família de receptores Toll/IL-1. Estruturalmente esta proteína é formada por três domínios extracelulares de imunoglobulinas-like (Ig-like), um único domínio transmembrana, um domínio intracelular Toll/IL-1R (TIR) e

uma longa calda C-terminal que interage com múltiplas moléculas de sinalização (Figura 16). Esta região C-terminal não é compartilhada com os outros membros da família de receptores IL-1. Em contraste aos outros membros desta mesma família, o *IL1RAPL1* é seletivamente expresso em alto nível nas estruturas cerebrais envolvidas no sistema de memória do hipocampo e está localizado nas sinapses excitatórias com um enriquecimento no compartimento pós-sináptico, desempenhando um importante papel na formação e estabilização da sinapse (RAMOS-BROSSIER *et al.*, 2014; LAINO *et al.*, 2016; MONTANI *et al.*, 2017).

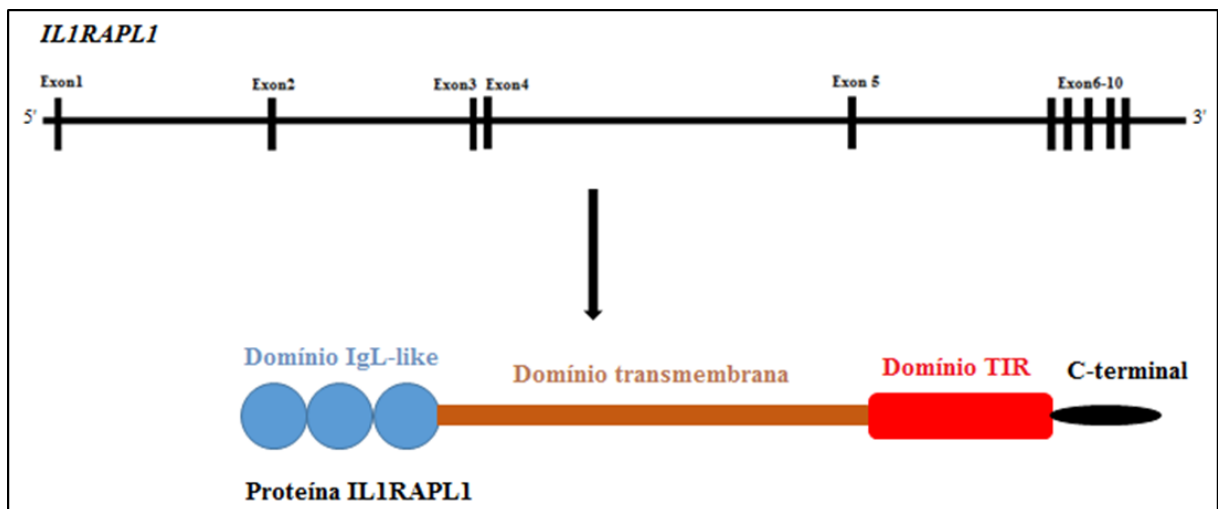


Figura 16. Representação esquemática do gene *IL1RAPL1* e da proteína transcrita. Os domínios da proteína estão apresentados com formatos diferentes.

Fonte: Adaptado de YOSHIDA *et al.*, 2011 e RAMOS-BROSSIER *et al.*, 2014.

A *IL1RAPL1* desempenha papel relevante na diferenciação pré-sináptica e na formação e estabilização da coluna dendrítica em neurônios corticais e hipocámpais. Além disso, a *IL1RAPL1* está envolvida na via de sinalização trans-sináptica e pode também regular a formação e estabilização de sinapses glutamatérgicas, no qual envolve o glutamato principal neurotransmissor excitatório do SNC, entre neurônios corticais (YOSHIDA *et al.*, 2011; LAINO *et al.*, 2016; MONTANI *et al.*, 2017). Yasumura e colaboradores (2014), em um estudo com ratos *knockout*, revelou que a remoção de *IL1RAPL1* afeta diversas funções cerebrais, incluindo o aprendizado, memória, comportamento, flexibilidade, atividade locomotora e ansiedade, demonstrando que múltiplas funções cerebrais são afetadas por mutações neste gene.

CNVs envolvendo o gene *IL1RAPL1*, se por grandes deleções envolvendo a região genômica do braço curto do cromossomo X ou pequenas deleções limitadas ao próprio gene, bem como mutações por perda de função em *IL1RAPL1*, têm sido responsáveis por formas de

DI não sindrômica e também pelo autismo (FRANEK *et al.*, 2011; YASUMURA *et al.*, 2014; RAMOS-BROSSIER *et al.*, 2014; LAINO *et al.*, 2016). Montani e colaboradores (2017), em seu estudo usando culturas neuronais primárias e ratos *knockout* para a caracterização do papel de *ILIRAPLI* na regulação da morfologia do dendrito, demonstrou em ratos e *in vivo*, uma função nova específica do gene na regulação da morfologia do dendrito, indicando que deleções no gene, além de alterar sinapse, também podem alterar o desenvolvimento dos dendritos, ajudando com isso a clarear como as mudanças nas vias reguladas por *ILIRAPLI* podem levar às desordens cognitivas.

g) Gene *DMD*

O gene da distrofina (OMIM300377), também conhecido como gene *DMD*, é um dos maiores genes do genoma humano com 11 kb (0,6%) de sequência codificante, abrange em torno de 2,2 Mb de tamanho, está localizado na região Xp21.1 (BASKIN *et al.*, 2014; OKUBO *et al.*, 2016). O *DMD* codifica uma grande proteína do citoesqueleto contendo um domínio N-terminal de ligação de actina e várias repetições de espectrina, sendo que esta proteína codificada forma um componente do complexo distrofina-glicoproteína (DGC) que faz uma ponte entre o citoesqueleto interno e a matriz extracelular (UCSC Genome Browser).

Este gene possui no total 89 exons e sete isoformas, sendo que sua isoforma maior (Dp427) consiste de 79 exons e 3685 resíduos de aminoácidos. Além da isoforma maior, este gene ainda possui mais quatro isoformas principais: Dp260, Dp140, Dp116 e Dp71. A expressão dessas isoformas é conduzida por promotores específicos para o cérebro, músculo e retina, portanto a expressão de cada uma das isoformas é direcionada para tecidos específicos. A isoforma Dp427 é subdividida em três isoformas, sendo que a Dp427M pode ser expressa em músculo esqueléticos e cardíaco, a Dp427B pode ser expressa no cérebro e, finalmente, a isoforma Dp427P pode ser expressa nas células de Purkinge. Além disso, a isoforma Dp260 é quase que exclusivamente expressa na retina, enquanto que a isoforma Dp140 é expressa no cérebro, retina e rim, e a isoforma Dp116 é expressa somente em nervos periféricos do adulto. Por fim, a menor isoforma, Dp71, é expressa na maioria dos tecidos, tais como cérebro, retina, rim, fígado, pulmão e coração, mas não é expressa no músculo (Figura 17) (DE BROUWER *et al.*, 2014; BANIHANI *et al.*, 2016; RICOTTI *et al.*, 2016; NCBI).

No músculo, a distrofina é essencial para a manutenção da integridade estrutural da fina camada do tecido conjuntivo que envolve a fibra muscular, conhecida como sarcolema. Nos neurônios, o complexo de glicoproteína associada à distrofina está envolvido no

agrupamento pós-sináptico dos canais de íons e dos receptores de membrana pós-sináptica durante a sinaptogênese. A presença de distrofina e do DGC em sinapses de áreas cerebrais distintas como o córtex cerebral, o hipocampo e o cerebelo, sugerem que a perda de função pode ser responsável pelo déficit cognitivo associado a mutações em *DMD* (DE BROUWER *et al.*, 2014; GAO; MCNALLY, 2015).

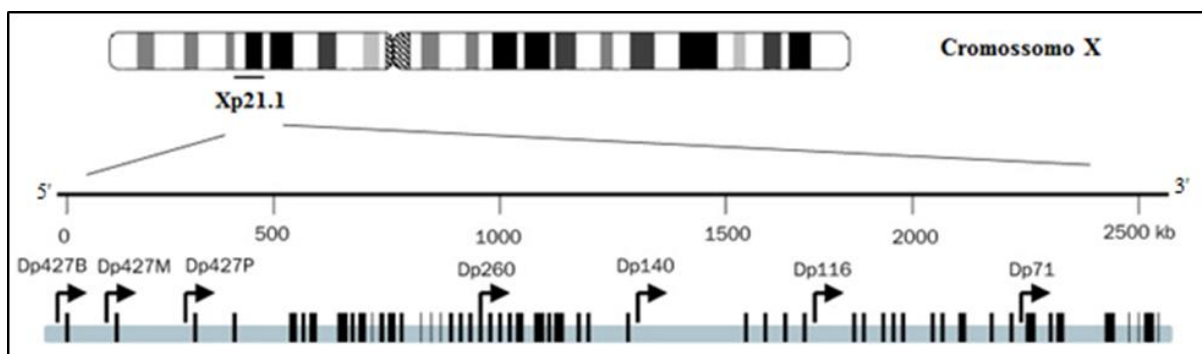


Figura 17. Organização genômica do gene *DMD* e representação das isoformas.

Fonte: Adaptado de RICOTTI *et al.*, 2016.

Mutações no gene *DMD* estão associadas com uma classe de doenças conhecidas como distrofinopatias, que são definidas como um conjunto de patologias neuromusculares debilitantes, incluindo a Distrofia Muscular de Duchenne (OMIM310200), a Distrofia Muscular de Becker (OMIM300376) e a Cardiomiopatia dilatada ligada ao X (OMIM302045). A Distrofia Muscular de Duchenne é uma doença hereditária ligada ao X, é caracterizada pela degeneração muscular progressiva e graus variáveis de comprometimento cognitivo. Indivíduos com BMD compartilham sinais e sintomas similares à DMD, mas está associada com curso clínico mais ameno, início tardio e progressão mais lenta da doença (COTTRELL *et al.*, 2010; GAO; MCNALLY, 2015; FUJIMOTO *et al.*, 2017).

h) Gene *PAK3*

O gene *PAK3*, conhecido como gene da Proteína Quinase Ativada por p21 (OMIM300142), presente em Xq23, é expresso em diferentes regiões do cérebro. Este gene codifica uma proteína serina/treonina quinase que pertence ao grupo I da família PAK (quinase ativada p21). A família PAK age a jusante (*downstream*) a família Rho GTPases (Rac/Cdc42), que regula múltiplos caminhos de transdução de sinal, além de estar envolvida no controle de múltiplos processos intracelulares, incluindo proliferação, progressão do ciclo celular e

remodelação do citoesqueleto (MAGINI *et al.*, 2014; HERTECANT *et al.*, 2017; RENKILARAJ *et al.*, 2017).

A PAK3 é caracterizada pela presença de motivos N-terminal ricos em prolina, que mediam a associação com várias proteínas que contêm o domínio SH3, um domínio N-terminal p21 GTPase ligante (PBD), que confere ligação às pequenas GTPases (Cdc42 ou Rac), e um domínio C-terminal proteína quinase (Figura 18) (HOFMANN *et al.*, 2004; REJEB *et al.*, 2008; VAN DEN BROEKE *et al.*, 2010; RENKILARAJ *et al.*, 2017).

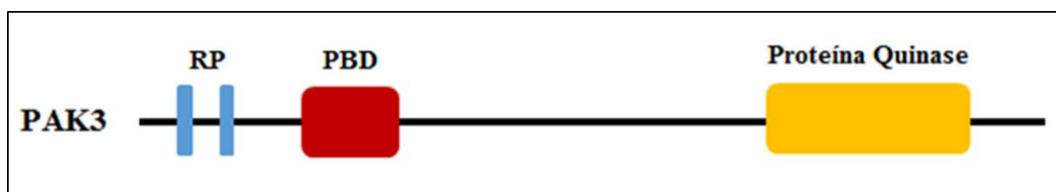


Figura 18. Representação esquemática da proteína PAK3. RP=motivos ricos em prolina; PBD=domínio N-terminal p21 GTPase ligante; Proteína Quinase=domínio C-terminal proteína quinase. **Fonte:** Adaptado de VAN DEN BROEKE *et al.*, 2010; RENKILARAJ *et al.*, 2017.

A PAK3 está envolvida na morfogênese, aperfeiçoamento e controle homeostático da coluna vertebral durante o desenvolvimento, na formação de sinapses neuronais, na plasticidade sináptica, bem como na regulação da dinâmica do citoesqueleto (PEIPPO *et al.*, 2007; BODA *et al.*, 2008; MORROW *et al.*, 2008; DUBOS *et al.*, 2012). Além disso, estudos feitos por Renkilaraj e colaboradores (2017), analisando a expressão de *PAK3* em células oligodendrogliais e os impactos da perda de função de *Pak3* no desenvolvimento de oligodendrócitos e mielinização, conseguiram demonstrar que o gene *PAK3* é altamente expresso em precursores oligodendrócitos, regulando a diferenciação destes precursores.

O gene *PAK3* é crítico nas funções cognitivas e foi descoberto por Allen e colaboradores (1998) como um gene causador de DI ligada ao X. Mutações por perda de função neste gene estão associadas com DI não síndrômica ligada ao X (OMIM300558), no qual pacientes afetados apresentam DI acompanhada por desenvolvimento cerebral grosseiramente normal ou com algumas alterações estruturais e poucos outros sinais ou sintomas (HOFMANN *et al.*, 2004; HERTECANT *et al.*, 2017; RENKILARAJ *et al.*, 2017). As causas da DI em pacientes com mutações em *PAK3* ainda não estão totalmente identificadas, apesar de ser bem aceito que mutações por perda de função neste gene alteram a densidade e morfologia dendrítica, afetam a organização e conectividade dos circuitos cerebrais, além de afetarem o

desenvolvimento de algumas estruturas cerebrais (DUBOS *et al.*, 2012; MAGINI *et al.*, 2014; VERPELLI *et al.*, 2014; RENKILARAJ *et al.*, 2017).

As CNVs de perda, identificadas pelo filtro reduzido, estão localizadas em regiões presentes nos cromossomos 7, cromossomo 22 e cromossomo X, sendo 86,6% destas CNVs de perdas encontradas no cromossomo X. Diversos autores demonstraram que, várias mutações em genes localizados no cromossomo X podem causar DI síndrômica ou DI não-sindrômica (PITON *et al.*, 2013; VERPELLI *et al.*, 2014) e mais de 100 genes presentes no cromossomo X estão associados com deficiência intelectual ligada ao X (TZSCHACH *et al.*, 2015). Das 15 CNVs de perda identificadas nos 12 pacientes com DI, 14 delas estavam presentes em regiões intrônicas dos genes correspondentes e uma CNV de perda presente em 22q13.33 englobava os exons 1, 2 e 3 do gene *SHANK3*.

A presença de CNVs de perdas nas regiões intrônicas pode ocasionar uma ruptura no ponto de ramificação ou na região próxima aos potenciadores ou silenciadores de *splice* intrônicos, prejudicando o processo de *splicing*, comprometendo o correto funcionamento dos mRNAs afetados. Nossos achados estão de acordo com os dados recentes e divulgados na literatura, que têm reportado o efeito deletério de mutações e pequenas perdas ou ganhos presentes em regiões intrônicas sobre o complexo processo de *splicing*, sendo ele constitutivo ou alternativo. Sendo assim, qualquer alteração neste processo pode levar à prejuízos funcionais do mRNA e, conseqüentemente, uma escassez de proteína funcional, que pode resultar em um metabolismo inapropriado e/ou função celular ineficiente, podendo ocasionar doenças genéticas (GRODECKÁ *et al.*, 2017; VAZ-DRAGO *et al.*, 2017; SU *et al.*, 2018).

A robustez e cobertura em todo o genoma dos marcadores polimórficos e não-polimórficos presentes no GeneChip® CytoScanHD™ utilizado no CMA, estabelecidas pela quantidade de marcadores e a distância entre eles, permite, além de uma análise mais precisa e informativa, uma confiabilidade nos resultados obtidos, mesmo quando esta análise é feita com filtro reduzido. De acordo com Nowakowska e colaboradores (2017) e com Harel e Lupski (2018), CNVs muito pequenas podem ser clinicamente importantes e pequenas variantes presentes somente em sequências intrônicas podem ter efeito na inativação do gene ou na modulação da expressão do gene e na sua função.

Dos casos analisados com filtro de tamanho reduzido, todos apresentaram uma cobertura de marcadores distribuídos na região analisada <1.000 pb, demonstrando uma ampla robustez na análise e uma confiabilidade nos resultados obtidos. A redução do filtro de tamanho foi uma alternativa para os pacientes com DI analisados pelo CMA, permitindo a identificação

de CNVs muito pequenas que abrigam genes importante relacionados à DI, explicando os sinais clínicos apresentados pelos nossos pacientes.

8. CONCLUSÃO

A deficiência intelectual é um sinal fenotípico que se manifesta concomitantemente com outras características fenotípicas, demonstrando ser uma condição fenotipicamente e genotipicamente complexa e heterogênea.

No presente estudo, a aplicação de metodologias biotecnológicas recomendadas como de primeira escolha no diagnóstico de pacientes com DI, TEA e ACM, permitiu um aumento de 64,5% no diagnóstico genético dos pacientes incluídos no estudo. Além disso, a redução do filtro de diagnóstico, utilizado na metodologia de CMA, foi útil na identificação de CNVs muito pequenas presentes em regiões intrônicas de genes relacionados à DI. A aplicação do sequenciamento de Sanger para identificação dos pontos de quebra destas CNVs muito pequenas, não permitiu chegar a uma análise conclusiva.

Sendo assim, foi possível:

- Identificar alterações numéricas e estruturais pelo cariótipo com bandeamento GTG, em 57,2% dos pacientes com DI clinicamente diagnosticada. Assim, o cariótipo continua sendo uma metodologia de triagem importante para o diagnóstico laboratorial da DI, sobretudo para os relacionados com síndromes bem estabelecidas e conhecidas. Além disso, a viabilidade do cariótipo por bandeamento GTG pelo SUS, reduz os custos do uso de novas metodologias e, portanto, facilitaria o acesso dos pacientes que realmente necessitam de testes mais resolutivos, como o CMA.
- Detectar alterações genômicas estruturais presentes em regiões que abrigam genes relacionados ao sinal fenotípico da DI e de síndromes recentemente identificadas. Portanto, com o CMA foi possível aumentar o diagnóstico em 29,8% dos pacientes com DI, com um incremento de 7,1%, permitindo chegar a um diagnóstico genético em pacientes que apresentavam o cariótipo por bandeamento GTG sem alteração, comprovando ser uma excelente metodologia de análise de rearranjos estruturais pequenos que estão relacionados com a manifestação clínica deste sinal.
- Verificar, através da análise do CMA com redução do filtro de tamanho, em 45,7% dos casos selecionados, a presença de CNVs em regiões que contém genes relacionados à DI, com o incremento na taxa de diagnóstico de 12%, permitindo o entendimento de que a redução do filtro de tamanho pode ampliar o espectro de identificação de CNVs presentes em regiões que

abrigam genes que são implicados com a manifestação clínica da DI, aumentando assim a possibilidade do diagnóstico genético.

