Bruna Lucheze Freire

**O sequenciamento completo do exoma deve ser
a primeira abordagem genético-molecular para
crianças com baixa estatura sindrômica sem
diagnóstico clínico**

Tese apresentada à Faculdade de Medicina
da Universidade de São Paulo para
obtenção do título de Doutor em Ciências

Programa de Pós-Graduação em
Endocrinologia

Orientador: Prof. Dr. Alexander Augusto de
Lima Jorge

São Paulo

2021

Bruna Lucheze Freire

**O sequenciamento completo do exoma deve ser
a primeira abordagem genético-molecular para
crianças com baixa estatura sindrômica sem
diagnóstico clínico**

Tese apresentada à Faculdade de Medicina
da Universidade de São Paulo para
obtenção do título de Doutor em Ciências

Programa de Pós-Graduação em
Endocrinologia

Orientador: Prof. Dr. Alexander Augusto de
Lima Jorge

São Paulo

2021

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Freire, Bruna Lucheze

O sequenciamento completo do exoma deve ser a primeira abordagem genético-molecular para crianças com baixa estatura sindrômica sem diagnóstico clínico / Bruna Lucheze Freire. -- São Paulo, 2021.

Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.

Programa de Endocrinologia.

Orientador: Alexander Augusto de Lima Jorge.

Descritores: 1.Retardo do crescimento fetal/genética 2.Insuficiência de crescimento/genética 3.Sequenciamento de nucleotídeos em larga escala 4.Sequenciamento completo do exoma 5.Mutação/genética 6.Genética

USP/FM/DBD-335/21

Responsável: Erinalva da Conceição Batista, CRB-8 6755

AGRADECIMENTOS

Agradeço primeiramente, aos meus familiares, principalmente meus pais Júlio César e Cristiane por todo apoio que recebo em todas as minhas decisões de vida pessoal e profissional, me oferecendo todo o suporte e tranquilidade necessários para eu alcançar meu objetivo.

Aqui, não podia deixar de exaltar meu companheiro César. É com ele que divido todas as minhas ansiedades e alegrias diariamente. Me apoia em todos os sentidos e me acalma nos piores momentos, principalmente nas pequenas crises que surgiram devido a pandemia.

Em especial, agradeço muito o meu orientador Alexander pela oportunidade oferecida e por todo suporte dos últimos 7 anos que tornaram essa parceria bem-sucedida. A relação orientador-orientando acaba aqui, mas tenho muito orgulho de ter feito parte deste time. E espero levar a amizade desta pessoa incrível que você é para a vida.

Agradeço todos os colegas de trabalho que fizeram parte de alguma forma deste projeto. Especialmente a Mariana pela ajuda nos experimentos, e a Thais por toda ajuda com a caracterização clínica dos pacientes. Agradeço também a Dra. Miriam e Dra. Luciana por todo o ensinamento me passado e dúvidas solucionadas. Agradeço ao SELA e seus funcionários pelo apoio no sequenciamento. Assim como todo pessoal do LIM 42 e LIM 25, que foram os

laboratórios envolvidos no projeto. E ao Dr. Antônio Lerário, responsável por toda a parte de bioinformática do projeto.

Sou grata às agências de fomento, Fapesp e Capes, pelo apoio financeiro oferecido para que o projeto fosse desenvolvido.

Um obrigado imenso aos meus amigos por fazerem parte da família que eu escolhi para vida, estamos juntos desde o começo da faculdade, e mesmo seguindo outros caminhos pelo mundo, estamos sempre presentes na vida um do outro. Obrigada, Barbe, Chicken, Messi, Chip, Paulista, Mococa, Buh, Ingrid, Rebelde, Nay, Mah, Bagas, Pushpop, pela amizade e apoio incondicional, quero vocês para sempre presentes.

Por fim, agradeço os amigos veganos queridos Anaju, Fe e Ju pelas conversas, jantas, compartilhamento de ideais e conversas profundas. Obrigada por serem tão especiais comigo e com os animais.

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de International Committee of Medical Journals Editors (Vancouver).

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. Guia de apresentação de dissertações, teses e monografias. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com List of Journals Indexed in Index Medicus. 3.1.11 Sumário

SUMÁRIO

Lista de abreviaturas	
Lista de siglas	
Lista de símbolos	
Lista de figuras	
Lista de tabelas	
Lista de anexos	
Resumo	
Abstract	

1 INTRODUÇÃO	01
1.1 Baixa Estatura Sindrômica	01
1.2 Investigação genético-molecular dos distúrbios do crescimento	01
1.3 Sequenciamento completo do exoma (WES)	04
1.4 Genética da Baixa Estatura	06
1.4.1 Defeitos nos genes do eixo <i>Gh-IGFs</i>	08
1.4.2 Defeitos em componentes da matriz extracelular da cartilagem de crescimento	09
1.4.3 Defeitos em fatores parácrinos que regulam a cartilagem de crescimento	09
1.4.4 Defeitos em vias de sinalização intracelular	10
1.4.5 Defeitos em mecanismos celulares fundamentais	11
1.4.6 Síndrome dos genes contíguos	12
1.4.7 Defeitos de <i>imprinting</i>	12
1.5 Diagnóstico Clínico	12
2 OBJETIVOS	15
2.1 Objetivo Principal	15
2.2 Objetivos Específicos	15
3 MATERIAL E MÉTODOS	16
3.1 Considerações éticas	16
3.2 Casuística	16
3.3 Estudos Moleculares	17
3.3.1. Extração do DNA genômico	17
3.3.2 Genotipagem por sequenciamento paralelo em larga escala	18

3.3.3 Cultura de fibroblastos	20
3.3.4 Confeção de cDNA e amplificação por PCR	20
3.3.5 MLPA	21
3.3.6 aCGH	22
4 ANÁLISE DOS RESULTADOS	23
4.1 Bioinformática	23
4.2 Análise de Variantes Número de Cópias (CNVs)	28
4.3 Classificação das variantes de interesse	29
5 RESULTADOS	31
5.1 Caracterização da casuística	31
5.2 Resultados genético-moleculares	34
5.3 Investigação de novos genes candidatos	42
6 DISCUSSÃO	49
6.1 Raridade e Sobreposição fenotípica	51
6.2 Estabelecimento de nova relação Genótipo-Fenótipo	60
6.3 Limitações do WES e análise de CNVs	62
6.3.1 Duplicação chr11 (p15.5-p15.4):2290980-3845365x3	64
6.3.2 Duplicação chr5 (q35.2-q35.3):175394113-177422934x3	65
6.3.3 Duplicação chr8 (q24.22-q24.3):133108784-146294098)x3 e Deleção chr15 (q26.2-q26.3):97439378-102383473)x1	66
7 CONCLUSÃO	70
8 ANEXOS	71
Anexo A - Protocolo para avaliação inicial de paciente com baixa estatura	71
Anexo B - Critérios de avaliação utilizados para calcular Escore-Z de acordo com as características informadas	73
Anexo C - Resumo das linhas de evidência utilizadas na classificação pelo guia ACMG-AMP	74
Anexo D - Critérios utilizados para classificação final de patogenici- dade segundo o guia ACMG-AMP	75
Anexo E - Heredogramas de todas as famílias onde foram encontra- das variantes Patogênicas ou Provavelmente Patogênicas	76
9 REFERÊNCIAS	79

LISTA DE ABREVIATURAS

aCGH	CGH <i>arrays</i>
ACMG-AMP	<i>American of Medical Genetics and Genomics and Association for Molecular Pathology</i>
ADNPM	Atraso do desenvolvimento neuro psicomotor
B	<i>Benign</i> (Benigna)
pb	<i>base pair</i>
cDNA	DNA complementar
CGH	<i>Comparative genomic hybridization</i> (Hibridação genômica comparativa)
CN	Comprimento ao nascimento
CNVs	Copy number variantions (Variações no número de cópias)
ddNPPs	Didesoxirribonucleotídeos
DNA	<i>Deoxyribonucleic Acid</i>
DNPM	Desenvolvimento neuro psicomotor
DO	Densidade ótica
<i>FGFs</i>	Fatores de crescimento dos fibroblastos
FISH	Hibridação <i>in situ</i> fluorescente
Fw	<i>Foward</i>
GDP	Guanosina difosfato
GEF	Fator de troca de nucleotídeos de guanina
GH	Hormônio do crescimento
GHD	Deficiência de GH
GTP	Guanosina trifosfato
Indels	Inserção e deleção
LoF	<i>Variante Loss of Function</i> (Perda de função)
M	Moderada
MAF	<i>Minor Allele Frequency</i> (Frequência do alelo menos comum)
MLPA	<i>Multiplex Ligation-dependent Probe Amplification</i> (Amplificação múltipla dependente da ligação)

n	Número
OMIM	<i>Online Mendelian Inheritance in Man</i>
P	<i>Pathogenic</i> (Patogênica)
PB	<i>Likely Benign</i> (Provavelmente benigna)
PCR	<i>Polimerase chain reaction</i> (Reação em cadeia da polimerase)
PIG	Pequeno para Idade Gestacional
pLI	Probabilidade de intolerância a variantes perda de função
PN	Peso ao Nascimento
PP	<i>Likely Pathogenic</i> (Provavelmente Patogênica)
<i>PTHrP-IHH</i>	<i>Peptídeo relacionado a hormônios da paratireoide – Indian hedgehog</i>
RGHC	Registro Geral no HC
rhGH	Hormônio recombinante de crescimento humano
RT-PCR	<i>Reverse Transcriptase PCR</i>
Rv	<i>Reverse</i>
S	Forte
SFH	Síndrome de Floating Harbour
SNVs	Variante em nucleotídeo único
SPLE	Sequenciamento paralelo em larga escala
<i>TGFβ</i>	fator de crescimento transformador Beta
UTR	<i>Untranslated region</i>
VCF	<i>Variant Call Format</i>
VS	Muito forte
VUS	<i>Variant of Unknown Significance</i> (Variante de significado incerto)
WES	<i>Whole Exome Sequencing</i> (Sequenciamento completo do exoma)
Z	Escore-Z

LISTA DE SÍMBOLOS

cm	centímetro
g	grama
kg	quilograma
mL	mililitro
mm	milímetro
nm	nanometro
nM	nanomolar
°C	grau Celsius
µg	micrograma
µL	microlitro
x	vezes
%	por cento
<	menor
>	maior

LISTA DE SIGLAS

CAPPesq	Comissão de Ética para Análise de Projetos de Pesquisa
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo
FMUSP	Faculdade de Medicina da Universidade de São Paulo
HCFMUSP	Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo

LISTA DE FIGURAS

Figura 1	Esquema das principais etapas do sequenciamento paralelo em larga escala. 1. Primeiramente ocorre a fragmentação do material genômico e a ligação a adaptadores. 2. Hibridação com sondas de biotina e captura dos exons com beads de estreptavidina. 3. Ligação dos adaptadores a pequenas sequencias que ficam em uma placa denominada flowcell, onde ocorre a amplificação clonal. 4. Ocorre o sequenciamento propriamente dito, que no caso é feito por síntese através de uma DNA polimerase, cada nucleotídeo adicionado emite um fluoróforo específico.....	19
Figura 2	Fluxo da estratégia de filtragem utilizada para selecionar as variantes com possibilidade de patogenicidade em 86 pacientes com baixa estatura síndrômica submetidos ao sequenciamento do exoma	25
Figura 3	Exemplo da visualização de uma variante não sinônima T>C (acima) e uma deleção de único nucleotídeo G (abaixo) no programa IGV (Integrative Genomics Viewer).	27
Figura 4	Exemplo da visualização de uma Duplicação (acima em azul) e uma Deleção (abaixo em vermelho), indicados por seta no programa IGV (Integrative Genomics Viewer)	29
Figura 5	Fluxograma dos pacientes com baixa estatura síndrômica do ambulatório do HCFMUSP. Dos 818 pacientes atendidos, 166 são síndrômicos sem diagnóstico clínico. Destes, 57 tiveram aCGH prévio com resultado negativo e foram encaminhados para WES. No total 86 realizaram exoma e 41 pacientes tiveram resultado positivo	33
Figura 6	Estado das 47 variantes patogênicas ou Provavelmente Patogênicas encontradas em 41 pacientes	40

Figura 7	<p>A: Transcrito <u>ENST00000550536.5</u> do gene <i>RAB3IP</i> utilizado para desenho dos primers. Setas Azuis, correspondem as regiões dos exons 1 e 2. Setas laranjas, as regiões dos exons 3 e 4. E setas vermelhas, as regiões dos exons 8 e 9 B: Foto do gel de eletroforese da amplificação do cDNA do paciente com variante no gene <i>RAB3IP</i>. A região correspondente aos exons 1 e 2 não foram amplificadas. Foram amplificadas as regiões correspondentes aos exons 3 e 4 e 8 e 9 em paciente e controles. Ladder: 1 kb da Thermo Fisher Scientific. C-:Controle negativo (DNA genômico);C1:Controle 1;C2:Controle 2 (cDNA de pacientes normais) 45</p>
Figura 8	<p>a: Transcrito ENST00000257904.11 do gene <i>CDK4</i> utilizado para desenho dos 2 pares de primers (Tabela 2) que cobrem o cDNA completo. A seta vermelha indica a variante c.218G>A. b: Foto do gel de eletroforese da amplificação do cDNA do paciente com variante no <i>CDK4</i>. A região correspondente ao par de primer 1 é menor no paciente (329pb) quando comparada ao controle (566pb). Ladder: 1 kb da Thermo Fisher Scientific. c: Eletroferograma do sequenciamento do cDNA do paciente e de um controle, mostrando o ponto novo de <i>splicing</i> 47</p>
Figura 9	<p>a: Visualização das duplicações chr11 (p15.5-p15.4): 2290980-3845365x3. b: chr5 (q35.2-q35.3): 175394113-177422934x3. c: chr8 (q24.22-q24.3): 133108784-146294098x3 e d: chr15 (q26.2-q26.3): 97439378-102383473x1 no IGV, indicadas por seta vermelha 63</p>
Figura 10	<p>MLPA ME030 para BWS/SRS realizado nos pais da paciente. a: MLPA da mãe - acima podemos ver a duplicação da região correspondente aos genes <i>KCNQ1</i> e <i>CDKN1C</i>. E abaixo podemos ver a hipometilação da região do cromossomo 11. b: MLPA do pai, normal 65</p>

Figura 11	Confirmações das 4 CNVs encontradas em 3 pacientes por aCGH. a: Duplicação chr11 (p15.5-p15.4): 2290980-3845365x3. b: Duplicação chr5 (q35.2-q35.3): 175394113-177422934x3. c: chr8 (q24.22-q24.3): 133108784-146294098)x3. d: Deleção chr15 (q26.2-q26.3): 97439378-102383473)x1	67
Figura 12	Sugestão de abordagem molecular para pacientes com baixa estatura sindrômica	69

LISTA DE TABELAS

Tabela 1	Principais grupos de defeitos genéticos que ocasionam quadros sindrômicos associados à baixa estatura	07
Tabela 2	Primers utilizados para amplificação do cDNA do paciente com variante no gene RAB3IP e CDK4. F:Foward; R:Reverse	21
Tabela 3	Característica da casuística de 86 pacientes sindrômicos sem diagnóstico clínico estudadas por WES	32
Tabela 4	Métricas médias das corridas de sequenciamento do exoma realizadas durante o projeto	34
Tabela 5	Variante patogênicas ou provavelmente patogênicas encontradas em 41 de 86 pacientes da casuística	35
Tabela 6	Variante em genes de interesse submetidas no GeneMatcher	43
Tabela 7	Características clínicas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas e síndromes associadas	52

LISTA DE ANEXOS

Anexo A	Protocolo para avaliação inicial de paciente com baixa estatura	72
Anexo B	Critérios de avaliação utilizados para calcular Escore-Z de acordo com as características informadas	73
Anexo C	Resumo das linhas de evidência utilizadas na classificação pelo guia ACMG-AMP	74
Anexo D	Critérios utilizados para classificação final de patogenicidade segundo o guia ACMG-AMP	75
Anexo E	Heredogramas de todas as famílias onde foram encontradas variantes Patogênicas ou Provavelmente Patogênicas	76

RESUMO

Freire BL. *O sequenciamento completo do exoma deve ser a primeira abordagem genético-molecular para crianças com baixa estatura sindrômica sem diagnóstico clínico*. [Tese]. São Paulo. Faculdade de Medicina, Universidade de São Paulo; 2021.