- A identificação dos pontos de quebra, por Sequenciamento de Sanger, das regiões de CNVs de perda identificadas pelo o filtro de tamanho reduzido, não foi eficiente muito provavelmente devido ao mecanismo de reparo, dificultando o sequenciamento exato do ponto de quebra devido à formação de sequências espúrias de alta similaridade com outras regiões do genoma (APÊNDICE 1 e 2). A clonagem pode ser uma estratégia que permitirá um melhor entendimento de todo o mecanismo que culminou na formação das CNVs de perda.

Avanços biotecnológicos estão permitindo um aumento no diagnóstico genético da DI e outras comorbidades e promovendo uma compreensão mais ampla dos aspectos genéticos relacionados à essas afecções, propiciando a identificação de novos rearranjos, mecanismos e genes que são responsáveis pela manifestação clínica deste sinal.

Finalmente, em decorrência desses avanços, atualmente o manejo adequado dessas famílias é mais amplo, com a possibilidade de acesso a diversas terapias que podem melhorar a qualidade de vida da criança e de acesso ao aconselhamento genético, que proporciona uma melhor compreensão das causas genéticas deste sinal fenotípico presente em suas crianças, as implicações familiares da contribuição genética para o aparecimento da doença e também a chance de recorrência, caso queiram ter outros filhos.

9. REFERÊNCIAS BIBLIOGRÁFICAS

AGHA, Zehra *et al.* Exome sequencing identifies three novel candidate genes implicated in intellectual disability. **PloS one**, v. 9, n. 11, p. e112687, 2014.

ALLEN, Kristina M. *et al.* PAK3 mutation in nonsyndromic X-linked mental retardation. **Nature genetics**, v. 20, n. 1, p. 25, 1998.

American Association on Intellectual and Developmental Disabilities (AAIDD). Disponível em: <https://aaid.org/intellectual-disability>. Acesso em: 10/10/2017.

AMERICAN PSYCHIATRIC ASSOCIATION - APA (tradução: Maria Inês Corrêa Nascimento *et al.*). Manual diagnóstico e estatístico de transtornos mentais. 5. ed. – Dados eletrônicos. – Porto Alegre: Artmed, 2014.

American Speech-Language-Hearing Association (ASHA). Disponível em: <http://www.asha.org/Practice-Portal/Clinical-Topics/Intellectual-Disability>. Acesso em: 10/10/2017.

ANNA, Abramowicz; MONIKA, Gos. Splicing mutations in human genetic disorders: examples, detection, and confirmation. **Journal of applied genetics**, v. 59, n. 3, p. 253-268, 2018.

BANIHANI, Rudaina *et al.* A Novel Mutation in DMD (c. 10797+ 5G> A) Causes Becker Muscular Dystrophy Associated with Intellectual Disability. **Journal of Developmental & Behavioral Pediatrics**, v. 37, n. 3, p. 239-244, 2016.

BARTNIK, Magdalena *et al.* Application of array comparative genomic hybridization in 256 patients with developmental delay or intellectual disability. **Journal of applied genetics**, v. 55, n. 1, p. 125-144, 2014.

BASKIN, Berivan *et al.* Complex genomic rearrangements in the dystrophin gene due to replication-based mechanisms. **Molecular genetics & genomic medicine**, v. 2, n. 6, p. 539-547, 2014.

BATTAGLIA, Agatino *et al.* Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. **European journal of paediatric neurology**, v. 17, n. 6, p. 589-599, 2013.

BERTELLI, Marco O. *et al.* “Intellectual developmental disorders”: reflections on the international consensus document for redefining “mental retardation-intellectual disability” in ICD-11. **Advances in mental health and intellectual disabilities**, v. 10, n. 1, p. 36-58, 2016.

BERTOK, Sara *et al.* Clinical and Molecular Cytogenetic Characterisation of Children with Developmental Delay and Dysmorphic Features/Klinična in Molekularna Citogenetska Obravnava Otrok Z Razvojnimi Zaostankom in Displastičnimi Znaki. **Slovenian Journal of Public Health**, v. 54, n. 2, p. 69-73, 2015.

- BODA, Bernadett; JOURDAIN, Lorena; MULLER, Dominique. Distinct, but compensatory roles of PAK1 and PAK3 in spine morphogenesis. **Hippocampus**, v. 18, n. 9, p. 857-861, 2008.
- BOWLING, Kevin M. *et al.* Genomic diagnosis for children with intellectual disability and/or developmental delay. **Genome medicine**, v. 9, n. 1, p. 43, 2017.
- BRETT, Maggie *et al.* Massively parallel sequencing of patients with intellectual disability, congenital anomalies and/or autism spectrum disorders with a targeted gene panel. **PloS one**, v. 9, n. 4, p. e93409, 2014.
- CAPALBO, Antonio; RIENZI, Laura; UBALDI, Filippo Maria. Diagnosis and clinical management of duplications and deletions. **Fertility and sterility**, v. 107, n. 1, p. 12-18, 2017.
- CARRIÉ, Alain *et al.* A new member of the IL-1 receptor family highly expressed in hippocampus and involved in X-linked mental retardation. **Nature genetics**, v. 23, n. 1, p. 25, 1999.
- CHABOT, Benoit; SHKRETA, Lulzim. Defective control of pre-messenger RNA splicing in human disease. **J Cell Biol**, v. 212, n. 1, p. 13-27, 2016.
- CHELLY, Jamel *et al.* Genetics and pathophysiology of mental retardation. **European Journal of Human Genetics**, v. 14, n. 6, p. 701, 2006.
- CHOI, Su-Yeon *et al.* Post-transcriptional regulation of SHANK3 expression by microRNAs related to multiple neuropsychiatric disorders. **Molecular brain**, v. 8, n. 1, p. 74, 2015.
- CHONG, Wilson Wai Sing *et al.* Performance of chromosomal microarray for patients with intellectual disabilities/developmental delay, autism, and multiple congenital anomalies in a Chinese cohort. **Molecular cytogenetics**, v. 7, n. 1, p. 34, 2014.
- CIACCIO, Claudia *et al.* Fragile X syndrome: a review of clinical and molecular diagnoses. **Italian journal of pediatrics**, v. 43, n. 1, p. 39, 2017.
- CLASSEN, Carl Friedrich *et al.* Dissecting the genotype in syndromic intellectual disability using whole exome sequencing in addition to genome-wide copy number analysis. **Human genetics**, v. 132, n. 7, p. 825-841, 2013.
- COSTALES, Jesse L.; KOLEVZON, Alexander. Phelan-McDermid syndrome and SHANK3: implications for treatment. **Neurotherapeutics**, v. 12, n. 3, p. 620-630, 2015.
- COTTRELL, Catherine E. *et al.* Unexpected detection of dystrophin gene deletions by array comparative genomic hybridization. **American Journal of Medical Genetics Part A**, v. 152, n. 9, p. 2301-2307, 2010.
- DE BROUWER, Arjan Pm *et al.* A 3-base pair deletion, c. 9711_9713del, in DMD results in intellectual disability without muscular dystrophy. **European Journal of Human Genetics**, v. 22, n. 4, p. 480, 2014.

DITTWALD, Piotr *et al.* NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. **Genome research**, v. 23, n. 9, p. 1395-1409, 2013.

DUBOS, Aline *et al.* Alteration of synaptic network dynamics by the intellectual disability protein PAK3. **Journal of Neuroscience**, v. 32, n. 2, p. 519-527, 2012.

ELLISON, Jay W.; ROSENFELD, Jill A.; SHAFFER, Lisa G. Genetic basis of intellectual disability. **Annual review of medicine**, v. 64, p. 441-450, 2013.

FONTANELLA, Bianca; RUSSOLILLO, Giorgio; MERONI, Germana. MID1 mutations in patients with X-linked Opitz G/BBB syndrome. **Human mutation**, v. 29, n. 5, p. 584-594, 2008.

FRANEK, Karl J. *et al.* Deletion of the immunoglobulin domain of IL1RAPL1 results in nonsyndromic X-linked intellectual disability associated with behavioral problems and mild dysmorphism. **American Journal of Medical Genetics Part A**, v. 155, n. 5, p. 1109-1114, 2011.

FREDERICKS, Alger *et al.* RNA-binding proteins: splicing factors and disease. **Biomolecules**, v. 5, n. 2, p. 893-909, 2015.

FUJIMOTO, Takahiro *et al.* Dp71 is regulated by phosphorylation and ubiquitin-proteasome system in neuronal cells. **Biochemical and biophysical research communications**, v. 492, n. 3, p. 349-355, 2017.

GALEJ, Wojciech P. Structural studies of the spliceosome: Past, present and future perspectives. **Biochemical Society Transactions**, v. 46, n. 6, p. 1407-1422, 2018.

GAMSIZ, Ece D. *et al.* Intellectual disability is associated with increased runs of homozygosity in simplex autism. **The American Journal of Human Genetics**, v. 93, n. 1, p. 103-109, 2013.

GAO, Quan Q.; MCNALLY, Elizabeth M. The dystrophin complex: Structure, function, and implications for therapy. **Comprehensive Physiology**, 2015.

GEETHA, Thenral S. *et al.* Targeted deep resequencing identifies MID2 mutation for X-linked intellectual disability with varied disease severity in a large kindred from India. **Human mutation**, v. 35, n. 1, p. 41-44, 2014.

GHOLKAR, Ankur A. *et al.* The Mid2 X-linked Intellectual Disability Ubiquitin Ligase Associates with Astrin and Regulates Astrin Levels to Promote Cell Division. **Cell reports**, v. 14, n. 2, p. 180, 2016.

GIGONZAC, Marc Alexandre Duarte *et al.* Standardization of capillary electrophoresis for diagnosis of fragile X syndrome in the Brazilian public health system. **Electrophoresis**, v. 37, n. 23-24, p. 3076-3078, 2016.

- GIJSBERS, Antoinet CJ *et al.* A new diagnostic workflow for patients with mental retardation and/or multiple congenital abnormalities: test arrays first. **European Journal of Human Genetics**, v. 17, n. 11, p. 1394, 2009.
- GIORGIO, Elisa *et al.* Exome sequencing in children of women with skewed X-inactivation identifies atypical cases and complex phenotypes. **European journal of paediatric neurology**, v. 21, n. 3, p. 475-484, 2017.
- GOLDFARB, Mitchell. Fibroblast growth factor homologous factors: evolution, structure, and function. **Cytokine & growth factor reviews**, v. 16, n. 2, p. 215-220, 2005.
- GÖRKER, I. *et al.* A 9-year-old-girl with Phelan McDermid Syndrome, who had been diagnosed with an autism spectrum disorder. **Balkan Journal of Medical Genetics**, v. 19, n. 2, p. 85-90, 2017.
- GRIESCHE, Nadine *et al.* Regulation of mRNA translation by MID1: A common mechanism of expanded CAG repeat RNAs. **Frontiers in cellular neuroscience**, v. 10, p. 226, 2016.
- GRIESI-OLIVEIRA, Karina; SERTIÉ, Andréa Laurato. Autism spectrum disorders: an updated guide for genetic counseling. **Einstein (São Paulo)**, v. 15, n. 2, p. 233-238, 2017.
- GRODECKÁ, Lucie; BURATTI, Emanuele; FREIBERGER, Tomáš. Mutations of pre-mRNA splicing regulatory elements: are predictions moving forward to clinical diagnostics? **International journal of molecular sciences**, v. 18, n. 8, p. 1668, 2017.
- GU, Wenli; ZHANG, Feng; LUPSKI, James R. Mechanisms for human genomic rearrangements. **Pathogenetics**, v. 1, n. 1, p. 4, 2008.
- GUILLEMOT, François; ZIMMER, Céline. From cradle to grave: the multiple roles of fibroblast growth factors in neural development. **Neuron**, v. 71, n. 4, p. 574-588, 2011.
- HARA, Munetsugu *et al.* De novo SHANK3 mutation causes Rett syndrome-like phenotype in a female patient. **American Journal of Medical Genetics Part A**, v. 167, n. 7, p. 1593-1596, 2015.
- HAREL, T.; LUPSKI, J. R. Genomic disorders 20 years on—mechanisms for clinical manifestations. **Clinical genetics**, v. 93, n. 3, p. 439-449, 2018.
- HARONY-NICOLAS, Hala *et al.* Phelan McDermid syndrome: from genetic discoveries to animal models and treatment. **Journal of child neurology**, v. 30, n. 14, p. 1861-1870, 2015.
- HARRIS, James C.; GREENSPAN, Stephen. Definition and nature of intellectual disability. In: **Handbook of evidence-based practices in intellectual and developmental disabilities**. Springer, Cham, p. 11-39, 2016.
- HASTINGS, Philip J. *et al.* Mechanisms of change in gene copy number. **Nature Reviews Genetics**, v. 10, n. 8, p. 551, 2009.

HERTECANT, Jozef *et al.* A de novo mutation in the X-linked PAK3 gene is the underlying cause of intellectual disability and macrocephaly in monozygotic twins. **European journal of medical genetics**, v. 60, n. 4, p. 212-216, 2017.

HOFMANN, Clemens; SHEPELEV, Mikhail; CHERNOFF, Jonathan. The genetics of Pak. **Journal of cell science**, v. 117, n. 19, p. 4343-4354, 2004.

HOWELL, Katherine B. *et al.* High resolution chromosomal microarray in undiagnosed neurological disorders. **Journal of paediatrics and child health**, v. 49, n. 9, p. 716-724, 2013.

HÜNING, Irina *et al.* Exon 2 duplication of the MID1 gene in a patient with a mild phenotype of Opitz G/BBB syndrome. **European journal of medical genetics**, v. 56, n. 4, p. 188-191, 2013.

INOUE, Ken; LUPSKI, James R. Molecular mechanisms for genomic disorders. **Annual review of genomics and human genetics**, v. 3, n. 1, p. 199-242, 2002.

IOUROV, Ivan Y. *et al.* Long contiguous stretches of homozygosity spanning shortly the imprinted loci are associated with intellectual disability, autism and/or epilepsy. **Molecular cytogenetics**, v. 8, n. 1, p. 77, 2015.

JANG, Woori *et al.* Chromosomal Microarray Analysis as a First-Tier Clinical Diagnostic Test in Patients With Developmental Delay/Intellectual Disability, Autism Spectrum Disorders, and Multiple Congenital Anomalies: A Prospective Multicenter Study in Korea. **Annals of laboratory medicine**, v. 39, n. 3, p. 299-310, 2019.

JARAMILLO, Thomas C. *et al.* Novel Shank3 mutant exhibits behaviors with face validity for autism and altered striatal and hippocampal function. **Autism Research**, v. 10, n. 1, p. 42-65, 2017.

Jl, Xing *et al.* A novel mutation in MID1 in a patient with X-linked Opitz G/BBB syndrome. **Gene**, v. 537, n. 1, p. 140-142, 2014.

KARAM, Simone M. *et al.* Genetic causes of intellectual disability in a birth cohort: A population-based study. **American Journal of Medical Genetics Part A**, v. 167, n. 6, p. 1204-1214, 2015.

KATAOKA, Naoyuki. Modulation of aberrant splicing in human RNA diseases by chemical compounds. **Human genetics**, v. 136, n. 9, p. 1237-1245, 2017.

KAUFMAN, Liana; AYUB, Muhammad; VINCENT, John B. The genetic basis of non-syndromic intellectual disability: a review. **Journal of neurodevelopmental disorders**, v. 2, n. 4, p. 182, 2010.

KEARNEY, Hutton M. *et al.* 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, 2011.

KIM, Philip M. *et al.* Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. **Genome research**, v. 18, n. 12, p. 1865-1874, 2008.

KIRIN, Mirna *et al.* Genomic runs of homozygosity record population history and consanguinity. **PLoS one**, v. 5, n. 11, p. e13996, 2010.

KRUSZKA, Paul *et al.* Down syndrome in diverse populations. **American Journal of Medical Genetics Part A**, v. 173, n. 1, p. 42-53, 2017.

LAINO, Luigi *et al.* Clinical and molecular characterization of a boy with intellectual disability, facial dysmorphism, minor digital anomalies and a complex IL1RAPL1 intragenic rearrangement. **European journal of paediatric neurology**, v. 20, n. 6, p. 971-976, 2016.

LEBLOND, Claire S. *et al.* Meta-analysis of SHANK mutations in autism spectrum disorders: a gradient of severity in cognitive impairments. **PLoS genetics**, v. 10, n. 9, p. e1004580, 2014.

LEE, Jennifer A.; CARVALHO, Claudia MB; LUPSKI, James R. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. **cell**, v. 131, n. 7, p. 1235-1247, 2007.

LEE, Yeunkum *et al.* Integrative analysis of brain region-specific Shank3 interactomes for understanding the heterogeneity of neuronal pathophysiology related to SHANK3 mutations. **Frontiers in molecular neuroscience**, v. 10, p. 110, 2017.

LI, Bijun; ZHOU, Tianhong; ZOU, Yi. Mid1/Mid2 expression in craniofacial development and a literature review of X-linked opitz syndrome. **Molecular genetics & genomic medicine**, v. 4, n. 1, p. 95-105, 2016.

LI, Ling-Hui *et al.* Long contiguous stretches of homozygosity in the human genome. **Human mutation**, v. 27, n. 11, p. 1115-1121, 2006.

LISKA, Adam *et al.* Homozygous loss of autism-risk gene CNTNAP2 results in reduced local and long-range prefrontal functional connectivity. **Cerebral Cortex**, p. 1-13, 2017.

LUPSKI, James R. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. **Trends in genetics**, v. 14, n. 10, p. 417-422, 1998.

MAGINI, Pamela *et al.* A mutation in PAK3 with a dual molecular effect deregulates the RAS/MAPK pathway and drives an X-linked syndromic phenotype. **Human molecular genetics**, v. 23, n. 13, p. 3607-3617, 2014.

MARIS, Angelica Francesca *et al.* Familial mental retardation: a review and practical classification. **Ciencia & saude coletiva**, v. 18, p. 1717-1729, 2013.

MICLEA, Diana *et al.* Genetic testing in patients with global developmental delay/intellectual disabilities. A review. **Clujul Medical**, v. 88, n. 3, p. 288, 2015.

- MIGLIORE, Chiara *et al.* Complex rearrangement of the exon 6 genomic region among Opitz G/BBB Syndrome MID1 alterations. **European journal of medical genetics**, v. 56, n. 8, p. 404-410, 2013.
- MILANI, Donatella; RONZONI, Luisa; ESPOSITO, Susanna. Genetic Advances in Intellectual Disability. **Journal of pediatric genetics**, v. 4, n. 03, p. 125-127, 2015.
- MILLER, David T. *et al.* Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. **The American Journal of Human Genetics**, v. 86, n. 5, p. 749-764, 2010.
- MNAYER, Laila *et al.* A structure–function study of MID1 mutations associated with a mild Opitz phenotype. **Molecular genetics and metabolism**, v. 87, n. 3, p. 198-203, 2006.
- MONTANI, Caterina *et al.* The X-linked intellectual disability protein IL1RAPL1 regulates dendrite complexity. **Journal of Neuroscience**, v. 37, n. 28, p. 6606-6627, 2017.
- MORROW, Eric M. *et al.* Sequence analysis of P21-activated kinase 3 (PAK3) in chronic schizophrenia with cognitive impairment. **Schizophrenia research**, v. 106, n. 2, p. 265-267, 2008.
- MOYSÉS-OLIVEIRA, Mariana *et al.* X-linked intellectual disability related genes disrupted by balanced X-autosome translocations. **American Journal of Medical Genetics Part B: Neuropsychiatric Genetics**, v. 168, n. 8, p. 669-677, 2015.
- MUTHUSAMY, Babylakshmi *et al.* Next-Generation Sequencing Reveals Novel Mutations in X-linked Intellectual Disability. **Omics: a journal of integrative biology**, v. 21, n. 5, p. 295-303, 2017.
- NAKAMURA, Takashi *et al.* Novel role of Rac-Mid1 signaling in medial cerebellar development. **Development**, v. 144, n. 10, p. 1863-1875, 2017.
- National Center for Biotechnology Information - NCBI. Disponível em <https://www.ncbi.nlm.nih.gov/gene/1756>. Acesso em: 02/01/2018.
- NICHOLSON, Linda K. Mechanism of midline defect-causing mutation P151L in MID1 revealed. **The FEBS journal**, v. 284, n. 14, p. 2167-2169, 2017.
- NIK, Sara; BOWMAN, Teresa V. Splicing and neurodegeneration: Insights and mechanisms. **Wiley Interdisciplinary Reviews: RNA**, p. e1532, 2019.
- NOWAKOWSKA, Beata. Clinical interpretation of copy number variants in the human genome. **Journal of applied genetics**, v. 58, n. 4, p. 449-457, 2017.
- OBERMAN, Lindsay M. *et al.* Autism spectrum disorder in Phelan-McDermid syndrome: initial characterization and genotype-phenotype correlations. **Orphanet journal of rare diseases**, v. 10, n. 1, p. 105, 2015.

OKUBO, Mariko *et al.* Genetic diagnosis of Duchenne/Becker muscular dystrophy using next-generation sequencing: validation analysis of DMD mutations. **Journal of human genetics**, v. 61, n. 6, p. 483, 2016.

PEIPPO, Maarit *et al.* PAK3 related mental disability: further characterization of the phenotype. **American journal of medical genetics Part A**, v. 143, n. 20, p. 2406-2416, 2007.

PEREIRA, Rodrigo Roncato *et al.* Screening for intellectual disability using high-resolution CMA technology in a retrospective cohort from Central Brazil. **PloS one**, v. 9, n. 7, p. e103117, 2014.

Pesquisa Nacional de Saúde (PNS): 2013: ciclos de vida: Brasil e grandes regiões / IBGE, Coordenação de Trabalho e Rendimento. - Rio de Janeiro: IBGE, p. 92, 2015.

PHILIPPE, Anne *et al.* Case report: an unexpected link between partial deletion of the SHANK3 gene and Heller's dementia infantilis, a rare subtype of autism spectrum disorder. **BMC psychiatry**, v. 15, n. 1, p. 256, 2015.

PITON, Amélie; REDIN, Claire; MANDEL, Jean-Louis. XLID-causing mutations and associated genes challenged in light of data from large-scale human exome sequencing. **The American Journal of Human Genetics**, v. 93, n. 2, p. 368-383, 2013.

POOT, Martin. Connecting the CNTNAP2 networks with neurodevelopmental disorders. **Molecular syndromology**, v. 6, n. 1, p. 7-22, 2015.

POOT, Martin. Intragenic CNTNAP2 deletions: a bridge too far. **Molecular syndromology**, v. 8, n. 3, p. 118-130, 2017.

PRATTE-SANTOS, Rodrigo *et al.* Analysis of chromosomal abnormalities by CGH-array in patients with dysmorphic and intellectual disability with normal karyotype. **Einstein (São Paulo)**, v. 14, n. 1, p. 30-34, 2016.

PREIKŠAITIENĖ, Egle *et al.* Identification of genetic causes of congenital neurodevelopmental disorders using genome wide molecular technologies. **Acta medica Lituanica**, v. 23, n. 2, p. 73, 2016.

PURANAM, Ram S. *et al.* Disruption of Fgf13 Causes Synaptic Excitatory–Inhibitory Imbalance and Genetic Epilepsy and Febrile Seizures Plus. **Journal of Neuroscience**, v. 35, n. 23, p. 8866-8881, 2015.

QUADERI, Nandita A. *et al.* Opitz G/BBB syndrome, a defect of midline development, is due to mutations in a new RING finger gene on Xp22. **Nature genetics**, v. 17, n. 3, p. 285, 1997.

RAFATI, Maryam *et al.* "Familial" versus "Sporadic" intellectual disability: contribution of common microdeletion and microduplication syndromes. **Molecular cytogenetics**, v. 5, n. 1, p. 9, 2012a.

- RAFATI, Maryam *et al.* "Familial" versus "sporadic" intellectual disability: contribution of subtelomeric rearrangements. **Molecular cytogenetics**, v. 5, n. 1, p. 4, 2012b.
- RAMOS-BROSSIER, Mariana *et al.* Novel IL1RAPL1 mutations associated with intellectual disability impair synaptogenesis. **Human molecular genetics**, v. 24, n. 4, p. 1106-1118, 2014.
- REDIN, Claire *et al.* Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. **Journal of medical genetics**, v. 51, n. 11, p. 724-736, 2014.
- REIM, Dominik *et al.* Proteomic Analysis of Post-synaptic Density Fractions from Shank3 Mutant Mice Reveals Brain Region Specific Changes Relevant to Autism Spectrum Disorder. **Frontiers in molecular neuroscience**, v. 10, p. 26, 2017.
- REJEB, Imen *et al.* A novel splice mutation in PAK3 gene underlying mental retardation with neuropsychiatric features. **European Journal of Human Genetics**, v. 16, n. 11, p. 1358, 2008.
- RENKILARAJ, Majistor Raj Luxman Maglorius *et al.* The intellectual disability protein PAK3 regulates oligodendrocyte precursor cell differentiation. **Neurobiology of disease**, v. 98, p. 137-148, 2017.
- REUTER, Miriam S. *et al.* Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. **JAMA psychiatry**, v. 74, n. 3, p. 293-299, 2017.
- RICOTTI, Valeria *et al.* Neurodevelopmental, emotional, and behavioural problems in Duchenne muscular dystrophy in relation to underlying dystrophin gene mutations. **Developmental Medicine & Child Neurology**, v. 58, n. 1, p. 77-84, 2016.
- RODENAS-CUADRADO, Pedro; HO, Joses; VERNES, Sonja C. Shining a light on CNTNAP2: complex functions to complex disorders. **European journal of human genetics**, v. 22, n. 2, p. 171, 2014.
- RODENAS-CUADRADO, Pedro *et al.* Characterisation of CASPR2 deficiency disorder-a syndrome involving autism, epilepsy and language impairment. **BMC medical genetics**, v. 17, n. 1, p. 8, 2016.
- RODRÍGUEZ-REVENGA, Laia *et al.* A parallel study of different array-CGH platforms in a set of Spanish patients with developmental delay and intellectual disability. **Gene**, v. 521, n. 1, p. 82-86, 2013.
- ROSELLÓ, Mónica *et al.* Phenotype profiling of patients with intellectual disability and copy number variations. **European journal of paediatric neurology**, v. 18, n. 5, p. 558-566, 2014.
- SCHWEIGER, Susann; SCHNEIDER, Rainer. The MID1/PP2A complex: a key to the pathogenesis of Opitz BBB/G syndrome. **Bioessays**, v. 25, n. 4, p. 356-366, 2003.

SCOTTI, Marina M.; SWANSON, Maurice S. RNA mis-splicing in disease. **Nature Reviews Genetics**, v. 17, n. 1, p. 19, 2016.

SHI, Yigong. Mechanistic insights into precursor messenger RNA splicing by the spliceosome. **Nature reviews Molecular cell biology**, v. 18, n. 11, p. 655, 2017.

SHORT, Kieran M. *et al.* MID1 and MID2 homo- and heterodimerise to tether the rapamycin-sensitive PP2A regulatory subunit, alpha 4, to microtubules: implications for the clinical variability of X-linked Opitz GBBB syndrome and other developmental disorders. **BMC cell biology**, v. 3, n. 1, p. 1, 2002.

SILVA, Marisa *et al.* European guidelines for constitutional cytogenomic analysis. **European Journal of Human Genetics**, p. 1, 2019.

SMOGAVEC, Mateja *et al.* Eight further individuals with intellectual disability and epilepsy carrying bi-allelic CNTNAP2 aberrations allow delineation of the mutational and phenotypic spectrum. **Journal of medical genetics**, p. jmedgenet-2016-103880, 2016.

STROM, Samuel P. *et al.* Assessing the necessity of confirmatory testing for exome-sequencing results in a clinical molecular diagnostic laboratory. **Genetics in medicine**, v. 16, n. 7, p. 510, 2014.

SU, Chun-Hao *et al.* Alternative splicing in neurogenesis and brain development. **Frontiers in molecular biosciences**, v. 5, p. 12, 2018.

SUZUKI, Makoto *et al.* MID1 and MID2 are required for *Xenopus* neural tube closure through the regulation of microtubule organization. **Development**, v. 137, n. 14, p. 2329-2339, 2010.

TAN, C. A. *et al.* Characterization of patients referred for non-specific intellectual disability testing: the importance of autosomal genes for diagnosis. **Clinical genetics**, v. 89, n. 4, p. 478-483, 2016.

TERRONE, Gaetano *et al.* A novel SHANK3 interstitial microdeletion in a family with intellectual disability and brain MRI abnormalities resembling Unidentified Bright Objects. **European Journal of Paediatric Neurology**, v. 21, n. 6, p. 902-906, 2017.

TZSCHACH, Andreas *et al.* Next-generation sequencing in X-linked intellectual disability. **European Journal of Human Genetics**, v. 23, n. 11, p. 1513, 2015.

University of California Santa Cruz - UCSC - Genome Browser. Disponível em: <https://genome.cse.ucsc.edu>. Acesso em: 02/01/2018.

UNTERBRUNER, Kristoffer *et al.* MicroRNAs miR-19, miR-340, miR-374 and miR-542 regulate MID1 protein expression. **PLoS one**, v. 13, n. 1, p. e0190437, 2018.

VAN BOKHOVEN, Hans. Genetic and epigenetic networks in intellectual disabilities. **Annual review of genetics**, v. 45, p. 81-104, 2011.

VAN DEN BROEKE, Celine *et al.* An emerging role for p21-activated kinases (Paks) in viral infections. **Trends in cell biology**, v. 20, n. 3, p. 160-169, 2010.

VASUDEVAN, Pradeep; SURI, Mohnish. A clinical approach to developmental delay and intellectual disability. **Clinical Medicine**, v. 17, n. 6, p. 558-561, 2017.

VAZ-DRAGO, Rita; CUSTÓDIO, Noélia; CARMO-FONSECA, Maria. Deep intronic mutations and human disease. **Human genetics**, v. 136, n. 9, p. 1093-1111, 2017.

VELTMAN, Joris A.; BRUNNER, Han G. De novo mutations in human genetic disease. **Nature Reviews Genetics**, v. 13, n. 8, p. 565, 2012.

VERMA RS, BABU A. Human chromosomes principles and techniques. 2nd edn. McGraw-Hill, New York. 1995.

VERPELLI, Chiara *et al.* Molecular basis for prospective pharmacological treatment strategies in intellectual disability syndromes. **Developmental neurobiology**, v. 74, n. 2, p. 197-206, 2014.

VISSERS, Lisenka ELM; GILISSEN, Christian; VELTMAN, Joris A. Genetic studies in intellectual disability and related disorders. **Nature Reviews Genetics**, v. 17, n. 1, p. 9, 2016.

WANG, Xiaoming *et al.* Transcriptional and functional complexity of Shank3 provides a molecular framework to understand the phenotypic heterogeneity of SHANK3 causing autism and Shank3 mutant mice. **Molecular autism**, v. 5, n. 1, p. 30, 2014.

WECKSELBLATT, Brooke; RUDD, M. Katharine. Human structural variation: mechanisms of chromosome rearrangements. **Trends in Genetics**, v. 31, n. 10, p. 587-599, 2015.

WESTERINEN, H. *et al.* The nationwide register-based prevalence of intellectual disability during childhood and adolescence. **Journal of Intellectual Disability Research**, v. 61, n. 8, p. 802-809, 2017.

WILL, Cindy L.; LÜHRMANN, Reinhard. Spliceosome structure and function. **Cold Spring Harbor perspectives in biology**, v. 3, n. 7, p. a003707, 2011.

WINTER, Jennifer *et al.* The MID1 protein is a central player during development and in disease. **Front Biosci**, v. 21, p. 664-682, 2016.

World Health Organization (WHO). Disponível em:
<http://www.who.int/mediacentre/factsheets/fs396/en/>. Acesso em: 10/10/2017.

WRIGHT, Katharine M.; DU, Haijuan; MASSIAH, Michael A. Structural and functional observations of the P151L MID1 mutation reveal alpha4 plays a significant role in X-linked Opitz Syndrome. **The FEBS journal**, 2017.

WU, Qing-Feng *et al.* Fibroblast growth factor 13 is a microtubule-stabilizing protein regulating neuronal polarization and migration. **Cell**, v. 149, n. 7, p. 1549-1564, 2012.

YAN, Chuangye; WAN, Ruixue; SHI, Yigong. Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. **Cold Spring Harbor perspectives in biology**, v. 11, n. 1, p. a032409, 2019.

YANG, Jing *et al.* FGF13 modulates the gating properties of the cardiac sodium channel Nav1.5 in an isoform-specific manner. **Channels**, v. 10, n. 5, p. 410-420, 2016.

YASUMURA, Misato *et al.* IL1RAPL1 knockout mice show spine density decrease, learning deficiency, hyperactivity and reduced anxiety-like behaviours. **Scientific reports**, v. 4, p. 6613, 2014.

YOSHIDA, Tomoyuki *et al.* IL-1 receptor accessory protein-like 1 associated with mental retardation and autism mediates synapse formation by trans-synaptic interaction with protein tyrosine phosphatase δ . **Journal of Neuroscience**, v. 31, n. 38, p. 13485-13499, 2011.

ZAHIR, Farah R. *et al.* Comprehensive whole genome sequence analyses yields novel genetic and structural insights for Intellectual Disability. **BMC genomics**, v. 18, n. 1, p. 403, 2017.

ZHANG, Xu *et al.* Roles of intracellular fibroblast growth factors in neural development and functions. **Science China Life Sciences**, v. 55, n. 12, p. 1038-1044, 2012.

ZHU, Li *et al.* Epigenetic dysregulation of SHANK3 in brain tissues from individuals with autism spectrum disorders. **Human molecular genetics**, v. 23, n. 6, p. 1563-1578, 2013.

ŽILINA, Olga *et al.* Chromosomal microarray analysis as a first-tier clinical diagnostic test: Estonian experience. **Molecular genetics & genomic medicine**, v. 2, n. 2, p. 166-175, 2014.

ZWEIER, Christiane. Severe intellectual disability associated with recessive defects in CNTNAP2 and NRXN1. **Molecular syndromology**, v. 2, n. 3-5, p. 181-185, 2011.

ANEXO 1 – TCLE DO PROJETO

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/Certidão nascimento _____,

Naturalidade _____, Idade _____,

Endereço _____,

Neste ato representado por mim:

Nome dos representantes legais:

Pai _____,

Naturalidade _____, Idade _____,

Estado Civil _____, Profissão _____,

Endereço _____

Mãe _____

Naturalidade _____, 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

ANEXO 2 – TCLE DO 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 exhibe 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)
..... (insira o nome da pessoa)

nascido(a) em (insira data de nascimento da pessoa)

para o(a) Dr(a)/Prof(a) APARECIDO DIVINO DA CRUZ do NÚCLEO DE PESQUISAS REPLICON, R. 235, n.40, Área IV, 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:

- o Meu nome/ o nome do meu filho(a) não será publicado.
- o 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 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.
- o 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.

o Os organizadores do DECIPHER não podem aceitar qualquer responsabilidade legal pelo uso indevido de informações.

o 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(a) geneticista solicitante do TCLE:.....