As doenças que comprometem o crescimento humano apresentam uma forte influência genética. O objetivo geral do projeto foi avaliar a utilidade do sequenciamento completo do exoma como primeira abordagem genético-molecular na investigação da etiologia da baixa estatura sindrômica de pacientes sem diagnóstico clínico reconhecido. Para tanto os pacientes com características sindrômicas foram submetidos a extensa avaliação clínica em nosso ambulatório e nos casos selecionados, foi utilizado o sequenciamento completo do exoma para estabelecer etiologia genética responsável pelo distúrbio de crescimento. Desta forma foram selecionados de forma prospectiva 86 pacientes com baixa estatura (escore-Z de altura menor ou igual a -2) ou baixa estatura em relação ao seu padrão familiar (escore-Z da altura do paciente menos escore-Z da altura alvo $\leq$ -2) associado a dismorfismos, malformações congênitas maiores, distúrbios de desenvolvimento neuropsicomotor ou déficit intelectual. Os pacientes selecionados não apresentam diagnóstico clínico ou a suspeita clínica inicial foi afastada por estudo molecular prévio baseado em gene candidato. Foram selecionados 86 pacientes que foram avaliados em nosso ambulatório. Estes foram submetidos ao sequenciamento do exoma com captura das regiões alvo pela tecnologia SureSelect (Agilent) ou xGen Exome Research Panel v2 (IDT). Em 61 casos apenas o probando foi inicialmente analisado em 25 houve a análise conjunta com outros familiares (trios). A análise das amostras gerou uma cobertura média de 152 vezes e com mais de 98,15% da região alvo com cobertura > 10 reads. Nestas amostras foram identificadas 47 variantes deletérias em 41 pacientes. Em 22 pacientes as variantes foram consideradas patogênicas e em 19 provavelmente patogênicas. _Dentre as variantes consideradas patogênicas ou provavelmente patogênicas houve uma grande heterogeneidade entre os genes, sendo identificadas variantes nos genes *ACTB*, *AFF4*, *ANKRD11*, *BCL11B*(x2), *BRCA1*, *CHD7*, *COL2A1*(x3), *DPH1*, *DYRK1A*, *FBN1*, *FOXP1*, *GINS1*, *HDAC8*, *IGF1R*, *INPP5K*, *KIF11*, *KDM5C*, *KMT2A*, *KIF11*, *KRAS*, *MVK*, *PCNT*, *PDHA1*, *POC1A*, *POLG1*, *PUF60*, *RPL5*, *RPS6KA3*, *SCN1A*, *SCUBE3*, *SETD5*, *SMARCA4*, *SRCAP*(x2), *STAT5B*, *TERT*, além de 3 duplicações e 1 deleção encontradas pela análise de CNVs. Podemos concluir que a técnica de sequenciamento paralelo em larga escala de exoma foi eficiente em estabelecer o diagnóstico molecular da baixa estatura sindrômica, e foi possível identificar a etiologia genética em 47,7% da casuística estudada.

Descritores: Retardo do crescimento/ genética; Insuficiência de crescimento/ genética; Sequenciamento de nucleotídeos em larga escala; Sequenciamento completo do exoma; Mutação/genética; Genética.

ABSTRACT

Freire BL. *Whole exome sequencing should be the first genetic-molecular approach for children with syndromic short stature without clinical diagnosis*. [Thesis]. São Paulo. "Faculdade de Medicina, Universidade de São Paulo"; 2021.

Diseases that compromise human growth have a strong genetic influence. The main aim of the current project was to evaluate the usefulness of complete exome sequencing as the first molecular-genetic approach in the investigation of the etiology of syndromic short stature in patients without a recognized clinical diagnosis. Therefore, patients with syndromic characteristics underwent extensive clinical evaluation in our clinic and, in selected cases, complete exome sequencing was used to establish the genetic etiology responsible for the growth disorder. Thus, 86 patients with short stature (Z-score for height less than or equal to -2) or short stature in relation to their family pattern (Z-score of patient height minus Z-score of target height ≤ -2) associated with dysmorphisms, major congenital malformations, neuropsychomotor development disorders or intellectual deficit were prospectively selected for WES. The selected patients do not have a clinical diagnosis or the initial clinical suspicion was ruled out by a previous molecular study based on a candidate gene. These were subjected to exome sequencing with capture of target regions by SureSelect (Agilent) or xGen Exome Research Panel v2 (IDT) technology. In 61 cases only the proband was initially analyzed and in 25 there was a joint analysis with other family members (trios). The analysis of the samples generated an average coverage of 152 times with more than 98.15% of the target region with coverage > 10 reads. In these samples, 47 deleterious variants were identified in 41 patients. In 22 patients the variants were considered pathogenic and in 19 probably pathogenic. Among the variants considered pathogenic or probably pathogenic, there was great heterogeneity between the genes, with variants being identified in *ACTB*, *AFF4*, *ANKRD11*, *BCL11B*(x2), *BRCA1*, *CHD7*, *COL2A1*(x3), *DPH1*, *DYRK1A*, *FBN1*, *FOXP1*, *GIN51* genes. , *HDAC8*, *IGF1R*, *INPP5K*, *KIF11*, *KDM5C*, *KMT2A*, *KIF11*, *KRAS*, *MVK*, *PCNT*, *PDHA1*, *POC1A*, *POLG1*, *PUF60*, *RPL5*, *RPS6KA3*, *SCN1A*, *SCUBE3*, *SETD5*, *SMARCA4*, *SRCAP*(x2), *STAT5B*, and 3 duplications and one deletion found by CNV analysis. We can conclude that the massive parallel sequencing technique of the exome was efficient in establishing the molecular diagnosis of syndromic short stature, and it was possible to identify the genetic etiology in 47,7% of the studied cohort.

Descriptors: Fetal growth retardation/genetics; Failure to thrive /genetics; High-throughput nucleotide sequencing; Whole exome sequencing; Mutation/genetics; Genetics.

1 INTRODUÇÃO

1.1 Baixa Estatura Sindrômica

Em medicina o termo síndrome (palavra originada do grego antigo συνδρομή, συν+δρόμος significando "con+correr") é utilizado para definir um conjunto de sinais e sintomas que se encontram associados a uma condição com etiologia conhecida ou não¹. A baixa estatura é uma das principais características fenotípicas de muitas síndromes dismórficas conhecidas¹ e é definida como altura menor ou igual a -2 desvios padrões em relação à média para idade, sexo e grupo populacional ²⁻⁴.

Dentre os pacientes com baixa estatura, uma parcela significativa apresenta outras características associadas, como dismorfismos faciais, malformações maiores (agenesia renal, defeito de linha média, defeitos cardíacos graves por exemplo), atraso do desenvolvimento neuropsicomotor (ADNPM), déficit intelectual, desproporcionalidade ou deformidades esqueléticas. Estes pacientes são classificados como tendo baixa estatura sindrômica e em muitos a causa etiológica reside em um defeito genético^{1,5}, sendo mais provável causas monogênicas^{5,6}.

1.2 Investigação genético-molecular dos distúrbios do crescimento

A identificação de defeitos moleculares associados a doenças genéticas sempre foi um grande desafio. Nas décadas de 50 e 60, a descrição do número

de cromossomos e do padrão de bandeamento dos mesmos⁷, respectivamente, possibilitou as primeiras associações entre alterações cromossômicas e algumas síndromes. Ainda em 1959, a síndrome de Down foi relacionada à trissomia do cromossomo 21⁸ e alterações no número de cromossomos sexuais foram associadas às síndromes de Turner e de Klinefelter^{9,10}. A partir de então, inserções, deleções, translocações e outros rearranjos complexos começaram a ser descritos e também foram relacionados a quadro clínicos sindrômicos.

Nas últimas décadas, observamos o desenvolvimento de novas técnicas utilizadas na pesquisa de variação no número de cópias de material genômico¹¹. Em nosso grupo, diversos estudos foram desenvolvidos utilizando as técnicas de hibridação *in situ* fluorescente (FISH) e de amplificação múltipla dependente da ligação de sondas (Multiplex Ligation-dependent Probe Amplification – MLPA) na investigação de doenças endocrinológicas que comprometem o crescimento¹²⁻¹⁴.

Na década de 90, foi descrita a técnica de hibridação genômica comparativa (*comparative genomic hybridization* - CGH)¹⁵, que possibilitou a detecção de ganhos e perdas de regiões cromossômicas. O aperfeiçoamento desta técnica foi obtido com o desenvolvimento do CGH *arrays* (aCGH), uma metodologia de alta resolução capaz de analisar todo o genoma a procura de variações no número de cópias (CNVs)¹⁶. Em nosso grupo, diversos trabalhos foram realizados utilizando esta técnica, onde foram avaliados pacientes com retardo de crescimento pré- e pós-natal associados a características dismórficas e/ou atraso no desenvolvimento, mas sem critérios para o diagnóstico de síndromes conhecidas^{16,17}. Um dos estudos avaliou 51 pacientes nascidos

pequenos para idade gestacional (PIG) e foram identificadas 9 CNVs patogênicas ou provavelmente patogênicas em 8 destes pacientes, correspondendo então a 16% de diagnóstico molecular na casuística estudada¹⁶. No outro estudo foram analisados 229 pacientes com distúrbio de crescimento e foram encontradas 32 CNVs patogênicas (n=25) ou provavelmente patogênicas (n=7), o que corresponde a 14% da casuística estudada¹⁷. O que mostra que as CNVs parecem apresentar um papel significativo em pacientes com baixa estatura sindrômica. Porém, a maioria destes pacientes ainda não apresentam aneuploidias ou CNVs, sendo importante avaliar alterações de ponto.

Na década de 70, foi descrito o sequenciamento genômico¹⁸ que permitiu analisar a sequência de nucleotídeos dos genes de interesse. A técnica de sequenciamento classicamente utilizada nas últimas quatro décadas é conhecida como técnica de Sanger¹⁸ e baseia-se no uso de sequências de nucleotídeos iniciadores, complementares a sequência alvo, iniciando o processo de replicação do DNA por ação de uma DNA polimerase. A presença de nucleotídeos modificados, didesoxirribonucleotídeos (ddNPPs), na reação provoca a parada aleatória da extensão da fita de DNA em pontos distintos. Esta técnica permite o sequenciamento de 400-800 pares de bases (pb) de uma determinada região alvo em cada reação. O Projeto Genoma Humano utilizou essa metodologia, sendo concluído em abril de 2003, 13 anos após o seu início¹⁹ envolveu mais de 500 pesquisadores e teve um custo estimado de US\$ 3,8 bilhões. Nestas últimas duas décadas a popularização da técnica de Sanger permitiu a identificação de diversos genes, que quando mutados eram responsáveis por doenças monogênicas. Na área de crescimento, nosso grupo

incorporou exames moleculares como uma extensão natural da investigação laboratorial comum. Nos últimos 20 anos foram mais de 22 genes estudados, muitos incorporados na rotina da investigação diagnóstica dos pacientes, abrangendo genes ligados a defeitos no eixo GH/IGF-1^{20,21}, displasias esqueléticas^{22,23} ou associados a síndromes genéticas^{24,25}.

Apesar do método de Sanger ter permitido a introdução do sequenciamento gênico como ferramenta para o diagnóstico em muitas doenças, a falta de regiões *hot spot*, a heterogeneidade genética e a necessidade de sequenciar extensas regiões em várias condições de origem genética dificultam a implantação desta metodologia na rotina diagnóstica. Por este motivo, recentemente novas metodologias de sequenciamento foram desenvolvidas visando a genotipagem simultânea de diversas regiões do genoma, com maior velocidade e menor custo. Desta forma, passou a ser possível que toda a região genômica de um determinado gene e até mesmo vários genes fossem analisados em uma única reação expandindo as possibilidades de uso das técnicas de sequenciamento para investigação de doenças genéticas e na aplicação desta ferramenta na rotina diagnóstica de diversas doenças conhecidas.

1.3 Sequenciamento Completo do Exoma (WES)

As novas tecnologias de sequenciamento enquadram-se no chamado sequenciamento paralelo em larga escala (SPLE), as quais combinam uma etapa que visa o preparo de bibliotecas de sequências de DNA a serem

analisados, o sequenciamento propriamente dito e uma análise de bioinformática para o alinhamento das sequências geradas a um genoma referência e priorização das variantes encontradas. Estes instrumentos permitem uma rápida preparação de amostras antes do sequenciamento, e gera uma grande quantidade de dados que são em parte processados de forma automatizada em ferramentas de bioinformática com significativa redução do custo por base sequenciada e no tempo de análise^{26,27}.

Uma das abordagens mais utilizadas nos últimos anos de SPLE para investigação de pacientes com suspeita de doença genética monogênica, é aquela que permite o sequenciamento completo de toda a parte codificadora de proteínas dos genes, aproximadamente 1-2% do genoma, conhecida como exons²⁸⁻³⁰. Essa técnica é conhecida como sequenciamento completo do exoma (WES - *Whole Genome Sequencing*). O WES permite a detecção de variantes genéticas raras e novas mutações que contribuem para a doença analisando os exons ou regiões codificantes de milhares de genes simultaneamente. Sequenciando o exoma de um paciente e comparando-o com uma sequência de referência, é possível identificar variações na sequência de DNA que pode ser analisadas com base no fenótipo do indivíduo^{29,30}. Yang, et al, em 2013, mostrou pela o potencial diagnóstico do WES em um grupo de 2000 pacientes com doenças de suspeita genética, porém sem diagnóstico clínico, com taxa diagnóstica de 25,2%²⁹. Nos últimos anos, o WES tornou-se cada vez mais disponível e utilizado para diagnóstico genético, o que levou a um grande progresso em relação a detecção de variantes genéticas e diagnóstico molecular

humano, com resultados identificados em até 46% da casuística estudada em diversos trabalhos³⁰⁻³⁵.