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

APÊNDICE 1 – BLAST DOS PRIMEIROS *PRIMERS* E DOS RESULTADOS DO SEQUENCIAMENTO

Gene *CNTNAP2*

Primer F

Homo sapiens contactin associated protein like 2 (CNTNAP2), RefSeqGene on chromosome 7
Sequence ID: [NG_007092.3](#) Length: 2309998 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Plus

Query 1 CAGGAATGCCAATGATGTTG 20
Sbjct 742274 CAGGAATGCCAATGATGTTG 742293

Primer R

Homo sapiens contactin associated protein like 2 (CNTNAP2), RefSeqGene on chromosome 7
Sequence ID: [NG_007092.3](#) Length: 2309998 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Minus

Query 1 TGCATGGGTGACAGAGGA 20
Sbjct 765862 TGCATGGGTGACAGAGGA 765863

Blast das seqüências obtidas

Homo sapiens BAC clone RP11-770K15 from 4, complete sequence
Sequence ID: [AC097533.2](#) Length: 174696 Number of Matches: 2

Score	Expect	Identities	Gaps	Strand
438 bits(237)	1e-118	268/283(95%)	1/283(0%)	Plus/Plus

Query 24 GCAGGCTTGT-AAGACATATACTCATGATCTTGCCTGCAAGATGGGTCCTT 82
Sbjct 89037 GCAGATCTGTAAAGCAAGATCTTACACAGGATCTTGTCTGAGGAGTGGCTCTT 89076
Query 83 TATCCTCATGCAAGCATAAGCCCATAGGAGGGGGGATTAGACCTTGAAGTAACTCTG 142
Sbjct 88977 TATCCTCATGCAAGCATAAGCCCATAGGAGGGGGGATTAGACCTTGAAGTAACTCTG 88918
Query 143 AAAACAATAGATGGCACAGCTCATGAATAAATGATCCAGTCCCTGTTTTCAAGTGGACAG 202
Sbjct 88917 AAAACAATAGATGGCACAGCTCATGAATAAATGATCCAGTCCCTGTTTTCAAGTGGACAG 88858
Query 203 TTCTGGAAGACATTATGCAAGATTTCCAGTGGAAATCAAGTATCTTACTGCTCATAGCAGT 262
Sbjct 88857 TTCTGGAAGACATTATGCAAGATTTCCAGTGGAAATCAAGTATCTTACTGCTCATAGCAGT 88798
Query 263 GATCCCTATAACATGCCATAGATAGGGCTTTTTCCTTTTCT 305
Sbjct 88797 GATCCCTATAACATGCCATAGATAGGGCTTTTTCCTTTTCT 88755

Homo sapiens BAC clone RP11-770K15 from 4, complete sequence
Sequence ID: [AC097533.2](#) Length: 174696 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
411 bits(222)	2e-110	230/234(98%)	0/234(0%)	Plus/Plus

Query 73 TAATGCTTCCAGAACTGTCACCTGAAACAGGGGACTGGATCATTATTATGAGCTG 152
Sbjct 88843 TAATGCTTCCAGAACTGTCACCTGAAACAGGGGACTGGATCATTATTATGAGCTG 88902
Query 133 TGCCATATTGTTTTCAGAGTAACTCAAGTATAATGCCCCACCCTCTGGGGCTTAT 192
Sbjct 88903 TGTCATATTGTTTTCAGAGTAACTCAAGTATAATGCCCCACCCTCTGGGGCTTAT 88962
Query 193 GCTTGCATGAGGATAAAGGAAGCCACCTCTCAACAACACATCTGTGTAAGATCTGT 252
Sbjct 88963 GCTTGCATGAGGATAAAGGAAGCCACCTCTCAACAACACATCTGTGTAAGATCTGT 89022
Query 253 CTTTACAAGATCGCCCACTTGAAGTGTCTGTCCACTGCTGACAGTCAACAT 306
Sbjct 89023 CTTTACAAGATCGCCCACTTGAAGTGTCTGTCCACTGCTGACAGTCAACAT 89076

Gene *SHANK3*

Primer F

Homo sapiens Chromosome 22q13 Cosmid Clone n66c4, complete sequence
Sequence ID: [AC000050.22](#) Length: 42167 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Plus

Query 1 GATCTGCTCACCTGGCTTC 20
Sbjct 1748 GATCTGCTCACCTGGCTTC 1767

Primer R

Homo sapiens Chromosome 22q13 Cosmid Clone n66c4, complete sequence
Sequence ID: [AC000050.22](#) Length: 42167 Number of Matches: 26

Score	Expect	Identities	Gaps	Strand
38.2 bits(19)	1.3	19/19(100%)	0/19(0%)	Plus/Minus

Query 1 GTGCACGCATGCAATCACT 19
Sbjct 13525 GTGCACGCATGCAATCACT 13507

Blast das seqüências obtidas

Homo sapiens BAC clone RP13-492P3 from 7, complete sequence
Sequence ID: [AC080064.6](#) Length: 82149 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
843 bits(456)	0.0	474/483(98%)	0/483(0%)	Plus/Plus

Query 162 GAGATACAGCTCAAGCAGAAAGAGTGAATTTTACTGAGATCTTTTGGCTTTCCAT 221
Sbjct 13418 GAGATAGCAGCTCAAGCCCTCAGGAGTGAATTTTACTGAGATCTTTTGGCTTTCCAT 13477
Query 222 TATGATTCGAAGATGGCTGCTACACACTGGTATATATCCATGTTCTGAAATTAAGAGT 281
Sbjct 13419 TATGATTCGAAGATGGCTGCTACACACTGGTATATATCCATGTTCTGAAATTAAGAGT 13357
Query 282 AAGGGTGAAGTGCAAAATAGCATCTCTCAGTCCCTTTAAGCTATTTTGAACAGGCTTT 341
Sbjct 13358 AAGGGTGAAGTGCAAAATAGCATCTCTCAGTCCCTTTAAGCTATTTTGAACAGTCTTT 13397
Query 342 GGTCCCATTAATCCACATAAATGACTCTGTTAATGTCACCTTACAATCTCTTCTATC 401
Sbjct 13398 GGTCCCATTAATCCACATAAATGACTCTGTTAATGTCACCTTACAATCTCTTCTATC 13657
Query 402 AGACCCAAATAGTTATGCCGCTGCTGAATCTTACTTGAACCTTCAATGTTATTAACAAT 461
Sbjct 13659 AGACCCAAATAGTTATGCCGCTGCTGAATCTTACTTGAACCTTCAATGTTATTAACAAT 13717
Query 462 GTCACCAATCTGCTTTCTTGAACAATTTCTTTCTTGGTTTCAAGACTTCTAATTTT 521
Sbjct 13718 GTCACCAATCTGCTTTCTTGAACAATTTCTTTCTTGGTTTCAAGACTTCTAATTTT 13777
Query 522 CTTTGTGTTAAGTGGACAGACCTTCTCAGTCTTCTTAACTGAATTCCTCATTTTGGCTC 581
Sbjct 13778 CTTTGTGTTAAGTGGACAGACCTTCTCAGTCTTCTTAACTGAATTCCTCATTTTGGCTC 13837
Query 582 AAGTTGATAGTAACCGCTTGGTCAATGTTGTCCTTCTTATTTGATGTCCTGAGTGA 641
Sbjct 13838 AAGTTGATAGTAACCGCTTGGTCAATGTTGTCCTTCTTATTTGATGTCCTGAGTGA 13897
Query 642 TTG 644
Sbjct 13898 TTG 13900

Homo sapiens BAC clone RP13-492P3 from 7, complete sequence
Sequence ID: [AC080064.6](#) Length: 82149 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
758 bits(410)	0.0	443/458(97%)	5/458(1%)	Plus/Minus

Query 172 GTAAGATTCAGACAGCGGGATAACTATCGGCTCTGATAAGAAAGAGATTGTAAGTACAT 231
Sbjct 13691 GTAAGATTCAGACAGCGGGATAACTATCGGCTCTGATAAGAAAGAGATTGTAAGTACAT 13682
Query 292 TAACAAGAGTCAATGAGTGGAGTAATGGGGACCAAGACTTCAAAATAGCTTAAAGA 291
Sbjct 13692 TAACAAGAGTCAATGAGTGGAGTAATGGGGACCAAGACTTCAAAATAGCTTAAAGA 13572
Query 292 GGGACTGAGAGATGCTATTTTGCACCTCACCTTACTCTTAATTCAGGAACATGGATAA 351
Sbjct 13571 GGGACTGAGAGATGCTATTTTGCACCTCACCTTACTCTTAATTCAGGAACATGGATAA 13512
Query 352 TAACTGAGTTGATGACGACCTTCTGAAATCATAATGAAAGACAAAAAGAAAAA 411
Sbjct 13511 TAACTGAGTTGATGACGACCTTCTGAAATCATAATGAAAGACAAAAAGAAAAA 13452
Query 411 AAAATCACTCTGACGGCTTGAAGCTGCTATCTTGGACTTCTAAGTGTATACCCAT 471
Sbjct 13451 AAAATCACTCTGACGGCTTGAAGCTGCTATCTTGGACTTCTAAGTGTATATTTAT 13392
Query 472 ATGCTCTATATTAATGAAACCTTCTGACGATAAAAAATATCTGTAACCAACAAAC 531
Sbjct 13391 ATGCTCTATATTAATGAAACCTTCTGACGATAAAAAATATCTGTAACCAACAAAC 13332
Query 532 AACACATATGGCTCTGTATCTAGTTTCTCAGCTTCACTTGA-TAAAAAATAG-CTGG 589
Sbjct 13331 AACACATATGGCTCTGTATCTAGTTTCTCAGCTTCACTTGA-TAAAAAATAG-CTGG 13272
Query 590 GTGTG-C-GCTTATGCCCCTAATC-AGCACTTTGGGA 624
Sbjct 13271 GTGTGCGGCTTATGCCGATATCCAGCACTTTGGGA 13234

Gene MID2

Primer F

Homo sapiens midline 2 (MID2), RefSeqGene on chromosome X
Sequence ID: [NG_011907.2](#) Length: 112784 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Plus

Query 1 CCCCATTTTTCAGATGAGGA 20
Sbjct 46085 CCCCATTTTTCAGATGAGGA 46104

Primer R

Homo sapiens midline 2 (MID2), RefSeqGene on chromosome X
Sequence ID: [NG_011907.2](#) Length: 112784 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Minus

Query 1 GAGGCTATCCAGGGGAAGAC 20
Sbjct 52465 GAGGCTATCCAGGGGAAGAC 52446

Blast das sequências obtidas

Homo sapiens chromosome 15 clone RP11-112N19 map 15q22, complete sequence
Sequence ID: [AC092754.4](#) Length: 173411 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
442 bits(239)	6e-120	257/265(97%)	4/265(1%)	Plus/Plus

Query 13 ACTTG-CCC-AGTCTATAGCTGATGATGGAGCCATGATTCAGCCAGGCCG-CAAAAT 69
Sbjct 84152 ACTTGCCCAGTCTATAGCTGATAAATGGAGCCATGATTCAGCCAGGCCGCTACT 84211
Query 70 TCCGCACTCTGGGAGCCTTACCAGCATACCACACTACTCCGAAAGACATTTGGCTTCCC 129
Sbjct 84212 TCCGCACTCTGGGAGCCTTACCAGCATACCACACTACTCCGAAAGACATTTGGCTTCCC 84271
Query 130 ACATAAGTGACTTCTAGTCCACGCACATCAGCCATATCTATATGAGCTTGGCTCTCGC 189
Sbjct 84272 ACATAAGTGACTTCTAGTCCACGCACATCAGCCATATCTATATGAGCTTGGCTCTCGC 84331
Query 190 CATCTGACTACCTCGCTTTTAAATGCCCGATTCCAAATAGCTAATGAGCAATCTAAT 249
Sbjct 84332 CATCTGACTACCTCGCTTTTAAATGCCCGATTCCAAATAGCTAATGAGCAATCTAAT 84391
Query 250 TAGCTTGTCAAAGCTCTTCCC-TGG 273
Sbjct 84392 TAGCTTGTCAAAGCTCTTCCCCTGG 84416

Homo sapiens chromosome 15 clone RP11-112N19 map 15q22, complete sequence
Sequence ID: [AC092754.4](#) Length: 173411 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
459 bits(248)	7e-125	263/270(97%)	2/270(0%)	Plus/Minus

Query 13 CTCATTACGATTTGGCAATGGGGCATTAAAAAGCGAGGGTAGTCAGATGGCAGGAGGCC 72
Sbjct 84381 CTCATTACGATTTGGCAATGGGGCATTAAAAAGCGAGGGTAGTCAGATGGCAGGAGGCC 84322
Query 73 AAGCTCATAAGATATGGCTGATGTCGGTGGCTAGCAAGTCACTTATGTGGGAAGCCAA 132
Sbjct 84321 AAGCTCATAAGATATGGCTGATGTCAGTGGCTAGCAAGTCACTTATGTGGGAAGCCAA 84262
Query 133 AATGCTTTCGGAGGTAGTGTGGTATGCTGTAAGGCTCCAGAGTGGGAAGTAGGCGGC 192
Sbjct 84201 AATGCTTTCGGAGGTAGTGTGGTATGCTGTAAGGCTCCAGAGTGGGAAGTAGGCGGC 84202
Query 193 CTGGCTTGAATCATGCTCCATTATCAGCTATAGGACTTGGGGCAAGTGTCTAACT 252
Sbjct 84201 CTGGCTTGAATCATGCTCCATTATCAGCTATAGGACTTGGGGCAAGTGTCTAACT 84142
Query 253 TTAGGTGACTCTGTTCCTC-TC-G33333 280
Sbjct 84141 TTAGGTGACTCTGTTCCTCATCTGTAAAA 84112

Gene FGF13

Primer F

Human DNA sequence from clone RP6-220J16 on chromosome Xq26.1-27.3, complete sequence
Sequence ID: [AL023798.1](#) Length: 108086 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Minus

Query 1 CAGACACGATTGGCATTTG 20
Sbjct 29094 CAGACACGATTGGCATTTG 29075

Primer R

Human DNA sequence from clone RP6-220J16 on chromosome Xq26.1-27.3, complete sequence
Sequence ID: [AL023798.1](#) Length: 108086 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
40.1 bits(20)	0.34	20/20(100%)	0/20(0%)	Plus/Plus

Query 1 GATTCGCTCTAGCCCTCT 20
Sbjct 8894 GATTCGCTCTAGCCCTCT 8913

Blast das sequências obtidas

Homo sapiens BAC clone RP11-67L14 from 2, complete sequence
Sequence ID: [AC079753.7](#) Length: 154214 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
257 bits(139)	2e-64	181/200(91%)	7/200(3%)	Plus/Plus

Query 131 ATTTCTAGGGCTAGGAGCAGGACATCTTTGGGGGCGCTA--ATTCGCGTACAAGGG 187
Sbjct 100804 ATTTCTGGGGTAAAGTCAAGATTTCTTTGGGGGCGCTAGGATGGC-ATTCCA...G 100839
Query 188 AATTCTCAAAGAAATCAACCCTGGTGGGACAGGAAGAGCTGAGCTGAGGCTCTCTGCTC 247
Sbjct 100860 AATTCTCAAAGAAATCAACCCTGGTGGGACAGGAAGAGCTGAGCTGAGGCTCTCTGCTC 100919
Query 248 ATGTGTACTTCTGAGATCATGGAGACAGGTGAGCCTGAGTGCACGCTCACCAGAGCC 307
Sbjct 100920 ATGTGTACTTCTGAGATCATGGAGACAGGTGAGCCTGAGTGCACGCTCACCAGAGCC 100979
Query 308 ACAGCAGAGGGCTAGGAGCC 327
Sbjct 100980 ACAGCAGAGGGGGAGGAGCC 100999

Homo sapiens BAC clone RP11-67L14 from 2, complete sequence
Sequence ID: [AC079753.7](#) Length: 154214 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
294 bits(159)	1e-75	168/172(98%)	1/172(0%)	Plus/Minus

Query 152 CCCATAAATGAATCATAG-AGATTTCAAATCAACCATCGTTTAACTACAGACTGG 210
Sbjct 100812 CCCAGAAATATGAATCATAGCAAATTTCAAATCAACCATCGTTTAACTACAGACTGG 100753
Query 211 GCACATCCAAAAACATATTTTCAGTTTTACAACAGTGCCTGGTGATATCGCACTATT 270
Sbjct 100752 GCACATCCAAAAACATATTTTCAGTTTTACAACAGTGCCTGGTGATATCGCACTATT 100693
Query 271 GTGGAGCAATAAATGACACGGAGCTGAAACACAAACAAATGCCAAATCGT 322
Sbjct 100892 GTGGAGCAATAAATGACACGGAGCTGAAACACAAACAAATGCCAAATCGT 100641

APÊNCIDE 2 – BLAST DOS *PRIMERS* REDESENHADOS E DA NOVA SEQUÊNCIA DO GENE *CNTNAP2*

Gene *CNTNAP2*

Primer F

Homo sapiens contactin associated protein like 2 (*CNTNAP2*), RefSeqGene on chromosome 7
Sequence ID: [NG_007092.3](#) Length: 2309998 Number of Matches: 2

Score	Expect	Identities	Gaps	Strand
44.1 bits(22)	0.022	22/22(100%)	0/22(0%)	Plus/Plus

Query 1 TGAGAGGTAAGAGACAGGCTT 22
Sbjct 742623 TGAGAGGTAAGAGACAGGCTT 742644

Primer R

Homo sapiens contactin associated protein like 2 (*CNTNAP2*), RefSeqGene on chromosome 7
Sequence ID: [NG_007092.3](#) Length: 2309998 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
44.1 bits(22)	0.022	22/22(100%)	0/22(0%)	Plus/Minus

Query 1 ATGCATGGGTATGACAGAGGAA 22
Sbjct 763883 ATGCATGGGTATGACAGAGGAA 763862

Blast das sequências obtidas

homo sapiens BAC clone VMRC66-509K18 from chromosome 8, complete sequence
Sequence ID: [AC278041.1](#) Length: 164228 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
671 bits(363)	0.0	423/451(94%)	8/451(1%)	Plus/Plus

Query 20 AGCGTAAATGATCGTGGTCTCCGCGGGCATGTGCTAATCTGAGC-C-TTGCCT-CGC 74
Sbjct 114498 AGCGTCATTGAATCTGGCTCCAGGCTGGC-TGTG-TAATCCTGAGCACATTGCTTAGGC 114555
Query 75 TCCCTGAA-CTGTTTCTTCAATAGTCACATTCCTCATGGTCAGTGTGAGAACGGGTCAAT 133
Sbjct 114356 TCTCTGAATCTGTTTCTTCAATAGTCACATTCCTCATGGTCAGTGTGAGAACGGTAAAGT 114615
Query 134 TAGTATCTATAAAAATGTTGATCATATAGCCAGGAGAGAGAGAACCTTGGTCTAATCA 193
Sbjct 114618 TAGTATCTATAAAAATGTTGATCATATAGCCAGGAGAGAGAACCTTGGTCTAATCA 114675
Query 194 TGCCATACATCAAAACAGAGGGGAACCTTTAAAACGACACCTCTTCATGTCGGGAGTGG 253
Sbjct 114678 TGCCATACATCAGACAGAGGGGAGCTTTAAAACGACACCTCTTCATGTCGGGAGTGG 114735
Query 254 AGAGTCTCTGAGAAGTAAAAACCACTCTCTTCAGATTAATGGGGAGTAATCCACCAAT 313
Sbjct 114738 AGAGTCTCTGAGAAGTAAAAACCACTCTCTTCAGATTAATGGGGAGTAATCCACCAAT 114795
Query 314 TAGTGAAAAATAGAAATATCAGTTTCACTTAATTTTATAATTTTCAATTACAGAACTAA 373
Sbjct 114798 TAGTGAAAAATAGAAATATCAGTTTCACTTAATTTTATAATTTTCAATTACAGAACTAA 114855
Query 374 TTAGGAAGTATATCCACCTTTTATTGTGTCTCCATTGACAAATTTTGTAGACCTTTT 433
Sbjct 114858 TTAGGAAGTATATCCACCTTTTATTGTGTCTCCATTGACAAATTTTGTAGACCTTTT 114915
Query 434 TAGCTTTTCACTTCTGTGCAATTTCTCTATCA 464
Sbjct 114918 TAGCTTTTCACTTCTGTGCAATTTCTCTGTCA 114946

homo sapiens BAC clone VMRC66-509K18 from chromosome 8, complete sequence
Sequence ID: [AC278041.1](#) Length: 164228 Number of Matches: 1

Score	Expect	Identities	Gaps	Strand
599 bits(324)	5e-167	395/431(92%)	4/431(0%)	Plus/Minus

Query 28 ATTGTCAATGTGTGACAC-CGGAAGGGGGATGA-CTTGTGTAATTAGTTTCTGTAGTT 85
Sbjct 114898 ATTGTCAATG-GAGACAAATAAAGGTGGATAACTT-CCTAATTAGTTTCTGTAAAT 114841
Query 86 GAAAGTCAAACTTCTAAGATCTGGGTATTCTATTTTCACTAATTGGCGATTACC 145
Sbjct 114840 GAAAAATTAAAAATTAAGATGAACCTGATTAATTTTCACTAATTGGCGATTACT 114781
Query 146 CCCGGTTAATTGATATAGAGTGGTTTTTATTCTCAGAGACTCTCCACTCCGGGACAT 205
Sbjct 114780 CCCCAATTAATCTGAAAGAGAGTGGTTTTTACTTCTCAGAGACTCTCCACTCCGGGACAT 114721
Query 206 GAAGAGGTGTCTGTTTAAAAGCTCCCTCTGTCTGATGATGGCATGATTAGGACCAAG 265
Sbjct 114720 GAAGAGGTGTCTGTTTAAAAGCTCCCTCTGTCTGATGATGGCATGATTAGGACCAAG 114661
Query 266 GTTCTCTCTCCAGGCTCTATGCTCAGCATTATAGATACTAATTTACAGTCTCT 325
Sbjct 114680 GTTCTCTCTCCAGGCTCTATGCTCAGCATTATAGATACTAATTTACAGTCTCT 114601
Query 326 ACACCTGACCATGAGGAATGTGACTATGAAAGAAACAGATTACAGAGCCTAAGCAATGTGC 385
Sbjct 114600 ACACCTGACCATGAGGAATGTGACTATGAAAGAAACAGATTACAGAGCCTAAGCAATGTGC 114541
Query 386 TCAGGATTACACAGCCAGCCTGGAGCCAGGATTAATGACGGTCTGTCTGACTCCAAA 445
Sbjct 114540 TCAGGATTACACAGCCAGCCTGGAGCCAGGATTAATGACGGTCTGTCTGACTCCAAA 114481
Query 446 GCCTGGTCTCT 456
Sbjct 114480 GCCTGGTCTCT 114470

CASE REPORT

Open Access



Mosaic Tetrasomy of 9p24.3q21.11 postnatally identified in an infant born with multiple congenital malformations: a case report

Irene Plaza Pinto^{1,2*}, Lysa Bernardes Minasi^{2,3}, Raphael Steckelberg^{4,5}, Claudio Carlos da Silva^{1,2,3,6,7} and Aparecido Divino da Cruz^{1,2,3,6}

Abstract

Background: Supernumerary Marker Chromosomes consist in structurally abnormal chromosomes, considered as an extra chromosome in which around 70% occur as a de novo event and about 30% of the cases are mosaic. Tetrasomy 9p is a rare chromosomal abnormality described as the presence of a supernumerary isochromosome 9p. Clinical features of tetrasomy 9p include a variety of physical and developmental abnormalities.