Em 2016, Stark et. al. comparou o uso desta técnica com o uso dos métodos tradicionais de investigação genética (Sequenciamento tradicional por Sanger, sequenciamento de painéis de genes candidatos ou mitocondriais e/ou estudos de metilação) em 80 pacientes com suspeita de doenças monogênicas, e com o uso do exoma foi possível estabelecer o diagnóstico genético-molecular em 46 pacientes, comparados a apenas 11 pacientes diagnosticados pelos métodos tradicionais. Mostrando que o uso desta tecnologia para investigação molecular de doenças raras é promissora e deve ser utilizada como primeira abordagem genético-molecular para suspeitas de doenças monogênicas raras, sem suspeita clínica³⁵.

1.4 Genética da Baixa Estatura

A variação normal na altura adulta é em grande parte devido a fatores genéticos herdados. Normalmente 80% ou mais da variação na altura é explicada por estes fatores, embora exista a influência de fatores ambientais para diferenças de altura entre indivíduos e populações. A contribuição genética para a altura é atribuída aos efeitos combinados de diversos genes, sendo a altura dentro da faixa de referência tipicamente uma característica poligênica^{6,36-38}.

Recentemente, estudos de associação em todo o genoma identificaram centenas de variantes genéticas comuns na população (frequência na população

>5%) que tem pequenos efeitos na altura. Esses efeitos somados explicam cerca de 24,6% da variação na altura adulta, e estima-se que variantes comuns adicionais ainda não identificadas representem outros aproximadamente 40% ou mais da variação da altura³⁹.

No caso da baixa estatura sindrômica, os pacientes têm em sua maioria, defeitos em apenas 1 gene, que resulta em efeitos importantes na altura e na determinação dos demais fenótipos. A associação de diversos defeitos moleculares que causam baixa estatura sindrômica pode ser didaticamente classificada conforme o mecanismo molecular pelo qual resulta em um fenótipo pleiotrópico^{5,37} (Tabela 1).

Tabela 1 - Principais grupos de defeitos genéticos que ocasionam quadros sindrômicos associados à baixa estatura

Grupos de defeitos genéticos	Exemplos
Eixo <i>GH-IGFs</i>	Deficiência de GH, síndrome de Laron
Componentes da Matriz extracelular da cartilagem de crescimento	Displasias esqueléticas (displasia espondiloepifisária)
Fatores parácrinos que regulam a cartilagem de crescimento	Displasias esqueléticas (Acondroplasia)
Vias de sinalização intracelular	Síndrome de Noonan
Mecanismos celulares fundamentais	Síndrome de Bloom, Anemia de Fanconi
Genes contíguos	Síndrome de DiGeorge
<i>Imprinting</i>	Síndrome de Silver-Russel

1.4.1 Defeitos nos genes do eixo *GH-IGFs*

O hormônio do crescimento (GH) desempenha um papel crítico no crescimento humano, principalmente através da regulação da produção de IGF-1. Distúrbios genéticos foram identificados em vários componentes do eixo *GH/IGFs*⁴⁰. A primeira doença associada à baixa estatura grave, foi a deficiência de GH causada por uma deleção em homozigose no gene *GH1* (MIM*139250)⁴¹. Subsequentemente, defeitos no gene do receptor de GH (*GHR*, MIM*600946) também foram descritos como causadores da Síndrome de Laron (MIM#262500), uma forma completa de Insensibilidade ao GH, além de formas parciais desta resistência hormonal (MIM#604271)^{38,42}.

Tem sido proposto também, que defeitos neste sistema envolvendo o *IGF1* (MIM*147440) e seu receptor *IGF1R* (MIM*147370), sejam um mecanismo de baixa estatura de início pré-natal e de persistência de déficit de crescimento pós-natal. Pacientes com defeitos no gene do *IGF1* foram previamente descritos em raros relatos de caso na literatura⁴³. Até o momento, poucas mutações inativadoras no gene *IGF2* (MIM*147470), pertencente à mesma via, foram descritas como causa de retardo de crescimento pré e pós-natal, causando um quadro de síndrome de Silver-Russel de herança autossômica dominante paterna^{44,45}.

Em contraste a raridade dos defeitos nas *IGFs*, mutações inativadoras e deleções do *IGF1R* em heterozigose vêm sendo descritas de forma crescente nos últimos anos em pacientes com de déficit de crescimento pré e pós-natal⁴⁶. É estimado que mutações no gene *IGF1R* sejam encontradas em aproximadamente 2% das crianças nascidas pequenas para idade gestacional⁴⁷.

Em nosso grupo, o estudo que avaliou a presença de variações no número de cópias por CGHa em um grupo de 671 crianças nascidas baixas observou em 8 pacientes CNVs envolvendo o gene *IGF1R*, além de alguns CNVs recorrentes à condição estudada, como o rearranjo 15q26, que compreende o gene *IGF1R*¹⁷.

1.4.2 Defeitos em componentes da matriz extracelular da cartilagem de crescimento

Os condrócitos secretam uma matriz extracelular que contém colágenos, proteoglicanos e outras proteínas essenciais para as propriedades da cartilagem, que são compressão, resiliência estrutural e interação com moléculas que participam na cartilagem de crescimento. O gene *ACAN* (MIM*155760) é um exemplo, cuja proteína codificada *Agrecan* é um dos componentes em maior quantidade na matriz extracelular e mutações neste gene podem ser responsáveis por um espectro fenotípico variável, desde baixa estatura proporcional com avanço da idade óssea quando em heterozigose até baixa estatura com displasias esqueléticas quando em homozigose⁴⁸.

1.4.3 Defeitos em fatores parácrinos que regulam a cartilagem de crescimento

Neste grupo são incluídos defeitos moleculares em genes participantes de vias que agem na cartilagem de crescimento, como a via de sinalização dos fatores de crescimento do fibroblasto (*FGFs*) e seus receptores, sendo o exemplo mais conhecido é o gene *FGFR3* (MIM*134934), que quando mutado pode casuar a acondroplasia (MIM#100800)⁴⁹.

Dentro desta categoria, ainda se encontram as vias de sinalização que determinam a morfologia, proliferação e diferenciação dos condrócitos da cartilagem de crescimento, que pertencem ao grupo fator de crescimento transformador Beta ($TGF\beta$), os fatores que participam da mobilidade e polaridade celular, como *ROR2* (MIM*602337), a via *PTHrP-IHH* (*Peptídeo relacionado a hormônios da paratireoide – Indian hedgehog*), a qual age como um feedback negativo para hipertrofia e proliferação dos condrócitos, e também a via *CNP* (MIM*123830) e seu receptor *NPR2* (MIM*108961), que apresenta papel fundamental na regulação do crescimento ósseo endocondral, promovendo a síntese da matriz cartilaginosa e estimulação da proliferação e diferenciação dos condrócitos^{38,50}.

1.4.4 Defeitos em vias de sinalização intracelular

Diversas vias de sinalização intracelular exibem um papel importante para o ciclo celular, incluindo a proliferação e diferenciação dos condrócitos da cartilagem de crescimento. Quando desreguladas, essas vias levam a profundas consequências no desenvolvimento⁵¹. Como exemplo, temos a via RAS/MAPK, a qual integra sinais de alguns fatores de crescimento, como GH, FGFs e EGF (Fator de crescimento epidérmico).

A hiperativação de alguns genes que participam desta via, como *PTPN11* (MIM*176876), *KRAS* (MIM*190070), *SOS1* (MIM*182530), *RAF1* (MIM*164760), *BRAF* (MIM*164757), *SHOC2* (MIM*602775), *NRAS* (MIM*164790), *MEK1* (MIM*176872), *CBL* (MIM*165360) e *LZTR1* (MIM*600574) levam a condições denominadas RASopatias, que incluem as

Síndromes de Noonan (MIM#163950), Síndrome de Leopard (MIM#151100), dentre outras³⁸. As RASopatias são caracterizadas por uma herança autossômica dominante na sua grande maioria, com exceção para o *LZTR1* (autossômica recessiva) e que tem como principais características fenotípicas a presença de anormalidades congênitas múltiplas, como alterações faciais típicas, baixa estatura, defeito cardíaco congênito, atrasos de DNPM e graus variáveis de deficiência intelectual⁵².

1.4.5 Defeitos em mecanismos celulares fundamentais

Processos celulares fundamentais envolvem reações que permitem o correto funcionamento celular, como transcrição, divisão, sinalização e localização celular, reparo de DNA e regulação de *splicing*. Mutações em genes que participam destes processos fundamentais podem ser responsáveis por deficiências de crescimento tanto pré- quanto pós-natal, em geral associado com múltiplas anormalidades sistêmicas³⁸.

Alguns exemplos são as síndromes de Coffin-Síris (MIM#135900), que tem como a principal causa defeitos no gene *ARID1B* (MIM*614556), e a síndrome de Floating-Harbour (MIM#136140), causada por defeitos no gene *SRCAP* (MIM*611421), ambos os genes tem papel importante na remodelação da cromatina^{53,54}. Assim como a síndrome Cornélia de Lange (MIM#122470), causada principalmente por defeitos moleculares no gene *NIPBL* (MIM*608667), gene que codifica uma proteína essencial para a coesão das cromátides irmãs durante a multiplicação do DNA⁵⁵. Outro exemplo é a síndrome de Bloom (MIM#210900), que tem como causa mutações no gene *BLM* ou *RECQL3*

(MIM*600537), responsável pelo correto funcionamento da recombinação genética⁵⁶.

1.4.6 Síndrome dos genes contíguos

A síndrome dos genes contíguos compreende a deleção ou duplicação de um grupo de genes que são responsáveis por fenótipos complexos, como podemos exemplificar com a deleção da região 22q11.21, associada à síndrome de DiGeorge (MIM#188400), caracterizada por baixa estatura, aplasia/hipoplasia do timo e da tireoide, anormalidades cardíacas e faciais⁵⁷.

1.4.7 Defeitos de *imprinting*

O processo de *imprinting* determina a expressão diferencial dos alelos presentes por meio da metilação de um dos alelos, impedindo-o de ser expresso, e as doenças causadas por defeitos neste processo são denominadas doenças de imprinting. Como exemplo, a hipometilação do alelo paterno da *região de controle de metilação H19/IGF2* (MIM*616186) é responsável pela síndrome de Silver-Russel (MIM#100860), caracterizada por baixa estatura, face triangular, assimetria corpórea, fronte ampla, e algumas outras deformidades físicas⁵⁸.

1.5 Diagnóstico Clínico

O diagnóstico clínico destas síndromes é um desafio devido ao grande número de condições distintas de grande raridade. São descritas no OMIM

(Online Mendelian Inheritance in Man – www.OMIM.org) mais de 1360 síndromes de causa genética que apresentam baixa estatura dentre suas características fenotípicas⁵⁹, o que dificulta o diagnóstico clínico por falta de conhecimento e contato com a doença^{1,60,61}. Muitas destas condições ainda apresentam características sobrepostas entre si ou são fenotipicamente heterogêneas e variáveis^{1,61,62}. Nestes casos o estudo genético é essencial para estabelecer o diagnóstico preciso, que se mostra extremamente importante para a decisão quanto o acompanhamento e tratamento, visto que para algumas síndromes (ex: Síndrome de Bloom e Anemia de Fanconi) ocorre a predisposição á neoplasias e o tratamento com hormônio recombinante de crescimento humano (rhGH) é contraindicado⁶³. Adicionalmente, somente com o estabelecimento da etiologia a nível molecular é possível oferecer um aconselhamento genético preciso para o paciente e familiares.

Recentemente, tem se proposto uma abordagem genômica para investigação destes pacientes sindrômicos. Principalmente naqueles sem hipótese diagnóstica na avaliação clínica, o uso de estudos através do sequenciamento do exoma tem permitido estabelecer com precisão o diagnóstico etiológico com maior agilidade e menor custo que a abordagem tradicional^{5,38,64,65}. Em 2018, Hauer et al, realizou o sequenciamento do exoma de 200 indivíduos com baixa estatura sem causa conhecida e encontrou variantes deletérias em 16,5% do grupo estudado⁶⁶.

Nos últimos anos, nosso grupo avaliou pacientes com baixa estatura de início pré-natal e identificou a causa genética em 34% dos pacientes sindrômicos estudados⁶⁷ e em 15% em pacientes com baixa estatura isolada⁶⁸. Estes estudos

utilizaram em sua maioria a tecnologia de sequenciamento de genes alvos, em alguns casos foi utilizado o sequenciamento do exoma completo, mostrando o potencial destas abordagens que utilizam o SPLE em realizar o diagnóstico de pacientes com suspeita de doença monogênica.

No meu projeto de Mestrado que antecedeu o presente estudo, foram avaliados geneticamente pacientes com baixa estatura pré-natal com persistência pós-natal, classificados com Pequenos para Idade Gestacional (PIG), tanto aqueles com baixa estatura isolada quanto sindrômica através de um painel amplo de genes (390 genes) já associados a baixa estatura pela literatura. Durante o mestrado tivemos a oportunidade de realizar alguns sequenciamentos de exoma completo em pacientes sindrômicos sem causa clínica reconhecida, já pensando em um futuro projeto de doutorado que veio a se concretizar. Estes dados estão incluídos nesta tese.

2 OBJETIVOS

2.1 Objetivo Principal

O presente estudo teve como objetivo avaliar a utilidade do sequenciamento completo do exoma como primeira abordagem genético-molecular na investigação da etiologia da baixa estatura sindrômica de pacientes sem diagnóstico clínico reconhecido.

2.2 Objetivos específicos

1) Identificação da etiologia genética responsável pelo distúrbio de crescimento de crianças sindrômicas sem diagnóstico clínico, por meio do sequenciamento do exoma.

2) Caracterização de novas síndromes associadas com distúrbio de crescimento por identificar novos genes ou expandir o fenótipo associado a genes já implicados a síndromes específicas por descrição de casos atípicos.

3 MATERIAL E MÉTODOS

3.1 Considerações éticas

Todos os estudos foram conduzidos de acordo com princípios éticos seguindo as orientações contidas na declaração de Helsinki e pelos termos descritos pela Portaria 466/2012, do Conselho Nacional de Saúde. A pesquisa atual (CAAE: 37868114.3.0000.0068) foi aprovada junto a Comissão de Ética para Análise de Projetos de Pesquisa (CAPPesq) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) conforme o parecer 1.175.278. Foi obtido consentimento por escrito de todos os pacientes ou pais/tutores antes que os procedimentos de pesquisas fossem iniciados. A confidencialidade dos resultados obtidos foi garantida a todos os pacientes. Os achados de defeitos genéticos relacionados a doenças monogênicas em investigação foram informados aos pacientes e/ou seus responsáveis legais que terão aconselhamento genético e segmento conforme o recomendado pela Sociedade Brasileira de Genética Médica no projeto Diretrizes do Conselho Federal de Medicina.