Case presentation: Herein, we reported a postnatal case of a newborn who died in early infancy with multiple congenital malformations due to a mosaic de novo tetrasomy 9p detected by Chromosomal Microarray Analysis. Conventional cytogenetics analysis of the proband was 47,XY,+mar[45]/46,XY[5]. The parental karyotypes presented no visible numerical or structural alterations. Microarray Analysis of the proband revealed that the marker chromosome corresponded to a mosaic de novo gain at 9p24.3q21.11.

Conclusions: Chromosomal Microarray Analysis was helpful to identify the origin of the supernumerary marker chromosome and it was a powerful tool to carry out genetic diagnostic, guiding the medical diagnosis. Furthermore, the CMA allowed observing at the first time in Central Brazil the tetrasomy 9p and partial tetrasomy 9q in mosaic, encompassing a large duplicated region with several morbid genes, in an infant with multiple congenital malformations.

Keywords: SMC, Chromosome 9, Congenital disorders, Genetics, CMA

Background

Supernumerary Marker Chromosomes (SMC) are structurally abnormal chromosomes whose origin cannot to be adequately established by G banding karyotyping [1, 2]. In general, SMC are seen as an extra chromosome in metaphase spreads with an incidence rate of 0.43/1000 in postnatal cases and about 1/1000 in prenatal testing.

Approximately 70% of SMC occur de novo and about 30% of the cases are mosaic [3, 4].

Tetrasomy 9p is a rare chromosomal abnormality described as the presence of a supernumerary isochromosome 9p, initially found as a SMC. Further investigation of the SMC commonly shows involvement of the entire 9p, or the entire 9p with part of the heterochromatic region of 9q, or yet the entire 9p with heterochromatic region of 9q, and part of the euchromatic material of 9q [5, 6]. Moreover, around 30% of known cases of tetrasomy 9p exhibits chromosome mosaicism. Both constitutive and mosaic SMC of chromosome 9 comprises a clinically noticeable syndrome [3, 7, 8].

Clinical phenotype of tetrasomy 9p includes a variety of physical and developmental abnormalities. Commonly,

* Correspondence: iplazapinto@gmail.com

¹Biotechnology and Biodiversity PhD Program, Federal University of Goiás, Rede Centro Oeste de Pós-Graduação de Pesquisa e Inovação, Rua 235, n. 40, Setor Leste Universitário, Goiânia, GO 74605-050, Brazil

²Replicon Research Group, Department of Biology, Pontifical Catholic University of Goiás, Rua 235, n. 40, Setor Leste Universitário, Goiânia, GO 74605-050, Brazil

Full list of author information is available at the end of the article



© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

patients have distinctive facial appearances with hypertelorism, cleft lip or palate, ear anomalies, and micrognathia [3, 9]. In addition, recurrent clinical features include developmental delay, central nervous system anomaly, limb defects, postnatal growth failure, congenital heart disease, renal anomalies, and short neck with excess nuchal skin [8].

Herein, we report a postnatal case of a newborn who had multiple congenital malformations due to a mosaic de novo tetrasomy of 9p and partial tetrasomy of 9q detected by Chromosomal Microarray Analysis (CMA).

Case presentation

The proband was a preterm newborn boy, the first child of non-consanguineous parents, born at 31 weeks gestation to a 44-year old father and a 43-year old mother by cesarean section. At birth, the child weighed 1,480 g, measured 44 cm in crown-to-heel length, and exhibited multiple congenital anomalies. The newborn was transferred to the Intensive Care Units (ICU) immediately after birth. His general health condition deteriorated progressively, leading to his death at 105th days after birth. The newborn had brain malformation, including ventriculomegaly and corpus callosum dysgenesis, cleft lip and palate, retrognathism, hypertelorism, clenched hands with overlapping fingers, and hypotonia. Additionally, he revealed mild heart septal hypertrophy, ambiguous genitalia, enlarged kidneys without corticomedullary differentiation, and gallbladder with tiny cystic formations (Fig. 1). His mother had three miscarriages from previous marriages and one miscarriage with her current husband. The remaining of his family history was otherwise unremarkable.



Fig. 1 Dismorphological facial features of newborn with multiple congenital malformations from Central Brazil

Both parents signed a written informed consent and the mother signed as the legal representative for the child. Peripheral blood was obtained to isolate genomic DNA for CMA using Qiagen QIAamp[®] DNA Mini kit (Hilden, Germany). Karyotyping was performed in a private laboratory through conventional cell culture, harvesting, and GTG banding with a > 550 bands resolution following standard procedures [10]. Chromosome analyses were done using Zeiss Axio Scope (Jena, Germany) and the software IKAROS[®] (Metasystems Corporation, Altlußheim, Germany). All laboratory procedures were carried out following international standardized protocols and consensual criteria of quality.

The CMA was carried out on proband and his biological parents using the GeneChip[®] CytoScanHD[™] (Affymetrix, Santa Clara, USA) following the manufacturer's recommendations without modifications. Chromosomal analyses were done using the Chromosome Analysis Suite (ChAS[®]) software (Affymetrix, Santa Clara, USA) and the CNVs found in the patient were analyzed in comparison with public databases, including Database of Genomic Variants (DGV), Database of Chromosomal Imbalance and Phenotype in Humans using Ensemble Resources (DECIPHER), and CytoScanHD[™] Array Database. Furthermore, CNVs were classified according to their nature, based on [11, 12].

The proband showed a male karyotype with a large submetacentric SMC in 90% of the analyzed metaphases after counting 50 metaphase spreads. His karyotype was 47,XY,+mar[45]/46,XY[5], suggesting 10% mosaicism. The parental karyotypes and CMA results had no visible numerical or structural alterations. The proband's CMA revealed the marker chromosome corresponded to a de novo 70.77 Mb gain at arr[GRCh37] 9p24.3q21.11(203,861_70,974,662)× 4[0.3] dn with 30% mosaicism, encompassing 286 genes, including 152 OMIM morbid genes (Fig. 2).

Discussion and conclusions

Chromosomal alterations associated with a spectrum of multiple congenital anomalies are most frequently numerical and are identified in 0.3–1% of newborns [13, 14]. Partial tetrasomy 9 is not yet a well recognizable clinical syndrome due to the limited number of postnatal cases described to date. Moreover, patients with partial tetrasomy 9 have variable phenotypes depending on the size and position of the duplicated region and the degree of mosaicism [6–8, 15].

Here we reported a case of a boy who died postneonataly in early infancy whose karyotype indicated an additional SMC which origin could not be determined by GTG-banding. Microarray analysis showed 47,XY,+mar[45]/46,XY[5].arr[GRCh37] 9p24.3q21.11(203,861_70,974,662)× 4[0.3] dn. Copy Number Variations (CNVs) within the region of gain

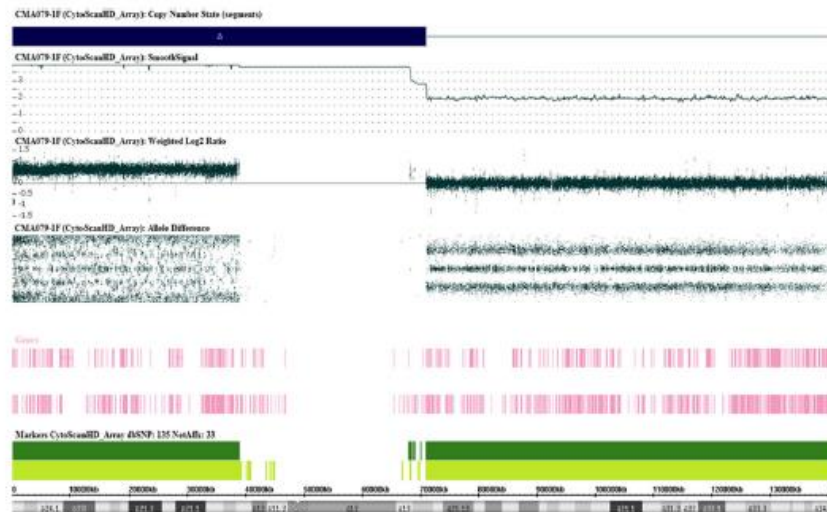


Fig. 2 ChAS* (Affymetrix, USA) interface indicating the results of Chromosomal Microarray Analysis showing the derivative chromosome 9 responsible for tetrasomy 9p extended distally to 9q21.11

indicated tetrasomy, revealing full tetrasomy 9p and partial tetrasomy 9q, including the proximal euchromatic region of the long arm in a mosaic state. Parental CMA were uneventful.

According to [4], about of 40% of all SMCs are derived from non-acrocentric autosomes and the risk of an abnormal phenotype is approximately 28%. Mosaic tetrasomy of 9p most commonly arises as a *de novo* chromosomal rearrangement resulting from an early sporadic error in embryonic development. Thus, the risk of recurrence is relatively low and this has important implications in genetic counseling for the family [16]. Furthermore, it is important to emphasize the mother history of miscarriage because either a cryptic translocation or some aberrant chromosomal constitution restricted to the mother's gonads, most likely as a gonadal mosaicism, could be responsible for her failure to conceive health offspring. However, the family refused to undergo genetic counseling and refused any additional genetic testing.

Tetrasomy 9p is a rare disorder firstly described by [17]. Up to now, at least 34 postnatal cases have been previously reported in medical literature, varying greatly in chromosomal range and phenotype severity. Although the affected population is small, males seemed to be slightly more affected than females [5]. The clinical severity of tetrasomy 9p varies from neonatal death to mild developmental delay and minor anomalies [9, 18, 19]. Furthermore, when the tetrasomic segment extend to 9q22-9q32 it is common to observe intrauterine growth retardation, skeletal anomalies, cleft lip/palate, and heart defects [15].

According to some authors, three different types of supernumerary isochromosome 9p have been described so far. There are tetrasomies containing exclusively the entire 9p, some includes the entire 9p and a small proportion of the heterochromatic region of 9q, and yet some harbors the entire 9p and a large portion of 9q extending to 9q21, including both heterochromatic and euchromatic regions of 9q [3, 5, 6]. The tetrasomy was categorized as the third type of SMC involving chromosome 9, with breakpoints at 9q21.11 and isochromosome 9p, justifying the severity of symptoms in the studied proband.

There is a positive correlation between degree of mosaicism and the severity of the phenotype in infants having tetrasomy 9p. However, the breakpoint of 9q extended distally to 9q21.11, which encompasses a large duplicated region with several OMIM morbid genes. This finding supports the claim the nature of the involved genes was more crucial than the level of mosaicism in defining phenotype's severity in our proband.

The CMA was helpful to identify the origin of the SMC in the proband born to parents with normal karyotypes. In addition, it was an effective method to identify the tetrasomy 9p and partial tetrasomy 9q in mosaic in an infant with multiple congenital malformations for the first time in Central Brazil. Rare and complex phenotypes must always be investigated to define subsets and allow the phenotype/genotype correlation. Furthermore, it was recommended to the family to undergo a non-directive genetic counseling to help understanding the familial implications of genetic contribution to disease and the chance of recurrence.

Abbreviations

ChAS: Chromosome Analysis Suite; CMA: Chromosomal Microarray Analysis; CNVs: Copy Number Variations; DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans using Ensemble Resources; DGV: Database of Genomic Variants; ICU: Intensive Care Units; OMIM: Online Mendelian Inheritance in Man; SMC: Supernumerary Marker Chromosomes; SNP: Single Nucleotide Polymorphism

Acknowledgments

The authors wish to thank Dr. Rinaldo W. Pereira and the Rede ExeGenes for supporting the betterment of genetic diagnosis in Central Brazil. We are also grateful for Mr. Sean M. Quail's contribution for proofreading the manuscript and assistance with the English language. This work was sponsored by a grant from CNPq (Edital 031/564465/2010-10) and FAPEG (2011.6002.19.1847.1134-03). IPP holds a scholarship from CAPES/UFPA.

Availability of data and materials

All data generated or analyzed during this study are included in this published article.

Authors' contributions

IPP, LBM, RS, CCS and ADC have made substantial contributions to conception and design the experiment. RS was the patient's attending physician. IPP performed CMA experiments. IPP, LBM, CCS, and ADC analyzed the data. IPP, ADC and LBM wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study was approved by the Pontifícia Universidade Católica de Goiás ethics approval (CAAE 0051.0.168.000-11) and the family (parents and child) provided written informed consent for publication of this report and the accompanying images. A copy of the written consent is available for review by the Editor-in-Chief of this journal.

Consent for publication

Written informed consent was obtained from the patient's parents for publication of this case report.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Biotechnology and Biodiversity PhD Program, Federal University of Goiás, Rede Centro Oeste de Pós-Graduação de Pesquisa e Inovação, Rua 235, n. 40, Setor Leste Universitário, Goiânia, GO 74605-050, Brazil. ²Replicon Research Group, Department of Biology, Pontifical Catholic University of Goiás, Rua 235, n. 40, Setor Leste Universitário, Goiânia, GO 74605-050, Brazil. ³Genetics Master's Program, Pontifical Catholic University of Goiás, Rua 235, n. 40, Setor Leste Universitário, Goiânia, GO 74605-050, Brazil. ⁴Maternity Hospital Amparo, Av T-12 n° 280 Setor Bueno, Goiânia, GO, Brazil. ⁵Dr Henrique Santillo Rehabilitation and Readjustment Center, Av. Ver. José Monteiro, 1655, Setor Negrão de Lima, Goiânia, GO, Brazil. ⁶Human Cytogenetics and Molecular Genetics Laboratory, Secretary of Goiás State for Public Health, Goiânia, GO, Brazil. ⁷State University of Goiás, ESEFFEGO, Goiânia, Goiás, Brazil.

Received: 23 February 2018 Accepted: 3 September 2018

Published online: 07 September 2018

References

- Liehr T, Cirkovic S, Lalic T, et al. Complex small supernumerary marker chromosomes - an update. *Mol Cytogenet.* 2013;3(16):46-51.
- Liehr T. Small supernumerary marker chromosomes. 2017. <http://smc-tl.com/SSMC.html>. Accessed 08 Feb 2017.
- Jafari-Ghahfarokhi H, Moradi-Chaleshtori M, Liehr T, Hashemzadeh-Chaleshtori M, Teimori H, Ghasemi-Dehkordi P. Small supernumerary marker chromosomes and their correlation with specific syndromes. *Adv Biomed Res.* 2015;27(4):140-6.

- Jang W, Chae H, Kim J, et al. Identification of small marker chromosomes using microarray comparative genomic hybridization and multicolor fluorescent in situ hybridization. *Mol Cytogenet.* 2016;8(9):61-70.
- El Khattabi L, Jaillard S, Andrieux J, et al. Clinical and molecular delineation of Tetrasomy 9p syndrome: report of 12 new cases and literature review. *Am J Med Genet A.* 2015;167(6):1252-61.
- Chen CP, Wang LK, Chen SR, et al. Mosaic tetrasomy 9p at amniocentesis: prenatal diagnosis, molecular cytogenetic characterization, and literature review. *Taiwan J Obstet Gynecol.* 2014;53(1):79-85.
- Düzcan F, Ergin H, Aytan M, Tepeli E, Özdemir Ö, Basaran S. Non-mosaic Tetrasomy 9p in an infant with multiple congenital anomalies. *Gynecol Obstet Reprod Med.* 2007;13(2):114-6.
- Ogino W, Takeshima Y, Nishiyama A, Yagi M, Oka N, Matsuo M. Mosaic tetrasomy 9p case with the phenotype mimicking Klinefelter syndrome and hyporesponse of gonadotropin-stimulated testosterone production. *Kobe J Med Sci.* 2007;53(4):143-50.
- Di Bartolo DL, El Naggar M, Owen R, et al. Characterization of a complex rearrangement involving duplication and deletion of 9p in an infant with craniofacial dysmorphism and cardiac anomalies. *Mol Cytogenet.* 2012;5(1):31-6.
- Verma RS, Babu A. Human chromosomes: Principles and Techniques. 2nd ed. New York: MacGraw-Hill; 1995.
- Miller DT, Adam MP, Aradhya S, et al. Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet.* 2010;86(5):749-64.
- Battaglia A, Doccini V, Bernardini L, et al. Confirmation of chromosomal microarray as a first-tier clinical diagnostic test for individuals with developmental delay, intellectual disability, autism spectrum disorders and dysmorphic features. *Eur J Paediatr Neurol.* 2013;17(6):589-99.
- Starke H, Nietzel A, Weise A, et al. Small supernumerary marker chromosomes (SMCs): genotype-phenotype correlation and classification. *Hum Genet.* 2003;114(1):51-67.
- Pratte-Santos R, Ribeiro KH, Santos TA, Cintra TS. Analysis of chromosomal abnormalities by CGH-array in patients with dysmorphic and intellectual disability with normal karyotype. *Einstein (Sao Paulo).* 2016;14(1):30-4.
- Zhou YC, Zhang C, Zhai JS, et al. A patient with unusual features and a 69.5 Mb duplication from a de novo extra der(9): a case report. *Mol Med Rep.* 2015;12(1):155-8.
- Campbell IM, Shaw CA, Stankiewicz P, Lupski JR. Somatic mosaicism: Implications for disease and transmission genetics. *Trends Genet.* 2015;31(7):382-92.
- Ghymers D, Hermann B, Distèche C, Frederic J. Partial tetrasomy of number 9 chromosome, and mosaicism in a child with multiple malformations (author's transl). *Humangenetik.* 1973;20(3):273-82.
- Frémond ML, Gitiaux C, Bonnet D, et al. Mosaic Tetrasomy 9p: a Mendelian condition associated with pediatric-onset overlap myositis. *Pediatrics.* 2015;136(2):e544-7.
- Oh M, Cho U, Shin S, Lee ST, Choi JR. Isolated 9p duplication with der(Y)t(Y;9)(q12;p13.2) in a male patient with cardiac defect and mental retardation confirmed by chromosomal microarray. *Ann Lab Med.* 2016;36(2):191-3.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions



APÊNDICE 4 – CAPÍTULO DE LIVRO PUBLICADO

Cytogenomic Microarray Testing

Irene Plaza Pinto, Alex da Cruz, Emília Costa,
Samara Pereira, Lysa Minasi and Aparecido da Cruz

Additional information is available at the end of the chapter

Abstract

Cytogenomic microarray testing allows the detection of submicroscopic genomic rearrangements, commonly denominated copy number variations (CNVs) that are implicated with many neurodevelopmental disorders, dysmorphic features, multiple congenital anomalies, hematological and solid tumors, and complex disorders and traits in both humans and animals. On the other hand, this approach is also widely used for the identification of structural variations that are applied as a biomarker in pharmacogenomics, agriculture, and animal selection and breeding. The chromosomal microarray analysis (CMA) has been applied for over a decade to screen for submicroscopic genomic gains and losses in DNA sample in both diagnostic and functional scenarios. Herein, we present an overview of the fundamental concepts of cytogenomics and its potential application in human genetic diagnosis, agrigenomics, mutagenesis, and pharmacogenomics.


Keywords: microarray analysis, SNP array, CNV, array CGH

1. Introduction

Since the use of high-resolution chromosome banding and molecular cytogenetic methodologies, several chromosomal aberrations have been identified. Despite genome-wide detection capacity of these technologies, the rearrangements still remain visibly undetectable, which can be explained by microscopic resolution limitations and the lack of knowledge regarding the regions under investigation.

In the 1990s, the array chromosome-based comparative genomic hybridization (array CGH) was established and began to be used for the detection of significant submicroscopic losses and gains with high sensitivity. It was initially applied to analyze copy number changes in

IntechOpen

© 2018 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

tumors, and it was able to detect changes as small as 50 kb in size, with a resolution up to 1000 times higher than the karyotype. Subsequently, the array CGH methodology was optimized and applied to detect unbalanced constitutional rearrangements [1–4]. Initially, bacterial artificial chromosomes (BACs) and fosmid clones were used in array CGH. However, it became clear that not only larger DNA insert constructed in BAC vector but also PCR products and oligonucleotide sequences were good targets for array CGH [5–7]. Microarray approaches offer a high resolution and relatively quick way for genome-wide analysis, increasing the potential possibilities of genomic scrutiny in clinical scenario as well as its potential application in many other distinct structural genome investigations. In addition, chromosomal microarray analysis (CMA) is useful to estimate the breakpoints of the DNA sequences that can reveal potential mechanisms and risk factors underlying the occurrence of chromosome rearrangement, especially in the case of recurrent rearrangements [8]. Nowadays, chromosomal microarray analysis (CMA) is used as a powerful tool to reveal copy number variants thought to play an important role in the pathogenesis of a variety of disorders or the development of complex traits. Thus, genomic variants can be used as a biological biomarker.

Herein, we present an overview of the fundamental concepts of cytogenomics and the potential application of this technology in human genetic diagnosis, agrigenomics, mutagenesis, and pharmacogenomics.

2. Fundamental concepts of cytogenomics: understanding the tool of arrays

Cytogenomic analysis comprises the use of microarray-based technologies for the investigation of specific loci and the entire genome [9]. It has been used for the detection of copy number variation (CNV), defined as genomic intervals that deviate from the normal diploid state that can vary in size ranging from a few base pairs to mega base pairs [10]. The microarray technologies are frequently nominated as chromosomal microarray analysis (CMA), known as comparative genomic hybridization (CGH) and SNP array (Figure 1). CMA could be used for a dual role in SNP (single nucleotide polymorphism) and CNV-based association studies and in humans for the evaluation of patients with various diseases and congenital malformations [11, 12].

Initially, the CMA was based on the same principles of chromosome-based comparative genomic hybridization. The CGH was developed in the early 1990s to screen for unbalanced rearrangements in whole genomes [3]. The CGH consists in the simultaneous hybridization of labeled test DNA and normal reference DNA onto normal metaphase chromosomes spread on glass slides. Metaphase CGH was widely used to identify chromosomal numerical alterations associated with solid tumors [13, 14]. Overall, the resolution of metaphase CGH is the same of G-banding cytogenetics, limited to 5–15 Mb. Moreover, experiments require cytogenetic expertise for the preparation of suitable metaphase chromosomes [3]. Due to its resolution limitations, the metaphase CGH became restricted to cancer research and did not demonstrate feasibility for analysis of genomic rearrangements in patients with developmental disorders. Subsequently, CGH was implemented as microarrays replacing the metaphases

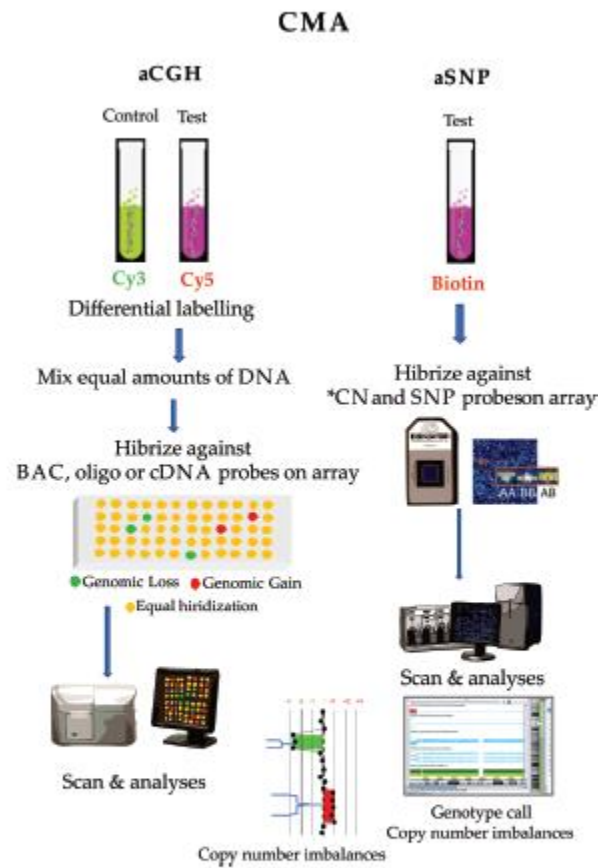


Figure 1. Fundamental concepts of chromosomal microarrays. Currently, the most widely used arrays are comparative genome hybridization arrays (aCGH) and single-nucleotide polymorphism arrays (aSNP). Both microarray-based technologies are able to detect genomic imbalances based on the spectral differences of the fluorescent dyes used to label test and reference DNA in aCGH and test DNA in aSNP. Sample analyses are possible with the aid of computational tools designed to call, view, summarize, and report chromosomal aberrations, including copy number gains or loss, across the genome. *CN: Copy Number.

CGH by cloned DNA segments as genomic reference to be used as targets for the hybridization, improving the potential for the detection of small chromosomal imbalances and increasing its resolution in more than tenfold [14].

The pieces of information regard the DNA sequence of the human genome, which was obtained by the Human Genome Project, making it possible for the construction and characterization of DNA libraries that could be cloned using bacterial artificial chromosomes (BACs) and fosmid as vectors [15, 16]. The array CGH was introduced having known clinically relevant microdeletion genomic DNA segments in BAC clones [10]. However, the relatively large size of the initial clones (170 kb for BACs and 40 kb for fosmids) limited the resolution of the arrays which is dependent on the distances between probes, the length of the clones, and how both probes and clones are distributed across the genome [14, 17]. With the emerging of new protocols, different probes, including small insert clones (1.5–2.5 kb), cDNA clones (0.5–2 kb), PCR products (0.1–1.5 kb), and oligonucleotides (25–85 bp), have been used in the arrays. However, the oligonucleotides have been more appropriated targets for array CGH [5–7, 18]. The oligonucleotide array offers higher resolution and is better than BAC in measuring size of CNVs, increasing the detection of small CNVs. Most CGH arrays available are designed with oligonucleotides ranging from 50 to 70 base pairs (bp) on the probe [19].

According to the International System for Human Cytogenetic Nomenclature [20], in contrast to array CGH, in which DNAs from patient and control are labeled with different dyes and hybridized to a single slide for array CGH and the alterations in the ratio of the two fluorescent dyes indicate a different quantity of DNA in the test sample as compared with the control corresponding locations of CNVs, in the SNP arrays, only the patient's DNA is hybridized to the microarray and compared by computer analysis to a pool of genomic DNA from reference healthy individuals. Additionally, in SNP arrays, the size of the oligonucleotides is about 20 bp and was designed initially to detect genotypes for thousands to hundreds of thousands of SNPs across the entire genome with the focus on genome-wide association studies [19, 21]. SNP arrays can also enable the detection of CNVs, but opposed to array CGH, each probe is located at an SNP and can determine the genotype of the corresponding SNP, and the current SNP arrays with median inter-SNP distances of <0.7 kb ensure the high density of genome coverage [22, 23]. There are many commercial platforms for microarray analysis. Array CGH allows the detection of non-polymorphic region and has been manufactured by Agilent Technologies (Santa Clara, CA) and NimbleGen (Roche Nimble-Gen Inc., Madison, WI), and more recently, Agilent Technologies is offering array CGH with inclusion of SNP markers. The SNP arrays, manufactured by Illumina (San Diego, CA) and by Affymetrix (Santa Clara, CA), have markers for the detection of polymorphic and non-polymorphic regions.

Besides the detection of CNVs, SNP arrays have some advantages in relation to array CGH. The SNP markers can also detect long contiguous stretches of homozygosity (LCSH), low-level mosaic aneuploidies, and chimerism. The detection of LCSH could indicate uniparental isodisomies (UPD) and consanguinity. LCSH distributed in several regions of chromosomes is characterized by genetic identical by descent; on the other hand, when LCSH was identified in a single chromosome, this observation may indicate UPD [24]. The major disadvantage of SNP array and array CGH is the inability to detect balanced chromosome rearrangements because balanced rearrangements show no copy number alterations. Thus, array methodologies do not replace G-banding karyotype for the detection of balanced structural rearrangements. However, they detect abnormalities that are cytogenetically cryptic by G-banded

chromosome analysis. It is important to remember that CMA testing cannot detect balanced karyotypic rearrangements such as reciprocal translocations that could be clinically significant if they disrupt a critical gene. For clinical indications with increased risk for a balanced chromosome rearrangement, such as recurrent pregnancy loss, G-banding chromosome analysis should remain a primary diagnostic test [23]. Additionally, in agrigenomics, SNP arrays have facilitated marker-assisted selections, genome-wide association studies, quantitative trait loci analyses, parentage, and traceability, helping in genomic selection programs.

The primary focus for microarrays has been biomedical-related analyses. However, applications for array technology have broadened to include such fields as plant and animal genotyping and pharmacogenomics. The number of CNVs identified has increased as a function of the increased resolution used by the array technologies. The wide use of arrays has allowed their application in agrigenomics providing a powerful and flexible range of genotyping calls useful for genomic selection programs for plants and animals, helping researchers and breeders to develop healthier and more productive crops and livestock [25, 26]. Besides, array technologies can be used in pharmacogenomic research for the investigation of potential associations between genomic variation and individual drug response. Several SNP-based microarrays are intended to provide information about specific polymorphisms associated with variable drug responses within individuals in a population, which could increase treatment's overall efficacy and decrease the incidence of adverse events [27, 28]. Moreover, the extensive use of CMA has not only contributed to the identification of CNVs and SNP related to human variability but also contributed to the identification of rearrangements implicated in a variety of diseases such as lifestyle diseases, cancer, autoimmune diseases, and neurodevelopmental disorders, including intellectual disability, autism spectrum disorder, global developmental delay, and neuropsychiatric disorders such as schizophrenia, creating a new field of investigation which has transformed the clinical practice [16, 24, 29].

2.1. Application in the diagnosis of human diseases

Genomic gains and losses, defined as CNVs, often cause a wide variety of specific and complex phenotypes, resulting from alterations in the normal dosage of genes, which cause multiple malformation syndromes, neurodevelopmental disorders (NDD), multiple congenital anomalies (MCA), and dysmorphic features. Nowadays, the improved resolution of the microarray technologies has allowed the identification of cryptic chromosomal alterations, increasing the knowledge of the etiology of genomic disorders and offering potential advantages in the patient's follow-up and management [3, 30].

Since 2010, CMA is widely recognized and recommended as the first-tier cytogenomic diagnostic test for individuals with unexplained developmental delay/intellectual disability (DD/ID), autism spectrum disorders (ASDs), or multiple congenital anomalies, increasing the diagnostic yield around 10–25% [16, 30–32]. DD and ID are defined as several significant delays in developmental areas, including cognitive, speech, social/personal, fine/gross motor, and daily activities. DD is described for children less than 5 years old, and ID is diagnosed at or after the age of 5 years old, with the intelligence quotient less than 70. ASD is a complex spectrum of neurodevelopmental disorders, including autism, Asperger syndrome, pervasive

developmental disorder, and childhood disintegrative disorder. MCAs are defined as multiple major structural malformations that cannot be explained by an underlying syndrome or sequence. These disorders might have a genetic etiology involving the gains and losses of CNVs and loss of heterozygosity (LOH), and the clinical consequences of these rearrangements are commonly associated with location, size, and the gene content (Figures 2 and 3) [32–34].

In a study with children with ID/DD, ASD, and/or MCAs from Hong Kong, the application of array CGH demonstrated a diagnostic yield of 11% for pathogenic and likely pathogenic CNVs [35]. Another study with a cohort of 339 patients with neurodevelopmental disorders and/or multiple congenital anomalies using the array CGH identified a detection rate of pathogenic CNVs of 20.6% [30]. Combining both array CGH and SNP array in a single platform, it is possible to make the most effective clinical diagnostic offering simultaneously

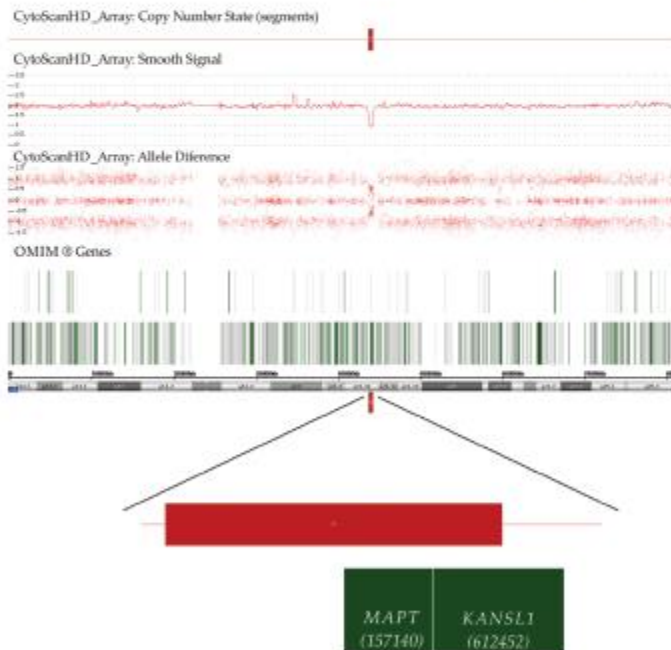


Figure 2 SNP array revealed a copy number loss of chromosome band 17q21.31 of approximately 0.56 Mb in size (arr[hg19] 17q21.31(43,703,801–44,212,416) × 1). This region involves 10 OMIM genes (LOC644172, CRHR1, MGC57346, C17orf69, MAPT-AS1, SPPL2C, MAPT, MAPT-IT1, 57H, and KANSL1) related to Koolen-De Vries Syndrome (MIM610443). The red bar indicates the deleted region and the green bars indicated morbid genes.

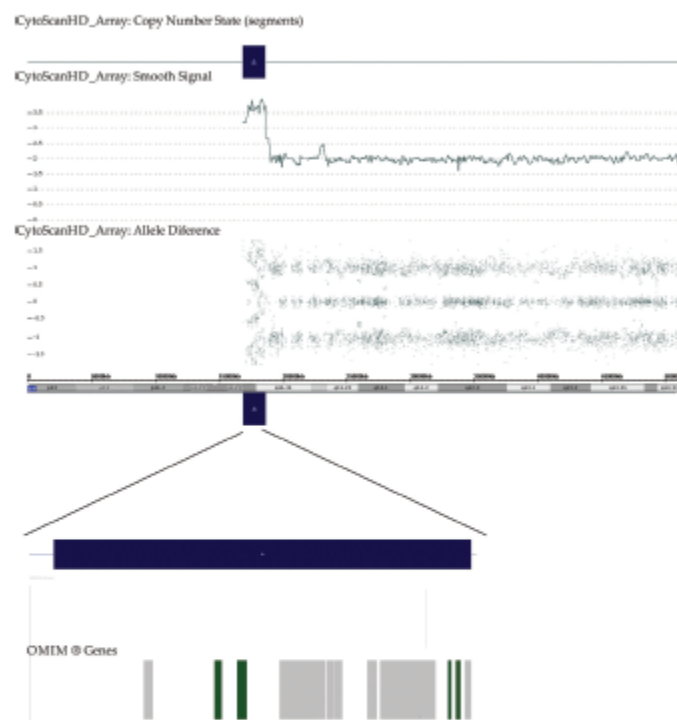


Figure 3. SNP array revealed a gain in copy number of chromosome band 22q11.1, spanning 1.75 Mb (*arr*[GRCh37] 22q11.1q11.21(16,888,899-18,644,773)×4). This gain as indicated in blue bars and comprise 12 OMIM genes (*XKR3*, *IL17RA*, *CECR1*, *CECR2*, *SLC25A18*, *ATP5V1E1*, *BID*, *MICAL3*, *MIR548*, *PEX26*, *TUBA8*, and *USP18*), indicated with gray and green bars, related with Chromosome 22q11.2 Duplication Syndrome (OMIM608363).

identification of CNVs and LOH, as reported by different researches. A comparative study between high-resolution SNP arrays and CGH microarrays revealed that the use of SNP arrays increased the diagnostic yield in children with ID/MCA because these platforms permitted the identification of LOH, which can unravel recessive disorders [36]. Using the combining SNPs with customized exon-targeted oligonucleotide array in a cohort of 3240 patients, Wiszniewska et al. [24] provided a comprehensive approach for the identification of clinically relevant copy number neutral changes in addition to CNVs in a single assay. A study using CMA for 42 Korean patients with unexplained DD, ID, ASD, and MCA identified clinically relevant CNVs in 66.6% of patients [33]. Therefore, microarray-based

technologies have become a powerful tool for the identification of genomic rearrangements smaller than 5 Mb that are associated with neurodevelopmental disorders.