3.2 Casuística

A casuística do projeto foi selecionada entre crianças submetidas à investigação clínica e laboratorial extensa conforme protocolos e formulários (Anexo I e II) de investigação de crianças com baixa estatura acompanhadas no

ambulatório da Unidade de Endocrinologia do Desenvolvimento do HCFMUSP. Estes pacientes constituem um grupo com alta probabilidade de apresentarem novos defeitos moleculares em genes ainda não estudados.

Os critérios de inclusão foram de pacientes com baixa estatura, caracterizada por escore-Z da altura inferior ou igual a menos -2^{69} ou baixa estatura em relação ao seu padrão familiar caracterizada por escore-Z da altura do paciente menos escore-Z da altura alvo inferior ou igual a menos 2, que apresentem outras características fenotípicas como dismorfismos faciais, malformações maiores, atraso de desenvolvimento neuropsicomotor e/ou déficit intelectual sem causa etiológica estabelecida.

Os critérios de exclusão utilizados foram a presença de doença que justifique o distúrbio de crescimento independentemente da condição sindrômica (ex: doenças crônicas, desnutrição, deficiência de GH, hipotireoidismo e outros); síndrome com diagnóstico clínico estabelecido; pacientes com displasia esquelética como principal manifestação do seu quadro sindrômico.

Foi dada preferência para análise pacientes com disponibilidade de amostra de DNA dos progenitores e/ou casos familiares.

3.3 Estudos Moleculares

3.3.1 Extração do DNA genômico

As amostras de DNA foram obtidas a partir de leucócitos de sangue periférico ou de células extraídas por *swab* da mucosa oral dos indivíduos

arrolados na pesquisa. O DNA foi extraído de acordo com procedimento padronizado no Laboratório de Hormônios e Genética Molecular da Unidade de Endocrinologia do Desenvolvimento, Faculdade de Medicina da USP (FMUSP), São Paulo. A concentração do DNA extraído foi obtida por leitura em espectrofotômetro no comprimento de onda de 260 nm (1 unidade DO (Densidade ótica) 260 = 50 µg/mL). A relação ideal entre as leituras em 260 e 280 nm para a caracterização da pureza do material é entre 1,80 e 2,00. As amostras foram mantidas congeladas a -20 °C até seu uso.

3.3.2 Genotipagem por sequenciamento paralelo em larga escala

O processo do sequenciamento para 86 pacientes seguiu os protocolos recomendados pelos fabricantes para o preparo de bibliotecas para enriquecimento das regiões alvos (exoma) e sequenciamento que foram realizados na facillite do SELA-FMUSP (Laboratórios de Sequenciamento em Larga Escala da FMUSP) (n= 41) com participação direta da aluna em todas as etapas ou nos seguintes serviços terceirizados BGI (n=2), Fleury (n=8) ou 3Billion Inc.(n= 35).

O kit de preparo das bibliotecas pelo SELA e pelo BGI foi o Sure Select Human All Exons V6 (Agilent Technologies, CA, USA) e pelo o Fleury e 3Billion o xGen Exome Research Panel v2 (Integrated DNA Technologies, Coralville, Iowa, USA). Já o sequenciamento propriamente dito ocorreu na plataforma HiSeq 2500 em modo *paired-end* quando realizado no SELA e na plataforma NovaSeq 6000 também em modo *paired-end* para os serviços terceirizados. Ambas as plataformas são da Illumina, CA, USA.

Todas as etapas que envolvem o sequenciamento completo do exoma foram previamente descritas na minha dissertação de mestrado no seguinte endereço: <https://teses.usp.br/teses/disponiveis/5/5135/tde-15082018-104927/pt-br.php> e resumidas na Figura 1.

Figura 1. Esquema das principais etapas do sequenciamento paralelo em larga escala. **1.** Primeiramente ocorre a fragmentação do material genômico e a ligação a adaptadores. **2.** Hibridação com sondas de biotina e captura dos éxons com beads de estreptavidina. **3.** Ligação dos adaptadores a pequenas sequências que ficam em uma placa denominada flowcell, onde ocorre a amplificação clonal. **4.** Ocorre o sequenciamento propriamente dito, que no caso é feito por síntese através de uma DNA polimerase, cada nucleotídeo adicionado emite um fluoróforo específico.

3.3.3 Cultura de fibroblastos

Para um dos casos, foi realizada cultura de fibroblastos a partir de biópsia de pele (*punch* de 3mm de diâmetro) em antebraço do paciente, realizada ambulatorialmente após assepsia e anestesia local. O produto da biópsia de pele foi fragmentado em aproximadamente 18 pequenos fragmentos que foram transferidos para frascos apropriados para cultura celular de 25cm², contendo AmnioMaxTM-C100 Basal com 20% de suplemento AmnioMaxTM-C100. Os frascos de cultura foram mantidos em estufa a 37°C com 5% de CO₂. O meio de cultura foi trocado a cada 2/3 dias e depois de 10 a 12 dias foi observado no microscópio de inversão os primeiros fibroblastos aderentes que resultaram do fragmento inicial. Com a proliferação dos fibroblastos, as células que ficaram aderentes à garrafa foram tratados com tripsina e transferidos para frascos maiores de 75cm², onde expandiram até atingir a confluência. Passado este período, os fibroblastos foram transferidos para um criotubo e congelados em nitrogênio líquido para estudos posteriores.

3.3.4 Confeção de cDNA e amplificação por PCR

Foi feita a extração do RNA total das células de fibroblastos e controles pelo método do trizol⁷⁰. O RNA extraído foi submetido a RT-PCR (*reverse transcription step*/transcrição reversa), no qual o cDNA (DNA complementar) é transcrito através do uso de *primers* randômicos presentes do kit High Capacity cDNA Archive Kit. Foram desenhados *primers* específicos para as regiões do cDNA (transcrito: ENST00000550536.5) correspondentes aos exons 1 e 2, 3 e 4, 8

e 9 do gene *RAB3IP*, e para o cDNA completo do gene *CDK4* (Tabela 2) para amplificação e comparação da expressão deles nos cDNAs do paciente e do controle.

Tabela 2 - *Primers* utilizados para amplificação do cDNA do paciente com variante no gene *RAB3IP* e *CDK4*. Fw:Forward; Rv:Reverse

Nome	Sequência
<i>RAB3IP_1e2Fw</i>	5' GCAGCATGTGAAGAGTTGCC 3'
<i>RAB3IP_1e2Rv</i>	5' AGCCTTCCAAGGGATCATTAGC 3'
<i>RAB3IP_3e4Fw</i>	5' TGTCTCGTTTACGAAGCCCA 3'
<i>RAB3IP_3e4Rv</i>	5' AGCTGTGAGTTCTTCCAATTCCT 3'
<i>RAB3IP_8e9Fw</i>	5' TGGCTTCAGCTGTTCTGGAG 3'
<i>RAB3IP_8e9Rv</i>	5' TAGTTGCTTGAGTCCCCTAAT 3'
<i>CDK4_1Fw</i>	5' GTGTATGGGGCCGTAGGAAC 3'
<i>CDK4_1Rv</i>	5' CCAGCTTGACTGTTCCACCA 3'
<i>CDK4_2Fw</i>	5' TGCGCCAGTTTCTAAGAGGC 3'
<i>CDK4_2Rv</i>	5' GAGGAGGACCCTCCATAGCC 3'

3.3.5 MLPA

Em um dos pacientes foi encontrada uma duplicação envolvendo a região de *imprinting* da Síndrome de Silver Russel (SSR). E para confirmação da herança da variante na família, foi realizado o SALSA MLPA® Probemix ME030 BWS/RSS (MRC-Holland, Amsterdã, Países Baixos). O protocolo de execução seguiu as instruções do fabricante.

3.3.6 aCGH

Os testes de aCGH (array *Comparative genomic hybridization*/ Hibridação genômica comparativa) foram realizados em 57 pacientes previamente ao WES. Adicionalmente, três pacientes com achados altamente sugestivos de CNVs na análise do WES realizaram aCGH para confirmação dos achados. Estas análises foram realizadas em colaboração com as Dras. Ana Krepischi e Carla Rosenberg no Instituto de biociências da USP em uma plataforma de alta resolução 180K (Agilent, CA, EUA).

4 ANÁLISE DOS RESULTADOS

4.1 Bioinformática

A partir dos arquivos gerados no processo de sequenciamento acima descritos, foi realizado a análise bioinformática seguindo as seguintes etapas em colaboração com o Dr. Antonio Lerário, responsável pelo processamento dos dados brutos até a geração da VCF. Sequenciamentos do exoma que tiveram cobertura aquém do considerado aceitável (92% ou menos das regiões alvo com pelo menos 10x de cobertura) foram desconsideradas da análise final.

I. Alinhamento das sequências geradas a uma sequência referência do genoma humano (hg19 UCSC ou b37 GRC/NCBI): nesta etapa, a origem de cada fragmento do DNA da amostra é determinada. Inicialmente o alinhamento é feito sequência a sequência individualmente. Após esta primeira etapa, é necessário recalibrar o alinhamento, dado que, em determinadas situações, diferentes sequências provenientes da mesma região cromossômica apresentem resultados ligeiramente diferentes de mapeamento (por exemplo, quando existem pequenas deleções, inserções ou repetições no genoma alvo). A recalibragem é considerada um realinhamento de forma contextualizada. Nesta etapa foi utilizado o programa BWA (<http://bio-bwa.sourceforge.net/>), e como resultado gera-se o arquivo com extensão “*.bam”.

II. Genotipagem: aqui é realizada a genotipagem da amostra a partir do melhor alinhamento possível. E consiste em determinar, com bases estatísticas,

as bases de cada par do cromossomo (alelos) para uma dada posição da amostra sequenciada. Freebayes foi utilizado para esta etapa. Esta ferramenta identifica SNVs e Indels;

III. Anotação dos genótipos variantes detectados: esta etapa consiste em (1) anotar as variantes encontradas utilizando o programa Annovar (<http://www.openbioinformatics.org/annovar/>) de acordo com bancos públicos de variantes (por exemplo, do gnomAD) e (2) estimar o efeito de cada variante, por exemplo, se neutro ou deletério na proteína sintetizada *in silico*. O resultado é apresentado em uma planilha em formato Excel chamada Variant Call Format (VCF);

IV. Priorização das variantes candidatas: com toda a base de dados gerada a partir das etapas anteriores, algoritmos heurísticos de priorização são empregados para gerar uma lista ordenada e, de preferência, com poucas variantes para análise do pesquisador. Não existem critérios fixos e padrão para esta priorização. Em geral, considera-se desde parâmetros de alinhamento, como quantidade de *reads* confirmando a variante até ser uma variante supostamente deletéria ou não. Nesta etapa há uma grande participação do pesquisador e depende do conhecimento do mesmo em relação ao fenótipo analisado.

A VCF, contém todas as variantes genéticas identificadas nos pacientes, que são analisadas pelo pesquisador. Neste banco de dados foram aplicados os seguintes filtros sequenciais para gerar uma lista limitada de variantes potencialmente relacionadas com o fenótipo (Figura 2).

Figura 2. Fluxo da estratégia de filtragem utilizada para selecionar as variantes com possibilidade de patogenicidade em 86 pacientes com baixa estatura síndrômica submetidos ao sequenciamento do exoma. LoF: Variante Loss of Function/Perda de função. MAF: Minor Allele Frequency/Frequência do alelo menos comum.

- I. Variantes presentes no paciente, em homozigose, hemizigose, heterozigose composta ou heterozigose simples de acordo com o modelo de herança sugerido pela análise de familiares;
- II. Foram analisadas apenas variantes raras, com frequência menor ou igual a 1% em bancos de dados públicos, como os bancos 1000 genomas (<http://browser.1000genomes.org/>) e ExAC (<http://exac.broadinstitute.org/>), além de um banco de controles da variação genética brasileira, obtidos em uma população de 609 idosos saudáveis (1218 alelos) (<http://abraom.ib.usp.br>)⁷¹. Além de variantes com frequência baixa (<1%) em banco de dados interno do laboratório (SELA: <https://intranet.fm.usp.br/sela/>), que contém 523 indivíduos, o que corresponde a 1046 alelos⁷²;
- III. Filtrou-se apenas as variantes presentes em Exons ou Splicing site (+5 e -5)
- IV. Foram selecionadas variantes com potencial efeito deletério sobre a função do gene: stop-códon, ganho ou perda, mudança de frame para a leitura correta dos aminoácidos e as não sinônimas. Para estas últimas, foram analisadas as predições de patogenicidade para as variantes de troca de um único nucleotídeo (SNVs – *single-nucleotide variants*) por diversas ferramentas *in silico*;
- V. Das variantes selecionadas até esta etapa, foi analisada a compatibilidade com a herança sugerida e a plausibilidade biológica em relação ao fenótipo: MIM: <http://www.MIM.org/>;
- VI. Avaliou-se ainda banco de dados de mutações associadas a doença:

4.2 Análise de Variantes Número de Cópias (CNVs)

Além da análise de variantes de ponto e pequenos indels, foram realizadas em colaboração com o Dr. Antonio Lerário, a análise de variantes de número de cópias (CNVs) através de dois novos sistemas. O ClinCNV⁷³, um novo método para detecção multiamostra de CNV em dados de NGS baseado em duas medidas de intensidade, a cobertura de leitura e a frequência de alelos B (BAF - medida normalizada da relação de intensidade alélica para cada SNP). E o DELLY⁷⁴, um método de predição de variante estrutural (VE) que pode identificar, genotipar e visualizar deleções, duplicações, inversões e translocações em resolução de nucleotídeo único em dados de sequenciamento massivo paralelo de leitura curta. Ele usa fragmentos *paired-ends*, *split-reads* e cobertura dos fragmentos para delinear com sensibilidade e precisão os rearranjos genômicos em todo o exoma. A filtragem dos dados foi feita através de uma planilha excel, e apenas as CNVs com frequência nula na casuística incluída foram consideradas. Os dados gerados são visualmente analisados pelo IGV (Figura 4).

Figura 4. Exemplo da visualização de uma Duplicação (acima em azul) e uma Deleção (abaixo em vermelho) identificado pelo programa ClinCNV, indicados por seta no programa IGV (*Integrative Genomics Viewer*).

4.3 Classificação das variantes de interesse

A classificação das variantes em relação ao fenótipo do paciente foi realizada seguindo o guia de recomendação da ACMG-AMP (American of Medical Genetics and Genomics e Association for Molecular Pathology)⁷⁵. Este guia propõe linhas de evidência para classificação de patogenicidade, que inclui frequência na população, dados de análises funcionais, análise *in silico*, dados de segregação, dentre outros. As evidências são classificadas como muito forte

(VS), forte (S), moderada (M) e pequena (P). Dependendo da combinação destas evidências, a variante é classificada dentre uma das 5 opções: patogênica (P), provavelmente patogênica (PP), variante de significado incerto (VUS), provavelmente benigna (PB) e benigna (B) (Anexos III e IV). Para apoiar a interpretação da segregação familiar, foi adotado um critério matemático quantitativo que foi adicionado recentemente ao guia de interpretação da ACMG-AMP, que promove uma quantificação estatística dos dados da segregação familiar⁷⁶. Utilizamos também o auxílio de um programa que classifica as variantes instantaneamente de acordo com este guia, com a possibilidade de editar pessoalmente os resultados para outras evidências, como compatibilidade fenotípica e segregação, denominado Varsome (<https://varsome.com>).

Para avaliar a possível patogenicidade das CNVs, os dados foram comparados com variantes já descritas no Decipher (<https://decipher.sanger.ac.uk/>) e no UCSC (<https://genome.ucsc.edu/>) e os genes incluídos nas CNVs foram individualmente analisados para associação com o fenótipo dos pacientes.

5 RESULTADOS

5.1 Caracterização da casuística

Para o projeto, foram incluídos pacientes selecionados pelo ambulatório do HCFMUSP. Dentre 166 pacientes classificados como baixa estatura sindrômica sem diagnóstico clínico, foram selecionados 86 pacientes para análise por WES, sendo 57 com resultado prévio negativo por análise de CNV por aCGH.

Dos 86 pacientes submetidos aos WES, 25 foram realizados em trios, ou seja, com 2 parentes, e 61 apenas o caso índice isolado. Desta casuística, 47 pacientes são do sexo masculino e 39 do feminino com uma média de 8,7 anos de idade na avaliação inicial. O Escore-Z médio de comprimento ao nascimento foi de $-2,7 (\pm 1,6)$ e de peso ao nascimento $Z = -1,5 (\pm 1,6)$. Dentre características adicionais à baixa estatura, 30 apresentam microcefalia e 54 apresentam retardo do desenvolvimento neuropsicomotor (DNPM) e/ou dificuldades intelectuais (Tabela 3).

Tabela 3 - Característica da casuística de 86 pacientes sindrômicos sem diagnóstico clínico estudadas por WES

Características	86 pacientes	Resultados positivos (n=41)	Resultados negativos (n=45)
Consanguinidade (%)	14 (16,1%)	7 (50%)	7 (50%)
Z-escore da altura alvo	-2,4 ± 1,5	-2,4 ± 1,4	-2,5 ± 1,5
Baixa estatura familiar	19 (21,8%)	9 (47,4%)	10 (52,6%)
Sexo (M:F)	47:39	24:17	23:22
Idade Gestacional (semanas)	36,6 ± 4	37,5 ± 3,3	35,8 ± 4,4
Prematuro (%)	31 (35,6%)	10 (32,2%)	21 (67,8%)
Z-escore de Peso ao nascimento	-1,5 ± 1,6	-1,1 ± 1,6	-1,7 ± 1,5
Z-escore de Comprimento ao nascimento SDS	-2,7 ± 1,6	-2,5 ± 1,5	-2,8 ± 1,7
Criança nascida pequena para idade gestacional	53 (60,9%)	11 (20,8%)	42 (79,2%)
Idade da 1ª avaliação (anos)	8,7 ± 6,0	10,2 ± 6,2	7,2 ± 5,5
Z-escore de Altura	-3,3 ± 1,4	-3,1 ± 1,2	-3,1 ± 1,7
Z-escore do IMC	-1,1 ± 2,6	-0,4 ± 1,8	-1,7 ± 3,1
Z-escore da altura sentada	0,6 ± 1,6	0,6 ± 1,6	0,7 ± 1,6
Microcefalia (%)	30 (34,5%)	14 (46,6%)	16 (53,4%)
Atraso do DNPM e/ou dificuldades intelectuais (%)	54 (62,1%)	27 (50%)	27 (50%)
Mal formações cardíacas	15 (17,2%)	10 (66,7%)	5 (33,3%)
Mal formações do SNC	18 (20,7%)	8 (44,4%)	10 (55,6%)

M: masculino; F: feminino

A Figura 5 mostra o Fluxograma dos pacientes com baixa estatura sindrômica do ambulatório do HCFMUSP.

Figura 5. Fluxograma dos pacientes com baixa estatura sindrômica do ambulatório do HCFMUSP. Dos 818 pacientes atendidos, 166 são sindrômicos sem diagnóstico clínico. Destes, 57 tiveram aCGH prévio com resultado negativo e foram encaminhados para WES. No total 86 realizaram exoma e 41 pacientes tiveram resultado positivo.

5.2 Resultados genético-moleculares

As métricas médias das corridas foram satisfatórias com uma cobertura média de 152x e aproximadamente 98,7% das regiões alvo tiveram 10 *reads* de cobertura (Tabela 4).

Tabela 4 - Métricas médias das corridas de sequenciamento do exoma realizadas durante o projeto

	Cobertura média por exoma (%)	Região alvo com ≥ 10 <i>reads</i> (%)	Região alvo com ≥ 20 <i>reads</i> (%)
Média	152,07	98,7	97,9
Mínimo	80,2	92,7	86,3
Máximo	272,2	99,7	99,2

Foram encontrados resultados positivos em 41 dos 86 pacientes (47,7%). Em 22 pacientes foram encontradas variantes consideradas patogênicas (P) e nos outros 19 foram identificadas variantes provavelmente patogênicas (PP). A maioria das variantes foi do tipo *missense* (n=23), seguida por variantes *frameshift* e *stopgain*, com 8 variantes cada, de alteração do sítio de *splicing* (n=3) e CNVs (n=4) (Tabela 5).

Tabela 5 - Variantes patogênicas ou provavelmente patogênicas encontradas em 41 de 86 pacientes da casuística

ID	Posição	Gene	Tipo	Variante	MAF	Estado	Herança	Classificação ACMG/AMP
1	chr7:5568098	ACTB	Missense	c.616C>T;p.Arg206Trp	0	Heterozigose	ND	Provavelmente Patogênica
2	chr5:132269985	AFF4	Missense	c.772C>T;p.Arg258Trp	0	Heterozigose	De novo	Patogênica ⁶⁷
3	chr16: 89350548	ANKRD11	Frameshift	c.2398_2401del; p.Glu800Asnfs*62	0	Heterozigose	De novo	Patogênica ⁶⁷
4	chr14:99723993	BCL11B	Frameshift	c.242delG;p.Cys81Leufs*76	0	Heterozigose	De novo	Patogênica ⁷⁹
5	chr14:99640721	BCL11B	Frameshift	c.2439_2452dupGCACCGGC; GGAGCC;	0	Heterozigose	De novo	Patogênica
6	chr17:41244839	BRCA1	Stopgain	p.His818ArgfsTer31 c.2709T>A;p.C903*	0	Homozigose	Herdada de pais consanguíneos	Patogênica ^{67,80}
7	chr8:61777565	CHD7	Sítio de splicing	c.8077-10T>A	0	Heterozigose	De novo	Provavelmente Patogênica
8	chr12: 48380136	COL2A1	Missense	c.1510G>A;p.Gly504Ser	0	Heterozigose	De novo	Patogênica ⁶⁷
9	chr12:48378364	COL2A1	Missense	c.1852G>A :p.Gly618Ser	0	Heterozigose	Herdada da mãe afetada	Provavelmente Patogênica ⁶⁷
10	chr12:48376911	COL2A1	Missense	c.2059G>A; p.Gly687Ser	0	Heterozigose	ND	Provavelmente Patogênica

continua

Tabela 5 - Variantes patogênicas ou provavelmente patogênicas encontradas em 41 de 86 pacientes da casuística - continuação

ID	Posição	Gene	Tipo	Variante	MAF	Estado	Herança	Classificação ACMG/AMP
11	chr17:1939344 chr17:1943063	DPH1	Missense	c.374T>C;p.Leu125Pro c.710G>A;p.Gly237Asp	0 0	Heterozigose composta	Herdada da mãe Herdada do pai	Provavelmente Patogênica
12	chr21:38853126	DYRK1A	Stopgain	c.514C>T;p.Gln172*	0	Heterozigose	De novo	Patogênica
13	chr15:48755320	FBN1	Missense	c.5183c>Tp.Ala1728Val	0	Heterozigose	ND	Patogênica ⁸¹
14	chr3:71027008	FOXP1	Stopgain	c.1319C>G;p.Ser440*	0	Heterozigose	ND	Patogênica
15	chr20:25394489 chr20:25398748	GINS1	Missense	c.139G>A;p.Val47Met c.247C>T;p.Arg83Cys	0 0.000887	Heterozigose composta	Herdada do pai Herdada da mãe	Patogênica ⁶⁷
16	chrX:71787820	HDAC8	Missense	c.356C>T;p.Thr119Met	0	Heterozigose	ND	Provavelmente Patogênica
17	Chr15:99467760	IGF1R	Stopgain	c.2629C>T;p.Arg877*	0	Heterozigose	Herdada do pai	Patogênica
18	chr17:1399973	INPP5K	Missense	c.1088T>C;p.Ile363Thr	0	Homozigose	ND	Provavelmente ⁶⁷ Patogênica
19	chrX: 53239931	KDM5C	Missense	c.1510G>A;p.Val504Met	0	Hemizigose	Ligada ao X recessivo	Provavelmente Patogênica
20	chr10:994405252 _994405254:CT>	KIF11	Frameshift	c.2402_2403del; p.Ser801*	0	Heterozigose	Herdada do pai Presente na irmã afetada	Patogênica ⁶⁷
21	chr11:118376518	KMT2A	Frameshift	c.9911delT;p.Leu3304Argfs*19	0	Heterozigose	De novo	Patogênica ⁶⁷

continua

Tabela 5 - Variantes patogênicas ou provavelmente patogênicas encontradas em 41 de 86 pacientes da casuística - continuação

ID	Posição	Gene	Tipo	Variante	MAF	Estado	Herança	Classificação ACMG/AMP
22	chr12:25398254	KRAS	Missense	c.65A>G;p.Gln22Arg	0	Heterozigose	ND	Patogênica
23	chr12:110029080 chr12:110034320	MVK	Missenses	c.803T>C;p.Ile268Thr c.1129G>A p.Val37Ile	0.00019 0.00159	Heterozigose composta	DNA do pai ausente Herdada da mãe	Patogênica
24	chr21:47851791 chr21:47856994	PCNT	Stopgain Missense	c.8413C>T;p.Gln2805* c.9099G>C;p.Gln3033His	0 0.00000399	Heterozigose composta	Herdada de pais consanguíneos	Patogênica
25	chrX:19375800	PDHA1	Missense	c.976C>T;p.Arg326Cys	0	Heterozigose	De novo	Patogênica
26	chr3:52183830	POC1A	Sítio de splicing	c.275+2T>G	0	Homozigose	Herdada de pais consanguíneos	Patogênica ⁶⁷
27	chr15:89868870 chr15:89873415	POLG1	Missenses	c.1760C>T;p.Pro587Leu c.752C>T;p.Thr251Ile	0.00150 0.00143	Heterozigose composta	ND	Provavelmente Patogênica
28	chr8:144906508	PUF60	Frameshift	c.83_86dupCAGC;p.Gly30SerfsTer111	0	Heterozigose	ND	Provavelmente Patogênica
29	chr1:93306144	RPL5	Stopgain	c.742C>T;p.R248*	0	Heterozigose	De novo	Patogênica
30	chrX:20183015	RPS6KA3	Sítio de splicing	c.1764+2T>C	0	Hemizigoto	De novo	Provavelmente Patogênica
31	chr2:166897755	SCN1A	Frameshift	c.2400delA;p.Thr801Glnfs*2	0	Heterozigose	ND	Provavelmente Patogênica
32	chr6:35200977	SCUBE3	Missense	c.611G>A;p.Gly204Asp	0	Homozigose	Herdada de pais consanguíneos	Provavelmente Patogênica ⁷⁸
33	chr3:9489439	SETD5	Stopgain	c.1852C>T;p.Arg618*	0	Heterozigose	ND	Patogênica

continua

Tabela 5 - Variantes patogênicas ou provavelmente patogênicas encontradas em 41 de 86 pacientes da casuística - continuação

ID	Posição	Gene	Tipo	Variante	MAF	Estado	Herança	Classificação ACMG/AMP
34	chr19:11130280	SMARCA4	Frameshift	c.2519_2542delCAAGACGGGC CTTTGTCCCCCAGCinsGGA;p. Ala840_Leu848delinsGlylle	0	Heterozigoto	De novo	Provavelmente patogênica
35 e 36	chr16:30748691	SRCAPx2	Stopgain	c.7330C>T;p.R2444*	0	Heterozigose	De novo	Patogênica ^{67,77}
37	chr17:40359693	STAT5B	Missense	c.1960C>T;p.Arg654Trp	0	Homozigose	ND	Provavelmente Patogênica
38	chr5:1264582 chr5:1260648	TERT	Missenses	c.2780T>G;p.Leu927Arg c.2911C>T;p.Arg971Cys	0 0	Heterozigose composta	Herdada do pai Herdada da mãe	Provavelmente Patogênica
39	chr11 (p15.5- p15.4):2290980- 3845365x3	CNV	Duplicação	Dup:1.554.385pb	0	Heterozigose	Herdada da mãe	Patogênica
40	chr5 (q35.2- q35.3):175394113- 177422934x3	CNV	Duplicação	Dup:2.028.821 pb	0	Heterozigose	ND	Patogênica
41	chr8 (q24.22- q24.3):133108784_1462 94098x3 chr15 (q26.2- q26.3):97439378_10238 3473x1	CNV CNV	Duplicação Deleção	Dup:13.200.000 pb Del:4.900.000 pb	0 0	Heterozigose Heterozigose	ND ND	Patogênica

ND: Não Disponível DNA dos pais; MAF = frequência do alelo menos comum (Minor allele frequency) baseado na frequência máxima observada no banco de dados no gnomAD (<https://gnomad.broadinstitute.org/>).

conclusão

Todas as variantes encontradas apresentam frequência nula ou muito baixa compatível com a apresentação da doença (critério PM2 pelo ACMG-AMP), tem predição *in silico* de serem patogênicas (PP3), as variantes perda de função são em genes que não toleram este tipo de variante (PVS1) e os pacientes apresentam fenótipo compatível com o achado molecular (PP4). E para alguns casos a segregação famílias da variante foi confirmada (PP1). O total de variantes SNVs (Variantes de nucleotídeo único) foi de 44, a maioria foi identificada em heterozigose simples ($n = 24$) e a patogenicidade foi analisada de acordo com modelos dominantes de herança, enquanto 12 variantes encontram-se em estado heterozigoto composto em 6 pacientes, e a análise foi realizada de acordo com modelo recessivo, assim como cinco variantes encontradas em homozigose. Duas variantes foram encontradas no cromossomo X de pacientes do sexo masculino, portando consideradas em hemizigose e foram analisadas conforme um modelo de herança ligada ao X recessivo.

Dos 86 pacientes submetidos ao WES, 57 tiveram a análise de CNVs pelos novos softwares ClinCNV e DELLY. E foram encontradas 4 CNVs em 3 pacientes, todas em heterozigose (5,3%), posteriormente confirmadas por aCGH (Figura 6).