In addition to the usefulness of CMA to help increase the diagnostic rate for ID, DD, ASDs, and MCA, a variety of human conditions, such as epilepsy, schizophrenia, bipolar disorder, and attention deficit hyperactivity disorder, have been reported in some individuals whose CNVs have an important causative role [23, 37–42]. Besides the role of CNVs in the development of neurodevelopmental disorders, CNVs might also contribute to genetic variation in stature in a general population. This conclusion was driven from the observation that in children with short stature lower-frequency copy number variants play a role in the genetic basis of height [43]. Additionally, studies of CNVs in idiopathic short stature (ISS) individuals demonstrated that CMA is a promising approach for the identification of pathogenic CNVs in patients with ISS and could contribute for the recognition of candidate genes associated with growth pathways in humans [44, 45].

CMA testing, especially using SNP-based array, has improved the understanding of genetic causes of many types of human disease and added new pieces of information about gains and losses in the genome in a large variety of hematological malignancies and solid tumors. Additionally, loss of heterozygosity, which is frequently implicated in the tumorigenesis of a variety of cancers, could be detected using SNP array. So, CMA has played an important role to help in diagnosis, prognosis, risk stratification, and therapy for cancer patients [46, 47].

CMA has become a widespread strategy of genetic diagnosis in postnatal settings, especially evaluating children with neurodevelopmental disorders and multiple congenital malformations [48]. Moreover, the implementation of CMA in prenatal settings has helped physicians to identifying chromosomal abnormalities in fetuses harboring anatomical anomalies in the ultrasound, influencing on healthcare delivery in many countries [49]. CMA achieves nearly 100% accuracy rates when applied to identify common aneuploidies in prenatal specimens compared to G-banding karyotyping. Overall, studies showed that in pregnancies with fetal structural anomalies and karyotype with no numerical or structural alterations CMA increased the diagnostic yield around 7%. On the other hand, for all other indications, the increment in the diagnostic yield by CMA has remained around 2% [50, 51]. Taking into consideration the aforementioned information, in 2016 the American Congress of Obstetricians and Gynecologists (ACOG) recommended CMA as the first-tier test for the diagnostic evaluation of fetal structural anomalies. However, the challenge of CMA in prenatal settings is the adequate classification of CNVs as pathogenic and variants of unknown significance (VOUS) [50, 52]. To minimize the reporting of uncertain findings, the practice guideline from Canada issued by the Society of Obstetricians and Gynecologists of Canada (SOGC) and the Canadian College of Medical Geneticists (CCMG) recommended not to report VOUS smaller than 500 kb or VOUS smaller than 1 Mb for losses and gains, respectively [52].

For the diagnosis of human diseases, the microarray platforms should use probes derived from closely spaced genomic loci and have probes concentrated in clinically relevant genes and genomic loci, allowing the detection of smaller CNVs within disease-associated regions. The identification of pathogenic or likely pathogenic CNV by CMA offers benefits for the patient and family, bringing information about prognosis, allowing for appropriate genetic counseling, and adequate patient's management and follow-up for future disclosures [23].

2.2. Application in agrigenomics

2.2.1. Livestock microarray analysis

SNP-based genotyping technologies have become the first-tier methodology in programs of animal and plant breeding for genomic selection [53]. In this context, the use of SNP arrays in organisms of economic interest has facilitated the association between a given SNP with desired productive phenotypes, promoting a positive impact in the economy and contributing to consolidate the technology as a powerful tool to select animals and plants with higher genomic value [54].

The animal genomic is a reality in breeding programs, and the application's impact of these methodologies can be noticed in several areas of animal production [55]. The commercial use of genomic markers is driven based on the need to develop efficient selection and production systems. For instance, selection assisted by genomic markers has been applied to swine breeding aiming for little size and higher meat quality. On the other hand, in bovine, animal selection is used for meat tenderness and higher milk production. Still, for sheep, animal selection is used to increase reproductive efficiency and muscle deposition [56].

Although microarray analysis may reflect a promising future in the agricultural setting to economically produce labor and commodities, the efficiency of microarrays and other innovative methodologies applied to livestock production will always be affected by situations of difficult control, including measurements of traits, phenotypic variance, and low heritability [57]. Inadequate phenotyping could negatively affect, limit, and hinder the usefulness of genetic information in breeding strategies. Improvements and the use of SNP array, followed by cost reduction for genotyping and genome sequencing by Sanger or NGS, have created the possibility to use genomic information for the creation of livestock and supported the emergence of genomic selection programs [58].

The development and application of genomic selection in livestock breeding programs have benefited from the consolidated knowledge generated by classical breeding programs, in particular, information derived from mapping of the Quantitative Trait Loci (QTLs). QTL is characterized as a region of the genome responsible for the expression of a phenotypic trait, which has a continue distribution [25]. This observation is supported by the fact that the main phenotypic characteristics targeted by genomic and genetic breeding programs are polygenic and controlled by several loci, each one adding to the final effect observed in the phenotype. In the aforementioned scenario, genomic selection must be conceived as a process of making decisions regarding the selection of the best-fit animals based on their estimated genomic values. Genomic estimated breeding values (GEBVs) are most commonly obtained by Bayesian models, and it is nowadays considered to be an important step for the success of genomic selection [59]. In summary, GEBVs are the result of the presence of meaningful genetic markers, identified through a dense array of SNPs equally spaced throughout the whole genome, contained within all known QTLs from previously studied livestock [60].

There are different SNP arrays available for livestock genomic analysis. Schaefer et al. [61] designed two different genotyping platforms and demonstrated the application of customized SNP array for domestic horse. Júnior et al. [62] observed the importance of identification

of genomic region associated with puberty and early pregnancy to females of Zebu cattle. Gutierrez et al. [63] used a high-density ovine chip (700 K) to search signatures of selection related to dairy production in sheep and demonstrated the usefulness of the array in the identification of regions of economic interest in dairy sheep. The application of SNP array for genomic selection has been useful to determine genetic attributes and contributed for the genomic selection of traits of economic interest.

The SNP arrays available to estimate genomic values in farm animals can be divided into two categories, according to the time of their development and use. First, there were the SNP arrays from 2000 to 2012, characterized by the development of arrays based on the sequence of reference genomes of farm animals. Most of the arrays were developed by multinational companies, including genome Illumina, Affymetrix, ARK Genomics, and Applied Biotechnologies. Secondly, the arrays used from 2012 on are based on the genome of selected animals, including the possibility of customizing the array for a given property.

In general, the positive aspects for using commercially available SNP arrays are as follows: (a) genetic polymorphism can be estimated in different breeds, including synthetic cross breeds, and (b) the SNPP panels are able to discriminate genomic variability even in animals with close genetic makeup, especially considering the elevated level of inbreeding, for instance, in pigs and poultry lineages. With respect to the negative aspect of using SNP arrays, the following are noteworthy: (a) high cost per genotyped animal, (b) unavailability of personnel and laboratories qualified to generate and analyze genomic data, and (c) in the same array makers for different QTLs and different aptitudes reduction of the availability of genomic data for the trait of interest. Carroll and Charo [55] called the attention upon an array for bovine selection. The markers in that array were chosen to select meat and milk production. The total of informative SNPs was reduced based on the total number of SNPs in the array. Goddard [64] reported that perhaps the two most critical issues that hinder the usefulness of genomic selection in the agribusiness industry are the lack of qualified personnel and the cost per animal for their genotyping.

Lately, the potential of array customization and the advancement of genome sequencing methodologies have boosted the applicability of genotyping farm animals and reduced the cost per animal. Moreover, NGS is a powerful tool to generate information on the whole genome of selected animals together with adequate animal phenotyping, which will contribute to adequately estimate true genomic breeding values for the livestock. Several authors have pointed out that the advancement of genome-wide association studies (GWAS) has providing excellent and efficient information to be used in livestock genomic selection programs (GSP) [54, 57, 65]. GWAS has allowed the identification of candidate genes potentially associated with phenotypes of economic interest (Figure 4). Thus, the piece of information generated for a given herd has become more powerful to predict its genomic merit and also to be used to assist adequate selection of the desired animal phenotype. Thereby, genomic breeding values have become even more useful and accurate, contributing to efficient decision-making by herd managers and producers.

Bosch et al. [67] have addressed the final cost for poultry genotyping considering two distinct arrays from the same company, namely, ChickenSNP50 and ChickenSNP600K, representing

both the early and the state-of-the-art arrays, respectively. The author reported that the genotyping challenge has remained the cost per animal, which has a negative impact on genomic selection. At first, genotyping as a whole has become cheaper. Nevertheless, the cost per animal has not reduced satisfactory. Manufactures claim that the new arrays will collect more significant SNPs to be used in breeding strategies (Figure 5), justifying the increase in the cost per animal. Similar arguments have been displayed by different authors [65, 67, 68] who also claim a steadfast increase in the use of high-density SNP arrays from both academic researchers and commercial facilities to assist with livestock breeding and genomic selection globally.

Much debate around the SNP arrays customarily is used for bovine genotyping, especially for selection of meat and dairy animals. Specialists have claimed that low-density arrays has a reduced capacity to predict the phenotype in *Bos taurus indicus* mostly because the markers in the array are more representative of *Bos taurus Taurus*, affecting mainly the minor allele

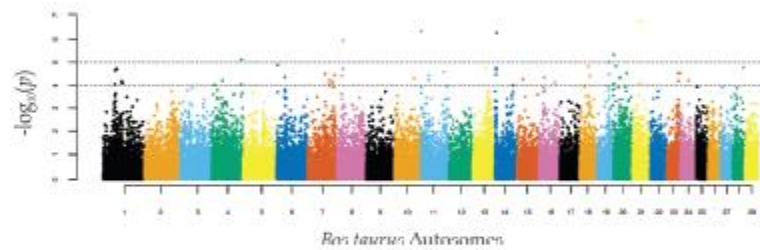
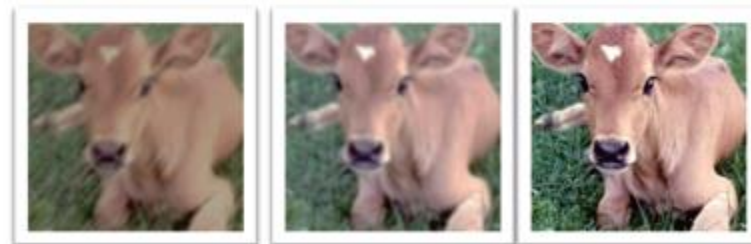


Figure 4 Manhattan plotting showing the distribution of the significant values of SNPs per bovine autosomal chromosomes with respect to 305-day milk yield in Girolando. The GWAS disclosed 7 SNPs associated with milk production trait in dairy cattle with a p value $<10^{-5}$ and a false discovery rate of 6.5% according to the study of milk production in Girolando [66].



3k (3,008 SNPs)

50k (54,608 SNPs)

HD (777,540 SNPs)

Figure 5. The resolution of the arrays. Illustration indicating the resolution of an array is based on the number of markers available to powerfully predict the breed genomic merit related to the desired economical trait.

frequency (MAF) of several SNPs for some breeds, rendering them non-useful for breeding selection [69]. Hickey [70], Auvray et al. [71], and Mrode et al. [68] have addressed the issue of development of future SNP arrays to be applied for bovine selection. They claim that new arrays must include a larger number of markers representative of distinct breeds and/or future arrays must be customized for a specific breed, having less markers, but all chosen based on their MAF. At last, it is also noteworthy to mention that array could be replaced by WGS, pending only on the cost per animal, a challenge that will be met by the manufacturers. Teng and Xiao [72], Bruford et al. [65], and Steyaert et al. [73] considered that as soon as WGS becomes economically competitive it will be accessible to promote a new revolution in the field of livestock genomic selection and breeding.

2.3. Application in research: from mutagenesis to pharmacogenomics

Microarray methodologies have been impacted in different fields of biological science, allowing the identification of genomic alterations applied in the area of mutagenesis, including toxicology, genetic toxicology, as well as pharmacogenomics [74, 75].

In the area of toxicology, gene expression profiles based on microarray analysis can help understand the multiple pathways and mechanisms on the action of toxic substances at the same time. Furthermore, microarray analyses are important to understand the effects of xenobiotics across the genome and the rapid identification of toxic risks of new drugs and chemicals. Thus, global analysis of gene expression has the importance of providing a more comprehensive view of toxicity than was previously possible, since toxicity usually involves changes not only in one or a few genes but is a cascade of gene interactions [74].

Understanding the function of genes is a major challenge in the post genomic era and in order to assign the role of genes in molecular networks, strategies such as proteomics, metabolomics, and transcriptomics have been implemented [76]. The gene expression profile of a cell determines its function, phenotype, and response to the environment. Thus, the analysis of gene expression becomes necessary for the in-depth study of biochemical pathways, regulatory mechanisms, and broader cellular function [77]. Some conventional analyses for the gene expression profile are optimized only for single-gene investigation. Microarrays have been developed as high-performance, efficient, and comprehensive tools for the simultaneous study of multiple genes [78]. Therefore, microarray methodologies are being used to study the transcriptional profile, leading to the research of new genes and molecular markers, having applicability in the field of pharmacogenomics for tracing changes in the expression of genes that are sensitive or resistant to a given drug; thus, it can be used to analyze differential profiles of gene expression that are induced or repressed by xenobiotics [79, 80].

Pharmacogenomic studies of genes and gene products (proteins) are essential for pharmacological or toxicological responses to pharmaceutical agents. In addition, it analyzes genetic determinants of enzymes, receptors, transporters, and targets that metabolize drugs and that influence drug efficacy, safety, and drug-related phenotypes [79]. A current focus of pharmacogenomic research explores the effect of interindividual genetic differences related to drug response by providing information that can be used to inform the appropriate selection of individual drugs or dosing regimens [27, 79, 81]. Pharmacogenomic research involves scanning the entire genome to find single nucleotide polymorphisms (SNPs) that may be

associated with drug responses [79]. Genetic polymorphism studies are performed to classify individuals according to their drug metabolism or disease response capabilities [81].

In a pilot study, Liljedahl et al. [82] developed a microarray genotyping system for multiplex analysis of a panel of SNPs in genes encoding proteins involved in the regulation of blood pressure, demonstrating their viability in response to hypertensive drugs. Therefore, microarray-based tests have provided a useful tool for simultaneous measurement of relative levels of expression of a large number of clinically relevant genes in the context of disease or drug responses [83]. Moreover, the application of the technique in the field of pharmacogenomics characterizes and validates new therapeutic targets, their mechanism of action, metabolic pathways, undesirable side effects, sensitivity, and toxicity to certain drugs [84].

To date, there is a scarcity of studies on the induction of germ line mutations in humans. However, SNP-based arrays can also be applied to monitor individuals exposed to ionizing radiation, and it has been proven to be a useful strategy to evaluate potential health risks related to environmental mutagens. Costa et al. [85] presented results of the analysis of the effect of accidental exposure to low doses of ionizing radiation on the formation of de novo, nonrecurrent CNVs in the progeny of a human population accidentally exposed to cesium-137 during the radiological accident in Goiânia, Brazil. The high-density SNP array used in that study allowed the observation of de novo mutations induced in the germ line of parents exposed to very low doses of ionizing radiation. Although the study of Costa et al. [85] is the pioneer in the field and requires validation, it shed light on the potential of SNP arrays to unravel CVS to be used as useful germ line biomarkers to characterize the exposure of biological systems to mutagenic agents. Thus, a new era of possibilities of using CMA to resolve a variety of biological questions is upon us and once again the future keep on looking promising.

Conflict of interest

The authors declare no conflict of interest.

Author details

Irene Plaza Pinto^{1,2}, Alex da Cruz¹, Emília Costa¹, Samara Pereira¹, Lysa Minazi¹ and Aparecido da Cruz^{1,2,3*}

*Address all correspondence to: acruz@pucgoias.edu.br

¹ Pontifical Catholic University of Goiás, School of Agricultural and Biological Sciences, Genetics Master Program, Replicon Research Group, Goiânia, Goiás, Brazil

² Federal University of Goiás, PhD Program in Biotechnology and Biodiversity, Goiânia, Goiás, Brazil

³ Human Cytogenetics and Molecular Genetics Laboratory, State Laboratory of Public Health Dr. Giovanni Cysneiros, Secretary of Goiás State for Public Health, Goiânia, Goiás, Brazil

References

- [1] Albertson DG, Pinkel D. Genomic microarrays in human genetic disease and cancer. *Human Molecular Genetics*. 2003;12:R145-R152. DOI: 10.1093/hmg/ddg261
- [2] Cai WW, Mao JH, Chow CW, Damani S, Balmain A, Bradley A. Genome-wide detection of chromosomal imbalances in tumors using BAC microarrays. *Nature Biotechnology*. 2002;20:393-396. DOI: 10.1038/nbt0402-393
- [3] Shaffer LG, Bejjani BA. A cytogeneticist's perspective on genomic microarrays. *Human Reproduction Update*. 2004;10:221-226. DOI: 10.1093/humupd/dmh022
- [4] Shaw-Smith C, Redon R, Rickman L, Rio M, Willatt L, Fiegler H, et al. Microarray based comparative genomic hybridization (array-CGH) detects submicroscopic chromosomal deletions and duplications in patients with learning disability/mental retardation and dysmorphic features. *The Journal of Medical Genetics*. 2004;41:241-248. PMID: FMC1735726.
- [5] Brennan C, Zhang Y, Leo C, Feng B, Cauwels C, Aguirre AJ, Kim M, Protopopov A, Chin L. High resolution global profiling of genomic alterations with long oligonucleotide microarray. *Cancer Research*. 2004;64:4744-4748. DOI: 10.1158/0008-5472.CAN-04-1241
- [6] Dharni P, Coffey AJ, Abbs S, Vermeesch JR, Dumanski JP, Woodward KJ, Andrews RM, Langford C, Vetrie D. Exon array CGH: Detection of copy-number changes at the resolution of individual exons in the human genome. *American Journal of Human Genetics*. 2005;76:750-762. DOI: 10.1086/429588
- [7] Wolf M, Mousses S, Hautaniemi S, Karhu R, Huusko P, Allinen M, et al. High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: Impact of copy number on gene expression. *Neoplasia*. 2004;6:240-247. DOI: 10.1593/neo.3439
- [8] Weckselblatt B, Rudd MK. Human structural variation: Mechanisms of chromosome rearrangements. *Trends in Genetics*. 2015;31:587-599. DOI: 10.1016/j.tig.2015.05.010
- [9] Liehr T, editor. *Fluorescence In Situ Hybridization (FISH) - Application Guide*. 2nd ed. Jena: Springer; 2017. 588 p
- [10] Harel T, Lupski JR. Genomic disorders 20 years on-mechanisms for clinical manifestations. *Clinical Genetics*. 2018;93:439-449. DOI: 10.1111/cge.13146
- [11] Haeri M, Gelowani V, Beaudet AL. Chromosomal microarray analysis, or comparative genomic hybridization: A high throughput approach. *MethodsX*. 2015;3:8-18. DOI: 10.1016/j.mex.2015.11.005
- [12] Wang K, Bucan M. Copy Number Variation Detection via High-Density SNP Genotyping. *CSH Protoc*. 2008; pdb.top46. DOI: 10.1101/pdb.top46