Figura 6. Estado das 47 variantes Patogênicas ou Provavelmente Patogênicas encontradas em 41 pacientes.

Doze das variantes foram publicadas no nosso artigo no *Journal of Pediatrics*⁶⁷, compreendendo pacientes sindrômicos que nasceram Pequenos para Idade Gestacional (PIGs) que não apresentaram recuperação de crescimento até o segundo ano de vida. Adicionalmente, dois pacientes que tiveram uma mesma variante em heterozigose de novo no gene *SRCAP*. Após o achado molecular foi confirmado o diagnóstico de síndrome de Floating Harbour (SFH) e foram incluídos em um artigo original que reuniu os pacientes do nosso serviço com esta rara condição⁷⁷.

O paciente (ID 32) com variante no gene *SCUBE3*, foi incluído em um estudo colaborativo, publicado no *American Journal of Human Genetics* descrevendo uma nova síndrome com fenótipo consistente caracterizado por retardo de crescimento, características esqueléticas, aparência craniofacial

distinta e anomalias dentárias⁷⁸. A colaboração neste estudo foi feita através da estratégia de divulgar genes candidatos onde variantes de interesse foram encontradas no *GeneMatcher* (<https://genematcher.org/>). O *Genematcher* é um banco de dados on-line que visa combinar pesquisadores que estudam pacientes com a apresentação de uma doença rara com base em genes de interesse. Quando dois ou mais pesquisadores enviam o mesmo gene ao banco de dados, o serviço os combina para permitir a comparação de casos. Neste caso, foi possível a parceria com pesquisadores de diversos grupos de pesquisa do mundo, que apresentavam pacientes com variantes em homozigose no gene *SCUBE3*, com fenótipo consistente.

Das 4 CNVs identificadas em 3 pacientes através da análise do WES, uma é uma duplicação de uma região do chr11 (p15.5-p15.4):2290980-3845365x3, envolvendo os genes *ASCL2*, *C11ORF21*, *TSPAN32*, *CD81*, *CD81-AS1*, *TSSC4*, *TRPM5*, *KCNQ1*, *KCNQ1OT1*, *KCNQ1 AS1*, *CDKN1C*, *SLC22A18AS*, *SLC22A18*, *PHLDA2*, *NAP1L4*, *CARS1*, *CARS1-AS1*, *OSBPL5*, *MRGPRG*, *MRGPRG-AS1*, *MRGPPE*, *ZNF195*, *ART5*, *ART1*, *CHRNA10*, *NUP98* e *PGAP2*. A outra, uma duplicação no chr5 (q35.2-q35.3):175394113-177422934x3, envolvendo os seguintes genes: *THOC3*, *FAM153B*, *SIMC1*, *KIAA1191*, *ARL10*, *NOP16*, *HIGD2A*, *CLTB*, *FAF2*, *RNF44*, *CDHR2*, *GPRIN1*, *SNCB*, *EIF4E1B*, *TSPAN17*, *UNC5A*, *HK3*, *UIMC1*, *ZNF346*, *FGFR4*, *NSD1*, *RAB24*, *PRELID1*, *MXD3*, *LMAN2*, *RGS14*, *SLC34A1*, *PFN3*, *F12*, *GRK6*, *PRR7-AS1*, *PRR7*, *DBN1*, *PDLIM7*, *DOK3*, *DDX41*, *FAM193B*, *TMED9*, *B4GALT7*, *FAM153A* e *PROP1*.

As outras duas CNVs foram encontradas em um único paciente. Uma é uma duplicação no chr8 (q24.22-q24.3):133108784-146294098x3 envolvendo 112 genes, entre eles genes já conhecidos como causa de baixa estatura síndrômica, como *PUF60* e *RECQL4*, e uma deleção no chr15 (q26.2-q26.3):97439378-102383473)x1 envolvendo os seguintes genes: *ARRDC4*, *FAM169B*, *IGF1R*, *PGPEP1L*, *SYNM*, *TTC23*, *LRRC28*, *MEF2A*, *LYSMD4*, *ADAMTS17*, *CERS3*, *LINS1*, *ASB7*, *ALDH1A3*, *LRRK1*, *CHSY1*, *SELENOS*, *SNRPA1*, *PCSK6*, *TM2D3*, *TARS3*, *OR4F6*, *OR4F15*, *OR4F4*. Todas as CNVs foram posteriormente confirmadas aCGH.

5.3 Investigação de novos genes candidatos

Além das variantes patogênicas ou provavelmente patogênicas, foram encontradas variantes consideradas de significado incerto (VUS) em genes sem fenótipo descrito em humanos até o momento, porém participam de vias moleculares essenciais. Estas variantes são preditas como deletérias em modelos *in silico* e possuem frequência em bancos de dados (gnomAD) nula ou muito baixa. Portanto, consideramos estes genes candidatos, para os quais utilizamos a estratégia de divulgar no GeneMatcher.

Dentro desta estratégia, foram submetidos 6 genes (Tabela 6). Além do *SCUBE3* descrito na sessão anterior, que foi associado a uma nova síndrome coma utilização desta estratégia. Para um deles, o *TBC1D32* tivemos resposta compatível e estamos escrevendo um artigo em esforço colaborativo com diversos centros de pesquisa no mundo, que apresentam famílias com casos de

variante em homozigose no gene *TBC1D32*, associada a um fenótipo de síndrome Orofacial Digital (OFD) Tipo IX, ainda não descrito no MIM. Neste estudo serão incluídas 4 famílias com variantes em homozigose neste gene em pacientes com fenótipo consistente de dismorfismos orofaciais distintos, anormalidades ósseas, microcefalia e retardo de DNPM.

Tabela 6 - Variantes em genes de interesse submetidas no GeneMatcher

Gene	Variante	MAF (gnomAD)	Estado	Genematcher
CDK4	c. 218G>A:p.Arg73Gln	0	Homozigose	Resposta positiva
IFT88	c.1945A>G:p.Ile649Val	0	Heterozigose de novo	Resposta negativa
MAP1A	c.4367_4372del; p.Ala1456_Leu1457del	0	Heterozigose de novo	Resposta negativa
RAB3IP	c.13A>T:p.Lys5*	0	Homozigose	Respostas negativas
SCUBE3	c.611G>A:p.Gly204Asp	0	Homozigose	Resposta positiva Estudo já publicado ⁷⁸
TBC1D32	c.3527A>G; p.His1176Arg	0.00000803	Homozigose	Resposta positiva Artigo em andamento

Um outro gene de interesse onde encontramos uma variante em homozigose de ganho de stop códon, foi o *RAB3IP* (c.13A>T:p.Lys5*) em um paciente do sexo masculino, filho de pais consanguíneos, com baixa estatura, microcefalia, retardo de desenvolvimento neuropsicomotor, TSH elevado e anemia. A variante é ausente em bancos populacionais em um gene que é fator

de troca de nucleotídeos de guanina (GEF) e promove a troca de GDP para GTP, convertendo proteínas Rab inativas ligadas ao GDP em sua forma ativa ligada ao GTP, quando ativados participam da organização da actina e do transporte de vesículas para a superfície celular⁸². Por se tratar de um gene ainda não associado a doença em humanos, e ser uma variante perda de função, foi realizado um estudo de avaliação do cDNA do paciente comparado a 2 controles. Para tanto, cultivamos fibroblastos extraídos de biópsia do antebraço do paciente e o RNA total foi extraído dessas células. Após isto, o cDNA foi confeccionado por transcrição reversa. Para avaliar o cDNA, foram desenhados 3 pares de primers específicos para 3 regiões diferentes do gene com o objetivo de observar se ocorreria RNA *decay* (degradação do mRNA mediada por mutação sem sentido), porém não observamos diferença e todos os fragmentos foram amplificados tanto no paciente quanto nos controles (Figura 7). Este resultado em associação com ausência de fenótipo semelhante com outros pacientes com variantes identificadas neste mesmo gene reportado no GeneMatcher, motivou a busca de outras potenciais alterações genéticas como responsáveis pelo fenótipo.

Figura 7. A: Transcrito [ENST00000550536.5](#) do gene *RAB3IP* utilizado para desenho dos primers. Setas Azuis, correspondem as regiões dos exons 1 e 2. Setas laranjas, as regiões dos exons 3 e 4. E setas vermelhas, as regiões dos exons 8 e 9. **B:** Foto do gel de eletroforese da amplificação do cDNA do paciente com variante no gene *RAB3IP*. A região correspondente aos exons 1 e 2 não foram amplificadas. Foram amplificadas as regiões correspondentes aos exons 3 e 4 e 8 e 9 em paciente e controles. **Ladder:** 1 kb da Thermo Fisher Scientific. **C-:** Controle negativo (DNA genômico); **C1:** Controle 1; **C2:** Controle 2 (cDNA de pacientes normais).

Neste mesmo paciente, foi feita a reanálise do WES, e foi encontrada uma variante missense em homozigose no gene *CDK4*: c.218G>A:p.Arg73Gln, ausente no gnomAD e com frequência muito baixa e compatível com variantes em homozigose no Abraom (MAF=0,000821), com predição de alteração de sítio de *splicing*. O gene *CDK4* codifica uma kinase dependente de ciclina que participa do ciclo celular, sendo assim considerado um gene do grupo de genes de mecanismo celular fundamental⁸³. Até o presente momento não existem fenótipos associados de variantes em homozigose no *CDK4* em humanos. Mas estudos de *Knockout* do *CDK4* em camundongos mostram um fenótipo consistente ao do nosso paciente e por isso o interesse pelo gene nos levou a

submete-lo ao *GeneMatcher*. Recebemos um contato de um pesquisador da Universidade de Edimburgo, que tem uma família com vários membros afetados compartilhando um ganho de *stogain* em homozigose no mesmo gene *CDK4*. Estes pacientes apresentam microcefalia e baixa estatura em semelhança ao nosso paciente, não apresentam dismorfismos faciais e estão em análise do perfil hormonal. No momento da redação desta tese, o processo para iniciar uma colaboração estava em andamento, porém acreditamos em uma possível nova associação genótipo-fenótipo.

Afim de confirmar uma alteração no RNA mensageiro resultante da alteração do sítio de *splicing* pela variante, foi realizada a mesma abordagem descrita anteriormente. A partir de células de fibroblastos do paciente, o RNA foi extraído e o cDNA confeccionado. A região correspondente ao *CDK4* foi amplificada por PCR utilizando os *primers* da Tabela 2 e o resultado foi comparado a um cDNA controle. Podemos observar que a banda correspondente ao amplificado pelo par de *primers* denominados 1 é menor no paciente (329pb) quando comparado ao controle (566pb). Já o amplificado 2 se mantém normal (Figura 8b). Foi realizado o sequenciamento destas regiões por Sanger para determinar a localização do sítio *splicing* alterado onde podemos observar a deleção de 237 pb com a perda do códon consenso de início pela deleção do exon 2 (Figura 8c).

Com este resultado podemos inferir que há alteração na proteína Kinase resultante de *CDK4* e provavelmente alteração de sua função no ciclo celular. Mas para estas confirmações mais estudos funcionais são necessários.

Pretendemos realizá-los em colaboração com o grupo Universidade de Edimburgo em um projeto futuro.

Figura 8. **a)** Transcrito ENST00000257904.11 do gene CDK4 utilizado para desenho dos 2 pares de primers (Tabela 2) que cobrem o cDNA completo. A seta vermelha indica a variante c.218G>A. **b)** Foto do gel de eletroforese da amplificação do cDNA do paciente com variante no CDK4. A região correspondente ao par de primer 1 é menor no paciente (329pb) quando comparada ao controle (566pb). **Ladder:** 1 kb da Thermo Fisher Scientific. **c:** Eletroferograma do sequenciamento do cDNA do paciente e de um controle, mostrando o ponto novo de *splicing*.

Outra variante de interesse foi encontrada em homozigose gene *PCNA* (c.443G>C:p.Cys148Ser) em um paciente do sexo masculino que nasceu prematuro (30,5 semanas) com PN de 1000g (Z=-2,4) e CN de 33cm (Z=-4,5). Os pais do paciente eram consanguíneos primos de 2º grau. O paciente foi avaliado ainda bebê pelo nosso ambulatório com 8 meses de idade com uma baixa estatura grave (Z=-8.2). Além disso possuía como fenótipo coartação da aorta já corrigida, fronte ampla, orelhas baixo implantadas com rotação incompleta, base nasal alargada, lábio superior fino, sinofris discreta, fosseta

sacra, déficit auditivo e alguns defeitos ósseos como desmineralização óssea difusa, desproporção craniofacial, deformidade em taça das extremidades distais do rádio, da ulna, da tíbia e da fíbula.

O gene *PCNA* faz parte do grupo de genes que participam de processos celulares fundamentais, fazendo parte do complexo proteico de replicação do DNA sendo essencial para o correto funcionamento deste processo³⁷. Variante em homozigose neste gene estão associadas a um transtorno semelhante à ataxia-telangiectasia, caracterizada por um fenótipo neurodegenerativo com atraso no desenvolvimento, ataxia, perda auditiva neurosensorial e baixa estatura⁸⁴. Ao longo do acompanhamento, no ano de 2019, o paciente com 4 anos iniciou quadros de ataxia e movimentos involuntários estando em avaliação pela neurologia. Além disso, foi realizado contato com o grupo de neurologia do HC da FMUSP, sob a chefia do Prof. Fernando Kok, e foi identificado dois outros pacientes brasileiros com quadro neurológico típico com a mesma variante em homozigose. Em paralelo iniciamos uma colaboração com o pesquisador Dr. Brian Kelch da Universidade de Massachusetts, USA, para realizar estudos funcionais da variante em fibroblastos do nosso paciente e controles. Até o momento não temos um resultado final comprovando a patogenicidade da variante. Mas de acordo com a frequência e compatibilidade fenotípica, a variante tem potencial de ser comprovada como patogênica, por enquanto a classificação é como VUS.

6 DISCUSSÃO

A baixa estatura sindrômica é uma condição que causa apreensão na família, por poder estar associada a condições graves, como predisposição a tumores, retardo de desenvolvimento neuropsicomotor, entre outras. Estabelecer um diagnóstico preciso para estas crianças é de fundamental importância para seguir um prognóstico correto, tomar decisões terapêuticas e prover o aconselhamento genético preciso para família e para o paciente^{5,79}. Entretanto não existe um consenso sobre a investigação molecular dessa condição, e os pacientes são muitas vezes submetidos a diversos tipos de exames, geralmente sem resultados positivos, o que leva a um elevado custo financeiro e emocional na busca do diagnóstico etiológico. Nas últimas duas décadas, o avanço da tecnologia permitiu a melhor compreensão dos mecanismos genéticos associados ao crescimento, com um número expressivo de novas causas monogênicas de baixa estatura^{37,38}. Neste projeto nós propusemos utilizar o sequenciamento do exoma para investigação molecular dos distúrbios de crescimento sindrômico em pacientes sem diagnóstico clínico estabelecido ou sem diagnóstico molecular anterior pela abordagem gene candidato. Foi possível estabelecer o diagnóstico em 47,7% da casuística (41 casos elucidados de 86 investigados), o que mostra uma taxa na faixa superior comparado aos estudos que avaliaram baixa estatura sindrômica (34%-57,5%)^{35,67}. Esta taxa de diagnóstico é 3,2 vezes maior do que nossas recentes descobertas em crianças nascidas pequenas para idade gestacional com baixa estatura isolada (15% vs. 47,7%)⁶⁸. Esse resultado corrobora estudos anteriores

que sugerem taxas mais altas de diagnóstico positivo entre pacientes com deficiência intelectual, malformação grave, dismorfismos faciais, displasia esquelética, problemas familiares e casos com consanguinidade dos pais^{36,85-87}.

Dos pacientes onde a causa genética foi encontrada, a maioria era do sexo masculino (n=24;58,5%), 27 (50%) apresentavam retardo de desenvolvimento neuropsicomotor e 14 (n=46,6%) apresentavam microcefalia. Dez eram nascidos PIGs, a minoria era prematuro (n=10; 32,2%) e a maioria apresentava alguma mal formação cardíaca (n=10; 66,7%) (Tabela 3). As características clínicas desses pacientes encontram-se na Tabela 7. E os heredogramas com segregação das variantes em familiares no Anexo V.

Os genes cujas variantes deletérias foram encontradas pelo sequenciamento do exoma, podem ser classificados em grupos, de acordo com a via molecular a qual faz parte. A maioria das síndromes encontradas apresentam defeitos moleculares em genes fundamentais de processos celulares, tais como replicação celular (*ACTB*, *DYRK1A*, *GINS1*, *KIF11*, *PCNA*, *PCNT*, *POC1A*, *POLG1*, *TERT*), transcrição (*AFF4*, *ANKRD11*, *BCL11B*, *CHD7*, *DPH1*, *HDAC8*, *KDM5C*, *KMT2A*, *RPL5*, *SETD5*, *SMARCA4*, *SRCAP*), reparo do DNA (*BRCA1*), diferenciação celular (*FOXP1*), biossíntese (*MVK*), regulação da expressão (*SMARCA4*) e maquinaria de *splicing* (*PUF60*). A maioria dos pacientes com defeitos nestes genes apresenta como fenótipo retardo de desenvolvimento psicomotor e/ou dificuldades intelectuais (19 de 23). Encontramos variantes em genes que podem ser classificados no grupo de vias de sinalização intracelular (*INPP5K*, *KRAS*, *PDHA1*, *RPS6KA3*, *SCN1A*, *SCUBE3*, *STAT5B*). Variantes em genes que participam da formação da matriz

extracelular da cartilagem de crescimento (*COL2A1*, *FBN1*) foram identificadas em dois pacientes que apresentam desproporção corpórea e alterações esqueléticas inespecíficas. Foi identificada também uma variante no gene *IGF1R* que faz parte da via GH-IGFs. Além de 4 CNVs que se enquadram no grupo das síndromes de genes contíguos.

6.1 Raridade e Sobreposição fenotípica

Grande parte destes genes já estão associados a síndromes previamente descritas (Tabela 7), porém por serem muito raras e/ou apresentarem características que se sobrepõe com outras síndromes o diagnóstico clínico não foi possível. Adicionalmente, alguns casos compuseram as primeiras casuísticas de uma nova condição, anteriormente não descrita.

Tabela 7 - Características clínicas e síndromes associadas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas

ID	Gene	Sexo	Z-escore altura	Atraso do DNPM	Microcefalia	Características adicionais	Doença (MIM#)
1	ACTB	F	-2,8	sim	não	Sinófrisis, raiz nasal rasa, boca com ângulo voltado para baixo, ângulo externo do olho voltado para cima	Síndrome de Baraitser-Winter (243310)
2	AFF4	F	-5,4	sim	sim	Face plana e arredondada, sobrancelha arqueada e unida, hipertelorismo, raiz nasal rasa, tórax com mamilo acessório a esquerda, obesidade, hirsutismo, xerose	Síndrome de CHOPS (616368)
3	ANKRD11	M	-3,4	sim	sim	Estrabismo, hérnia umbilical, convulsão, hiperatividade, polidactilia, Legg-Calves-Perthes	Síndrome de KBG (148050)
4	BCL11B	M	-2,4	sim	não	Distfunção de células betas	Déficit de desenvolvimento intelectual com fácies dismórfica, atraso na fala e anormalidades de células T (618092)
5	BCL11B	F	-2,2	sim	sim	Palato ogival, orelhas de baixa implantação, sobrancelhas ralas, epicando, ptose palpebral corrigida, fenda palpebral pequena, hipermetropia e astigmatismo de alto grau (7,4), base nasal achatada, filtro longo, lábio superior fino, atopia, histórico de infecções respiratórias, hipotireoidismo	Déficit de desenvolvimento intelectual com fácies dismórfica, atraso na fala e anormalidades de células T (618092)
6	BRCA1	F	-5,5	sim	sim	Manchas hiperocrômicas, lábio superior cheio, clinodactilia, anemia de fanconi	Anemia de Fanconi-grupo S complementar (617883)
7	CHD7	M	-2,7	não	não	Epicanito, agenesia de nervo auditivo, Tetralogia de Fallot	Síndrome de CHARGE (214800)
8	COL2A1	F	-3,9	não	não	Displasia esquelética, desproporção corpórea	Displasia espondiloepifiseal (271700)
9	COL2A1	M	-2,9	não	não	Osteonecrose de epífise de cabeça de fêmur, redução da altura dos corpos vertebrais com irregularidade da superfície dos platôs	Doença de Legg-Calve- Perthes (150600)
10	COL2A1	M	-2,9	não	não	Pé torto congênito, escoliose grave, ectasia e coarctação de aorta	Displasia Czech (609162)

continua

Tabela 7 - Características clínicas e síndromes associadas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas - continuação

ID	Gene	Sexo	Z-score altura	Atraso do DNPM	Microcefalia	Características adicionais	Doença (MIM#)
11	DPH1	F	-4,5	sim	não	Ptose palpebral, epicantero, hipertelorismo, miopia, fronte ampla, rarefação de sobrancelhas, dentes pequenos na arcada inferior, palato ogival, cabelos finos e esparsos, narina antevertida, base nasal larga, baixa impantação das orelhas, alteração ungueal, malformação cardíaca	Retardo de neurodesenvolvimento com baixa estatura, fâcies distintas e cabelos esparsos (616901)
12	DYRK1A	F	-3,1	sim	não	Pectus excavatum, clinodactilia, hiperlordose lombar, episódio de crise convulsiva febril	Retardo mental (614104)
13	FBN1	M	-3,9	não	sim	Baixa estatura desproporcional, braquidactilia, genu varo, voz metálica	Displasia Acromioclavicular (102370)
14	FOXP1	M	-2,6*	sim	não	Hipertelorismo aparente, sobrancelhas rarefeitas e arqueadas, palato ogival, raiz nasal curta, hipoplasia de aleta nasal, sinófrisis, sindactilia, Clinodactilia, braquidactilia, fosseta pré-sacral, criptorquidia bilateral, fimose, hipospádia	Retardo mental com deficiência de linguagem e com ou sem características autistas (613670)
15	GINS1 GINS1	F	-3,4	sim	sim	Infecções de repetição respiratórias, úvula bifida, hipogonadismo hipergonadotrófico, agenesia mamária	Imunodeficiência 55 (617827)
16	HDAC8	F	-2,0	sim	não	Olhos fundos, epicantero, astigmatismo, retrognatía, nariz bulboso, obesidade truncal, acantose nigricans nas axilas, IGF1 elevado (1156 µg/L)	Síndrome Cornélica de Lange (300882)
17	IGF1R	M	-2,3	sim	não	Fronte proeminente, orelhas pouco antevertidas, epicantero, hipertelorismo, ptose discreta, ponte nasal rebaixada, boca em arco de cupido, telangiectasia em face, prega palmar única direita, clinodactilia, obesidade	Resistência a IGF-I (270450)
18	INPP5K	M	-4,1	sim	sim	Dentinogênese imperfeita, Catarata congênita, alterações em vértebras torácicas (hemivértebra e vértebra em H), criptorquidia, micropenis	Distrofia muscular congênita com catarata e deficiência intelectual (617404)

continua

Tabela 7 - Características clínicas e síndromes associadas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas - continuação

ID	Gene	Sexo	Z-score altura	Atraso do DNPM	Microcefalia	Características adicionais	Doença (MIM#)
19	KDM5C	M	-3,6	sim	não	Face plana, Pavilhão auricular de abano, base do nariz rasa, dedos afilados, mamilo antevertido	Retardo mental ligado ao X, síndrome, tipo Claes-Jensen (300534)
20	KIF11	M	-2,3	sim	sim	Estrabismo convergente, pavilhão auricular proptótico, crises convulsivas, mal formações cerebrais	Microcefalia com ou sem coriorretinopatia, linfedema ou retardo mental (152950)
21	KMT2A	M	-3,0	sim	sim	Microftalmia, retrognatia, pavilhão auricular pequeno, criptorquidia, mamilo desviado, cifose torácica	Síndrome Wiedemann-Steiner (605130)
22	KRAS	F	-4,6	sim	sim	Pavilhão auricular de baixa implantação, mal formação cardíaca	Síndrome de Noonan 3 (609942)
23	MVK MVK	M	-3,1	sim	não	Orelha reduzida e espessada, sinofris, base nasal alargada, epicaneto, clinodactilia	Aciduria mevalonica (610377)
24	PCNT	F	-6,2	sim	sim	Fechamento prematuro das fontanelas, fronte curta, nariz proeminente, ponte nasal rasa, face plana, fendas palpebrais para cima, orelhas pequenas, clinodactilia, CIA	Nanismo primordial osteodisplásico microcefálico, tipo II (210720)
25	PDHA1	F	-3,5	sim	sim	Estrabismo, sobrelanceiras largas, ponte nasal deprimida, lábios grossos, dentes acavalados, orelhas displásicas, presença de sardas no rosto, face triangular, clinodactilia	Deficiência de piruvato desidrogenase E1-alfa (312170)
26	POC1A	M	-5,1	sim	sim	Epicaneto, sobrelanceiras pouco arqueadas, hipermetropia e astigmatismo, palato ogival, dentes grandes e afastados, prognatismo, face triangular, magreza acentuada, escoliose, dedos curtos	Baixa estatura, oncodisplasia, dismorfismo facial e hipotricose (614813)
27	POLG	F	-4,3	sim	não	Hipotonia na infância, obesidade, puberdade precoce e déficit intelectual importante	Síndrome de depleção de DNA mitocondrial 4B -tipo MNGIE (613662)
28	PUF60	M	-2,2	sim	não	Encurtamento do ombro esquerdo, mal formação cardíaca	Síndrome de Verheij (615583)

continua

Tabela 7 - Características clínicas e síndromes associadas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas - continuação

ID	Gene	Sexo	Z-escore altura	Atraso do DNPM	Microcefalia	Características adicionais	Doença (MIM#)
29	RPL5	M	-2,7	não	não	Epicanto, ângulo olho esquerdo voltado para baixo, assimetria facial, pectus <i>carinatum</i> , palato ogival, retromicrognatia, anemia congênita grave	Anemia Diamond-Blackfan (612561)
30	RPS6K43	M	-2,7	sim	sim	Ângulo externo do olho voltado para baixo, lábios grossos, macrodontia, palato ogival, clinodactilia, teleangectasias em face e abdomen, escoliose, luxação de quadril, hipotonia ao nascimento	Síndrome Coffin-Lowry (303600)
31	SCN1A	M	-2,6	sim	não	Sinófris, problemas dentários	Síndrome Dravet (607208)
32	SCUBE3	M	-3,8	não	sim	Craniossinostose, micrognatia, palato ogival, epicanto, 4° e 5° metacarpos curtos, não consegue flexionar as falanges médias, surdez, fez gastrostomia	Nova síndrome (Não MIM)
33	SETD5	F	-2,5	sim	não	Face triangular, ombrancelha arqueada e sinófris, baixa implantação do pavilhão auricular, pescoço curto	Retardo mental, autossômico dominante 23 (615761)
34	SMARCA4	M	-3,7	sim	não	Olhos amendoados, pescoço alargado e curto, palato ogival, ombrancelha arqueada, cílios longos, baixa implantação de cabelos e orelhas, orelhas anteriorizadas e simplificada, criptorquidismo bilateral, hernia inguinal, hernia diafragmática, acidose tubular renal proximal tipo 2, pectus <i>carinatum</i> , escoliose pés com hipoplasia ungueal, desproporção dos dedos dos pés, clinodactilia, pés planos	Síndrome Coffin-Siris (614609)
35	SRCAP	M	-2,4	sim	não	Defeito de linha média - lábio leporino, fenda palatina corrigidos, clinodactilia, epípfises em forma de cone nas falanges, crise convulsiva	Síndrome (136140)
36	SRCAP	F	-2,7	sim	sim	Filtro nasal curto, columela, hernia umbilical 5° dedo curto, hiperinsulinemia, IGF-I elevado	Síndrome (136140)
37	STAT5B	F	-6,1	não	não	Ictiose, dermatite atópica, hipotireoidismo de Hashimoto	Insensibilidade ao hormônio do crescimento com desregulação imunológica 1 (245590)

continua

Tabela 7 - Características clínicas e síndromes associadas dos 41 pacientes com variantes Patogênicas ou Provavelmente Patogênicas - continuação

ID	Gene	Sexo	Z-score altura	Atraso do DNPM	Microcefalia	Características adicionais	Doença (MIM#)
38	TERT	M	-4,0		sim	Sinofris, baixa implantação de cabelos, aplasia de medula óssea constitucional, hipertrícore, pele seca, leucoplasia, mal formações cerebrais	Disqueratose congênita (613989)
39	dup11p15.5-p15.4	F	-2,2	não	não	Olhos amendoados, palato ogival, diastema, dentes acavalados, base nasal achatada, narina antevertida, 4º e 5º dedo dos pés curtos, olites médias de repetição	Síndrome de Silver Russel (180860)
40	dup5q35.2-q35.3	M	-3,1	sim	sim	Valva aórtica bivalvar	Síndrome de Baixa estatura, microcefalia e retardo de desenvolvimento (não OMIM)
41	dup8q24.22-q24.3 del15q26.2-q26.3	M	-3,5	sim	ND	Fronte proeminente, cílios longos, hipertelorismo, epicanto, sinofris, palato ogival, filtro curto, polidactilia, clinodactilia, hipertrícore, defeitos ósseos, idade óssea atrasada	Resistência a IGF-I (270450)

M: Masculino; F: Feminino; ND: dado não disponível; *Baixa estatura em relação ao padrão familiar; dup: Duplicação; del: Deleção

conclusão

A variante c.772C>T:p.Arg258Trp no gene *AFF4* na paciente ID2 é um exemplo de sobreposição de fenótipos entre síndromes que dificultam o diagnóstico clínico. A paciente teve como suspeita clínica a Síndrome de Cornélia de Lange, a qual apresenta diversas características que se sobrepõe à Síndrome de CHOPS, como baixa estatura, microcefalia, pescoço curto, cílios longos, sobrancelhas unidas, hirsutismo, raiz nasal deprimida, sopro cardíaco, atraso do DNPM, obesidade, dentre outras características que também estavam presentes na paciente. Esta variante já estava descrita na literatura como deletéria, associada a síndrome de CHOPS e então foi considerada patogênica⁸⁸.