- [13] Kallioniemi A, Kallioniemi OP, Sudar DA, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*. 1992;258:818-821. DOI: 10.1126/science.1359641
- [14] Redon R, Carter NP. Comparative genomic hybridization: Microarray design and data interpretation. *Methods in Molecular Biology*. 2009;529:37-49. DOI: 10.1007/978-1-59745-538-1_3
- [15] Morozova O, Marra MA. From cytogenetics to next-generation sequencing technologies: Advances in the detection of genome rearrangements in tumors. *Biochemistry and Cell Biology*. 2008;86:81-91. DOI: 10.1139/O08-003
- [16] Nowakowska B. Clinical interpretation of copy number variants in the human genome. *Journal of Applied Genetics*. 2017;58:449-457. DOI: 10.1007/s13353-017-0407-4
- [17] Manning M, Hudgins L, Professional Practice and Guidelines Committee. Array-based technology and recommendations for utilization in medical genetics practice for detection of chromosomal abnormalities. *Genetics in Medicine*. 2010;12:742-745. DOI: 10.1097/GIM.0b013e3181f8baad
- [18] Carter NP. Methods and strategies for analyzing copy number variation using DNA microarrays. *Nature Genetics*. 2007;39:S16-S21. DOI: 10.1038/ng2028
- [19] Beaudet AL. The utility of chromosomal microarray analysis in developmental and behavioral pediatrics. *Child Development*. 2013;84:121-132. DOI: 10.1111/cdev.12050
- [20] McGowan-Jordan J, Simons A, Schmid M, editors. *An International System for Human Cytogenomic Nomenclature (ISCN)*. 1st ed. Switzerland: Karger; 2016. 139 p
- [21] Ko JM. Chromosomal microarray: Application for congenital heart diseases. *Korean Circulation Journal*. 2018;48:233-235. DOI: 10.4070/kcj.2018.0032
- [22] Keren B. The advantages of SNP arrays over CGH arrays. *Molecular Cytogenetics*. 2014; 7(Suppl1):I31. DOI: 10.1186/1755-8166-7-S1-I31
- [23] Rosenfeld JA, Patel A. Chromosomal microarrays: Understanding genetics of neurodevelopmental disorders and congenital anomalies. *Journal of Pediatric Genetics*. 2017; 6:42-50. DOI: 10.1055/s-0036-1584306
- [24] Wiszniewska J, Bi W, Shaw C, Stankiewicz P, Kang SH, Pursley AN, et al. Combined array CGH plus SNP genome analyses in a single assay for optimized clinical testing. *The European Journal of Human Genetics*. 2014;22:79-87. DOI: 10.1038/ejhg.2013.77
- [25] Boichard D, Chung H, Dassonneville R, David X, Eggen A, Fritz S, et al. Design of a bovine low-density SNP array optimized for imputation. *PLoS One*. 2012;7:e34130. DOI: 10.1371/journal.pone.0034130
- [26] Bovine HapMap Consortium, Gibbs RA, Taylor JF, Van Tassell CP, Barendse W, Eversole KA, et al. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*. 2009;324:528-532. DOI: 10.1126/science.1167936

- [27] Berm EJ, Loeff MD, Wilffert B, Boersma C, Annemans L, Vegter S, Boven JF, Postma MJ. Economic evaluations of pharmacogenetic and pharmacogenomic screening tests: A systematic review. Second update of the literature. *PLoS One*. 2016;11:e0146262. DOI: 10.1371/journal.pone.0146262
- [28] Matson RS. *Applying genomic and proteomic microarray technology in drug discovery*. 1st ed. Boca Raton: CRC Press: Taylor & Francis e-Library; 2005. 72 p.
- [29] Takumi T, Tamada K. CNV biology in neurodevelopmental disorders. *Current Opinion in Neurobiology*. 2018;48:183-192. DOI: 10.1016/j.conb.2017.12.004
- [30] Maini I, Ivanovski I, Djuric O, Caraffi SG, Errichiello E, Marinelli M, et al. Prematurity, ventricular septal defect and dysmorphisms are independent predictors of pathogenic copy number variants: A retrospective study on array-CGH results and phenotypical features of 293 children with neurodevelopmental disorders and/or multiple congenital anomalies. *Italian Journal of Pediatrics*. 2018;44:34. DOI: 10.1186/s13052-018-0467-z
- [31] Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, Carter NP, et al. Consensus statement: Chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *The American Journal of Human Genetics*. 2010;86:749-764. DOI: 10.1016/j.ajhg.2010.04.006
- [32] Pratte-Santos R, Ribeiro KH, Santos TA, Cintra TS. Analysis of chromosomal abnormalities by CGH-array in patients with dysmorphic and intellectual disability with normal karyotype. *Einstein (Sao Paulo)*. 2016;14:30-34. DOI: 10.1590/S1679-45082016AO3592
- [33] Lee SH, Song WJ. Chromosomal microarray testing in 42 Korean patients with unexplained developmental delay, intellectual disability, autism spectrum disorders, and multiple congenital anomalies. *Genomics and Informatics*. 2017;15:82-86. DOI: 10.5808/GI.2017.15.3.82
- [34] Ooki S. Multiple congenital anomalies after assisted reproductive Technology in Japan (between 2004 and 2009). *ISRN Epidemiology*. 2013:1-9. DOI: 10.5402/2013/452085
- [35] Tao VQ, Chan KY, Chu YW, Mok GT, Tan TY, Yang W, et al. The clinical impact of chromosomal microarray on paediatric care in Hong Kong. *PLoS One*. 2014;9:e109629. DOI: 10.1371/journal.pone.0109629
- [36] D'Amours G, Langlois M, Mathonnet G, Fetni R, Nizard S, Srour M, et al. SNP arrays: Comparing diagnostic yields for four platforms in children with developmental delay. *BMC Medical Genomics*. 2014;7:70. DOI: 10.1186/s12920-014-0070-0
- [37] Chen J, Calhoun VD, Ferrone-Bizzozero NI, Pearson GD, Sui J, Du Y, Liu J. A pilot study on commonality and specificity of copy number variants in schizophrenia and bipolar disorder. *Translational Psychiatry*. 2016;6:e824. DOI: 10.1038/tp.2016.96
- [38] Chen YH, Lu RB, Hung H, Kuo PH. Identifying potential regions of copy number variation for bipolar disorder. *Microarrays (Basel)*. 2014;3:52-71. DOI: 10.3390/microarrays3010052
- [39] Lowther C, Merico D, Costain G, Waserman J, Boyd K, Noor A, et al. Impact of IQ on the diagnostic yield of chromosomal microarray in a community sample of adults with schizophrenia. *Genome Medicine*. 2017;9:105. DOI: 10.1186/s13073-017-0488-z

- [40] Mefford HC. CNVs in epilepsy. *Current Genetic Medicine Reports*. 2014;2:162-167. DOI: 10.1007/s40142-014-0046-6
- [41] Thygesen JH, Wolfe K, McQuillin A, Viñas-Jornet M, Baena N, Brison N, et al. Neurodevelopmental risk copy number variants in adults with intellectual disabilities and comorbid psychiatric disorders. *The British Journal of Psychiatry*. 2018;212:287-294. DOI: 10.1192/bjp.2017.65
- [42] Vlaskamp DRM, Callenbach PMC, Rump F, Giannini LAA, Dijkhuizen T, Brouwer OF, van Ravenswaaij-Arts CMA. Copy number variation in a hospital-based cohort of children with epilepsy. *Epilepsia Open*. 2017;2:244-254. DOI: 10.1002/epi4.12057
- [43] Dauber A, Yu Y, Turchin MC, Chiang CW, Meng YA, Demerath EW, et al. Genome-wide association of copy-number variation reveals an association between short stature and the presence of low-frequency genomic deletions. *American Journal of Human Genetics*. 2011;89:751-759. DOI: 10.1016/j.ajhg.2011.10.014
- [44] Hu G, Fan Y, Wang L, Yao RE, Huang X, Shen Y, Yu Y, Gu X. Copy number variations in 119 Chinese children with idiopathic short stature identified by the custom genome-wide microarray. *Molecular Cytogenetics*. 2016;9:16. DOI: 10.1186/s13039-016-0225-0
- [45] Singh H, Tiwari P, Bhavi V, Chaudhary PS, Suravajhala P, Mohan MK, Mathur SK. Application of chromosomal microarray for evaluation of idiopathic short stature in Asian Indian children: A pilot study. *Indian Journal of Endocrinology and Metabolism*. 2018;22:100-106. DOI: 10.4103/ijem.IJEM_202_17
- [46] Mukherjee S, Sathanoori M, Ma Z, Andreatta M, Lennon PA, Wheeler SR, et al. Addition of chromosomal microarray and next generation sequencing to FISH and classical cytogenetics enhances genomic profiling of myeloid malignancies. *Cancer Genetics*. 2017;216-217:128-141. DOI: 10.1016/j.cancergen.2017.07.010
- [47] Nowak D, Hofmann WK, Koeffler HP. Genome-wide mapping of copy number variations using SNP arrays. *Transfusion Medicine and Hemotherapy*. 2009;36:246-251. DOI: 10.1159/000225372
- [48] de Jong A, Dondorp WJ, Macville MV, de Die-Smulders CE, van Lith JM, de Wert GM. Microarrays as a diagnostic tool in prenatal screening strategies: Ethical reflection. *Human Genetics*. 2014;133:163-172. DOI: 10.1007/s00439-013-1365-5
- [49] Callaway JL, Shaffer LG, Chitty LS, Rosenfeld JA, Crolla JA. The clinical utility of microarray technologies applied to prenatal cytogenetics in the presence of a normal conventional karyotype: A review of the literature. *Prenatal Diagnosis*. 2013;33:1119-1123. DOI: 10.1002/pd.4209
- [50] Levy B, Wapner R. Prenatal diagnosis by chromosomal microarray analysis. *Fertility and Sterility*. 2018;109:201-212. DOI: 10.1016/j.fertnstert.2018.01.005
- [51] Wapner RJ, Martin CL, Levy B, Ballif BC, Eng CM, Zachary JM, et al. Chromosomal microarray versus karyotyping for prenatal diagnosis. *The New England Journal of Medicine*. 2012;367:2175-2184. DOI: 10.1056/NEJMoa1203382
- [52] Armour CM, Dougan SD, Brock JA, Chari R, Chodirker BN, DeBie I, et al. Practice guideline: Joint CCMG-SOGC recommendations for the use of chromosomal microarray

- analysis for prenatal diagnosis and assessment of fetal loss in Canada. *Journal of Medical Genetics*. 2018;55:215-221. DOI: 10.1136/jmedgenet-2017-105013
- [53] Bhat SA, Malik AA, Ahmad SM, Shah RA, Ganai NA, Shafi SS, Shabir N. Advances in genome editing for improved animal breeding: A review. *Veterinary World*. 2017;10:1361-1366. DOI: 10.14202/vetworld.2017.1361-1366
- [54] Gonen S, Jenko J, Gorjanc G, Mileham AJ, Whitelaw CB, Hickey JM. Potential of gene drives with genome editing to increase genetic gain in livestock breeding programs. *Genetics, Selection, Evolution*. 2017;49:3. DOI: 10.1186/s12711-016-0280-3
- [55] Carroll D, Charo RA. The societal opportunities and challenges of genome editing. *Genome Biology*. 2015;16:242. DOI: 10.1186/s13059-015-0812-0
- [56] Blendon RJ, Gorski M, Benson J. The public and the gene-editing revolution. *The New England Journal of Medicine*. 2016;374:1406-1411. DOI: 10.1056/NEJMp1602010
- [57] Murray JD, Maga EA. Genetically engineered livestock for agriculture: A generation after the first transgenic animal research conference. *Transgenic Research*. 2016;25:321-327. DOI: 10.1007/s11248-016-9927-7
- [58] Fleming A, Emhümad AA, Christian M, Christine FB. Invited review: Reproductive and genomic technologies to optimize breeding strategies for genetic progress in dairy cattle. *Archives Animal Breeding*. 2018;61:43-57. DOI: 10.5194/aab-61-43-2018
- [59] Davoudi F, Abdollahi-Arpanahi R, Nejati-Javaremi A. The impact of QTL allele frequency distribution on the accuracy of genomic prediction. *Archives Animal Breeding*. 2018;61:207-213. DOI: DOI.org/10.5194/aab-61-207-2018
- [60] de Los Campos G, Vazquez AI, Fernando R, Klimentidis YC, Sorensen D. Prediction of complex human traits using the genomic best linear unbiased predictor. *PLoS Genetics*. 2013;9:e1003608. DOI: 10.1371/journal.pgen.1003608
- [61] Schaefer RJ, Schubert M, Bailey E, Bannasch DL, Barrey E, Bar-Gal GK, et al. Developing a 670k genotyping array to tag ~2M SNPs across 24 horse breeds. *BMC Genomics*. 2017;18:565. DOI: 10.1186/s12864-017-3943-8
- [62] Júnior GAO, Perez BC, Ferraz JBS. Genomics applied to puberty in beef cattle (*Bos Indicus*). *Revista Brasileira de Reprodução Animal*. 2017;41:264-269
- [63] Gutiérrez-Gil B, Arranz JJ, Pong-Wong R, García-Gómez E, Kijas J, Wiener P. Application of selection mapping to identify genomic regions associated with dairy production in sheep. *PLoS One*. 2014;9:e94623. DOI: 10.1371/journal.pone.0094623
- [64] Kemper KE, Goddard ME. Understanding and predicting complex traits: knowledge from cattle. *Human Molecular Genetics*. 2012;21(R1):R45-51. DOI: 10.1093/hmg/ddz332
- [65] Bruford MW, Ginja C, Hoffmann I, Joost S, Orozco-terWengel P, Alberto FJ, et al. X. Prospects and challenges for the conservation of farm animal genomic resources, 2015-2025. *Frontiers in Genetics*. 2015;21:314. DOI: 10.3389/fgene.2015.00314

- [66] da Cruz, Alex Silva. Estudo de associação ampla do genoma bovino para lactação ajustada em 305 dias em Girolando [thesis]. Goiânia: Universidade Federal de Goiás; 2015. p.106
- [67] Bosch P, Forcato DO, Alustiza FE, Alessio AP, Fili AE, Olmos Nicotra MF, Liaudat AC, Rodriguez N, Talluri TR, Kues WA. Exogenous enzymes upgrade transgenesis and genetic engineering of farm animals. *Cellular and Molecular Life Sciences*. 2015;72:1907-1929. DOI: 10.1007/s00018-015-1842-1
- [68] Mrode R, Tarekegn GM, Mwacharo JM, Dijkeng A. Invited review: Genomic selection for small ruminants in developed countries: How applicable for the rest of the world? *Animal*. 2018;12:1333-1340. DOI: 10.1017/S1751731117003688
- [69] Uemoto Y, Sasaki S, Kojima T, Sugimoto Y, Watanabe T. Impact of QTL minor allele frequency on genomic evaluation using real genotype data and simulated phenotypes in Japanese black cattle. *BMC Genetics*. 2015;16:134. DOI: 10.1186/s12863-015-0287-8
- [70] Hickey JM. Sequencing millions of animals for genomic selection 2.0. *Journal of Animal Breeding and Genetics*. 2013;130:331-332. DOI: 10.1111/jbg.12054
- [71] Auvray B, McEwan J, Newman SA, Lee M, Dodds K. Genomic prediction of breeding values in the New Zealand sheep industry using a 50K SNP chip. *Journal of Animal Science*. 2014;92:4375-4389. DOI: 10.2527/jas.2014-7801
- [72] Teng X, Xiao H. Perspectives of DNA microarray and next-generation DNA sequencing technologies. *Science in China. Series C, Life Sciences*. 2009;52:7-16. DOI: 10.1007/s11427-009-0012-9
- [73] Steyaert W, Callens S, Coucke P, Dermaut B, Hemelsoet D, Terryn W, Poppe B. Future perspectives of genome-scale sequencing. *Acta Clinica Belgica*. 2018;73:7-10. DOI: 10.1080/17843286.2017.1413809
- [74] Aardema MJ, MacGregor JT. Toxicology and genetic toxicology in the new era of "toxicogenomics": Impact of "-omics" technologies. *Mutation Research*. 2002;499:13-25. DOI: 10.1016/S0027-5107(01)00292-5
- [75] Guindalini C, Tufik S. Use of microarrays in the search of gene expression patterns - application to the study of complex phenotypes. *Revista Brasileira de Psiquiatria*. 2007;29:370-374. DOI: 10.1590/S1516-44462007000400014
- [76] Saei AA, Omidi Y. A glance at DNA microarray technology and applications. *BioImpacts*. 2011;1:75-86. DOI: 10.5681/bi.2011.011
- [77] Chavan P, Joshi K, Patwardhan B. DNA microarrays in herbal drug research. *Evidence-based Complementary and Alternative Medicine*. 2006;3:447-457. DOI: 10.1093/ecam/nel075
- [78] Bumgarner R. DNA microarrays: Types, applications and their future. *Current Protocols in Molecular Biology*. 2013;22:1-17. DOI: 10.1002/0471142727.mb2201s101

- [79] Ventola CL. Role of Pharmacogenomic biomarkers in predicting and improving drug response part 1: The clinical significance of pharmacogenetic variants. *P&T*. 2013;38: 545-551. PMID: 24273401
- [80] Wang Z, Du Q, Wang F, Liu Z, Li B, Wang A, Wang Y. Microarray analysis of gene expression on herbal glycoside recipes improving deficient ability of spatial learning memory in ischemic mice. *Journal of Neurochemistry*. 2004;88:1406-1415. DOI: 10.1046/j.1471-4159.2003.02258.x
- [81] Shahandeh A, Johnstone DM, Atkins JR, Sontag JM, Heidari M, Daneshi N, Freeman-Acquah E, Milward EA. Advantages of array-based technologies for pre-emptive pharmacogenomics testing. *Microarrays*. 2016;5:1-11. DOI: 10.3390
- [82] Liljedahl U, Karlsson J, Melhus H, Kurland L, Lindersson M, Kahan T, Nyström F, Lind L, Syvänen AC. A microarray minisequencing system for pharmacogenetic profiling of antihypertensive drug response. *Pharmacogenetics*. 2003;3:7-17. PMID: 12544508
- [83] Anderson DC, Kodukula K. Biomarkers in pharmacology and drug discovery. *Biochemical Pharmacology*. 2014;87:172-188. DOI: 10.1016/j.bcp.2013.08.026
- [84] Meloni R, Khalfallah O, Biguet NF. DNA microarrays and pharmacogenomics. *Pharmacological Research*. 2004;49:303-308. DOI: 10.1016/j.phrs.2003.06.001
- [85] Costa EOA, Pinto IP, Gonçalves MW, da Silva JF, Oliveira LG, da Cruz AS, et al. Small de novo CNVs as biomarkers of parental exposure to low doses of ionizing radiation of caesium-137. *Scientific Reports*. 2018;8:1-13. DOI: 10.1038/s41598-018-23813-5