Outro exemplo que leva a dificuldade de diagnosticar clinicamente é do gene *BRCA1*. Inicialmente não era associado a nenhum fenótipo com variantes em homozigose neste gene. Após a identificação da nossa paciente foi possível rever as características clínicas de 2 outros casos da literatura e uma publicação derivada desta análise permitiu delinear que alterações em homozigose no *BRCA1* estão associados a Anemia de Fanconi Grupo S complementar, que nossa descrição de caso ajudou a estabelecer no MIM, pois apresenta muitas das características da síndrome de Anemia de Fanconi, como baixa estatura, cardiopatias congênitas, microcefalia dentre outras⁸⁰.

É sabido que variantes em heterozigose no *BRCA1*, causam câncer de mama e ovário. E até o momento deste estudo, nenhuma variante em homozigose tinha sido descrita nesse gene, apenas em heterozigose composta. As duas pacientes destes estudos prévios apresentavam fenótipo semelhantes à da nossa paciente, porém o diagnóstico foi estabelecido na idade de adulto

jovem por terem apresentado câncer ^{89,90}. A nossa paciente na qual foi descoberta a variante foi apresentada em nossa clínica com 2,5 anos de idade com uma baixa estatura grave, microcefalia, atraso no desenvolvimento neuropsicomotor, uma cardiopatia congênita e características faciais dismórficas. Na sua história familiar encontramos que seus pais eram primos de terceiro grau. A variante em homozigose no *BRCA1* (c.2709T > A; p. Cys903*) causa um ganho de stop códon prematuro, levando a perda de domínios importantes da proteína e consequentemente anulando sua função. A doença caracterizada pela mutação em homozigose no *BRCA1*, foi então associada a um novo fenótipo.

Retrospectivamente, analisando os pais da paciente, confirmamos que cada alelo da mutação encontrada foi herdado do pai e da mãe. Sabendo da importância desse achado, a mãe da paciente foi submetida à análise clínica, e foi descoberto um linfonodo maligno na região axilar próximo a mama esquerda. Ela foi encaminhada ao tratamento e hoje segue acompanhando rotineiramente com médicos para observar o desenvolvimento de mais tumores. Assim como toda a família por parte de pai e mãe foi orientada a procurar acompanhamento médico para rastrear o câncer.

Em relação a variante no gene *SRCAP* aqui descrita, que leva a uma perda de função no exon 34, já foi reportada como a variante mais comum encontrada nesta síndrome, foi encontrada em 2 pacientes e publicada em um artigo que relata as características clínicas de todos os pacientes com a Síndrome de Floating-Harbor avaliados em nosso ambulatório⁷⁷. O diagnóstico desta condição fora de centros especializados é dificultado por sua raridade e

por poder apresentar sobreposição com outros distúrbios de crescimento sindrômico. Por esses motivos, a análise de sequenciamento do exoma usado foi essencial para o diagnóstico genético destes 2 casos. Um destes pacientes com SFH (ID35) diagnosticado por WES apresentava um defeito de linha média, hipotireoidismo central e deficiência de GH (DGH), o que impediu o reconhecimento da SFH. A associação de SFH e DGH isolada foi relatado em 2 casos adicionais na literatura^{91,92} e, semelhante ao nosso paciente, nenhum deles apresentava defeito estrutural hipotalâmico-hipofisário. Outro estudo sugeriu a presença de disfunção neurosecretora de GH entre os pacientes com SFH devido aos baixos valores séricos de IGF-1, resposta variável ao GH durante testes provocativos, e uma boa resposta ao tratamento com GH⁹³. Contudo, esse fenótipo bioquímico claramente não reflete todos os pacientes com SFH que parecem ter uma variedade de mecanismos envolvidos em seu distúrbio de crescimento. Na nossa casuística apenas este paciente de um total de sete teve o diagnóstico de DGH.

O outro paciente com SFH (ID36) apresentava baixa estatura com microcefalia e concentrações elevadas de IGF-1, sendo inicialmente diagnosticado como resistência para este hormônio. Por esta razão ele foi inicialmente selecionado para sequenciamento do *IGF1R* que não identificou alterações. Posteriormente, em nosso estudo, observou-se que outros pacientes com SFH apresentaram valores basais elevados de IGF-1 ou no limite superior da normalidade. De fato, 4 de 7 de nossos pacientes com SFH preencheram os critérios clínicos observados em pacientes com insensibilidade parcial ao IGF-1 e que direcionam o estudo genético para a análise de *IGF1R*⁹⁴. Desta forma, nosso resultado indica que a SFH deva ser um diagnóstico diferencial nos

pacientes com suspeita de insensibilidade ao IGF-1 e sugere-se deva haver um comprometimento da via de sinalização do IGF-1 nesses pacientes⁹⁵.

6.2 Estabelecimento de nova relação Genótipo-Fenótipo

Ao identificar variantes raras potencialmente deletérias em genes ainda sem descrição de fenótipo associado em humanos e compartilhar estes dados em plataformas colaborativas como o *GeneMatcher* permitiu caracterizar novas síndromes.

Por esta estratégia participamos da descrição de um novo transtorno do desenvolvimento não reconhecido anteriormente causado por defeitos bialélicos no *SCUBE3* (paciente ID32). Neste caso, foi relatado que as variantes bialélicas inativadoras no *SCUBE3* causam um transtorno de desenvolvimento não reconhecido anteriormente com crescimento reduzido (com escore-z da altura variando entre -4,9 e -2,0), defeitos esqueléticos, anomalias dentárias e uma aparência craniofacial distinta como características principais⁷⁸. Neste artigo foram descritas oito famílias consanguíneas não relacionadas, com um ou mais indivíduos afetados por uma doença não reconhecida e compartilhando variantes possivelmente patogênicas em homozigose no *SCUBE3*, incluindo variantes *stopgains*, *frameshift*, *missense*, alterações em sítio de splicing e um rearranjo intragênico complexo. No caso da variante c.611G>A;p.Gly204Asp encontrada no nosso paciente, modelos *in silico* sugerem que ela prejudica o dobramento adequado da proteína, impactando assim a interação com os fatores de crescimento necessários⁷⁸.

Genes da família SCUBE codificam glicoproteínas ancoradas na superfície celular funcionando como receptores auxiliares para modular a sinalização através de uma variedade de fatores de crescimento. Estudos funcionais realizados pelos grupos de pesquisadores envolvidos no artigo, mostraram que o *SCUBE3* especificamente é altamente expresso em osteoblastos, ossos longos e regiões alta concentração cartilaginosa contribuindo para o ajuste da formação de osso endocondral dependente de BMP, regulando a diferenciação condrogênica, a proliferação de condrócitos e a diferenciação osteogênica. Sendo assim Scube3 é necessário para o desenvolvimento e crescimento esquelético adequado⁷⁸.

Além deste caso positivo, ainda estamos em colaboração com o artigo do *TBC1D32* e em fase inicial de conversas em relação ao *CDK4*, podendo futuramente estabelecer mais 2 novos genótipos-fenótipos. Mesmo porque nossos experimentos mostraram que há alteração no cDNA do paciente, podendo então ter alguma alteração na função da proteína. Mas para esta afirmação precisamos de estudos adicionais.

O sucesso destes casos mostra a importância do compartilhamento de dados entre pesquisadores para acelerar e contribuir para a compressão de mecanismos genéticos responsáveis por doenças até então sem causa conhecida.

6.3 Limitações do WES e análise de CNVs

A ampla aplicação do sequenciamento paralelo em larga escala para a detecção de variantes genéticas vem mostrando as diferenças genéticas entre os indivíduos e solucionado casos de doenças com maior eficiência e rapidez, além do menor custo comparado a abordagem tradicional. Porém este tipo de abordagem ainda apresenta suas limitações. Com a utilização do WES, é possível apenas analisar as variantes nas regiões codificadoras de proteínas dos genes e em regiões próximas de *splicing*. Solucionando até aproximadamente 57% dos casos estudados, o que significa que em resultados negativos, podemos ter modificações epigenéticas, mosaicismos somáticos, rearranjos cromossômicos não identificados e variações nas regiões não analisadas pelo WES, como as regiões intrônicas e reguladoras⁶⁷.

Em relação as variantes estruturais, como as variantes em número de cópias, translocações, inversões e rearranjos complexos, a identificação é limitada devido a maioria das técnicas de SPLE utilizar *reads* de leitura pequenos (150-300 pb) e requer inferência computacional em algoritmos que não são muito eficientes até o presente momento. Além disso, estas variantes podem estar localizadas em regiões de baixa complexibilidade com sequências de repetição do DNA que não tem boa leitura e mapeamento⁹⁶.

Apesar destas limitações, novos modelos de análise de CNVs, como ClinCNV⁷³ e DELLY⁷⁴, vem aperfeiçoando a análise de CNVs a partir dos resultados brutos do WES. Porém para este tipo de análise ser eficiente, é preciso um grande número de amostras processadas pela mesma metodologia

e que apresentem uma cobertura homogênea, já que estes programas são baseados em diferenças de cobertura de determinadas regiões para estabelecer uma CNV.

Utilizando estes 2 programas conseguimos identificar quatro novas CNVs que explicam os fenótipos dos pacientes ID39 (Figura 9a), ID40 (Figura 9b) e ID41 (Figuras 9c e 9d). Como controles positivos na avaliação de CNVs por essa tecnologia utilizamos dados de pacientes com CNVs patogênicas previamente identificadas por outras técnicas. Assim foi possível elevar o diagnóstico positivo de 44,2% (38 de 86 pacientes) para 47,7% (41 de 86 pacientes).

Figura 9. a) Visualização das duplicações chr11 (p15.5-p15.4):2290980-3845365x3. b) chr5 (q35.2-q35.3):175394113-177422934x3. c) chr8 (q24.22-q24.3):133108784-146294098x3 e d) chr15 (q26.2-q26.3):97439378-102383473x1 no IGV, indicadas por seta vermelha.

6.3.1 Duplicação chr11 (p15.5-p15.4):2290980-3845365x3

A duplicação no chr11 (p15.5-p15.4):2290980-3845365x3, envolvendo 28 genes (Figura 9a), foi encontrada em uma paciente do sexo feminino (ID39), que apresenta histórico de baixa estatura familiar, com irmã e primos maternos com fenótipo semelhante de baixa estatura associado a oligodrâminio durante a gestação, parto prematuro e baixa estatura pós-natal.

A CNV envolve a região de *imprinting* associada a Síndrome de Silver Russell com CNVs que sobrepõe a região já descritas. Sabe-se que o gene *CDKN1C*, localizado na região centromérica do cromossomo 15, codifica a proteína p57(KIP2), que age como um regulador negativo da proliferação celular, o qual sofre *imprinting* paterno, com consequente expressão preferencial do alelo materno^{97,98}. É sabido que duplicações maternas da região centromérica do cromossomo 15 são causa da Síndrome de Silver-Russel⁹⁸. Para confirmar a herança materna da CNV, foi realizado o MLPA ME-030 e a duplicação foi confirmada como herdada da mãe (Figura 10a). O MLPA do pai da paciente foi normal (Figura 10b). A duplicação na paciente foi confirmada por aCGH (Figura 11a).

Figura 10. MLPA ME030 para BWS/SRS realizado nos pais da paciente. **a)** MLPA da mãe - acima podemos ver a duplicação da região correspondente aos genes *KCNQ1* e *CDKN1C*. E abaixo podemos ver a hipometilação da região do cromossomo 11. **b)** MLPA do pai, normal.

6.3.2 Duplicação chr5 (q35.2-q35.3):175394113-177422934x3

A duplicação chr5 (q35.2-q35.3):175394113-177422934x3, envolvendo 41 genes (Figura 9b), nunca foi descrita na literatura e foi encontrada em um paciente do sexo masculino, filho de pais com altura normal. É descrito que microduplicações envolvendo o gene *NSD1*, compreendido nesta CNV, são responsáveis por um fenótipo síndrome conhecido como “Sotos reverso”, consistente com baixa estatura, microcefalia, retardo de DNPM e características faciais distintas^{99,100}, compatível com o fenótipo paciente ID40 (Tabela 7). Para confirmar a CNV como *de novo* no paciente, foi realizado o aCGH pelo laboratório de biociências da USP (Figura 11b).

6.3.3 Duplicação chr8 (q24.22-q24.3):133108784-146294098)x3 e Deleção chr15 (q26.2-q26.3):97439378-102383473)x1

As CNVs aqui descritas foram encontradas em um paciente (ID41) do sexo masculino, nascido PIG de comprimento e peso, filho de pais de altura normal (Tabela 7). A duplicação no cromossomo 8 que envolve 112 genes (Figura 9c), dentre os quais encontram-se alguns genes já associados a síndromes com baixa estatura sindrômica, como *PUF60* e *RECQL4*. A região da duplicação envolve 80% ou mais do tamanho de duplicações já descritas como patogênicas no DECIPHER.

Já a deleção no cromossomo 15 (Figura 9d), compreende o gene *IGF1R*, onde sabe-se que mutações inativadoras e deleções em heterozigose estão associadas a déficit de crescimento pré e pós-natal pela resistência aos IGFs^{46,101}. Portanto a combinação das 2 CNVs foi considerada patogênica para o fenótipo do paciente. Ambas as CNVs foram confirmadas por aCGH (Figuras 11c e 11d).

Figura 11. Confirmações das 4 CNVs encontradas em 3 pacientes por aCGH. **a)** Duplicação chr11 (p15.5-p15.4):2290980-3845365x3. **b)** Duplicação chr5 (q35.2-q35.3):175394113-177422934x3. **c)** chr8 (q24.22-q24.3):133108784-146294098x3. **d)** Deleção chr15 (q26.2-q26.3):97439378-102383473x1.

Estudos recentes mostram que a taxa de detecção de variantes por WES pode ser até 8,3x maior do que por aCGH, sugerindo que o WES deve ser a primeira abordagem genômica a ser realizada em pacientes com suspeita de doença genética¹⁰². Em nosso estudo o WES foi eficaz em detectar CNVs em 5,3% dos pacientes que foram submetidos a análise deste tipo de variantes através de dados de WES. Porém a taxa de positividade geral é até 3x maior do que quando o aCGH é utilizado como primeira abordagem (47,7% vs ~15%) na casuística de pacientes com baixa estatura^{16,17}.

A eficiência em detectar CNVs por WES vem sendo aperfeiçoada, mas ainda assim o WES ainda é limitado em identificar variantes de origem

epigenética, mosaicismos somáticos e outras variantes estruturais, podendo estes defeitos explicarem uma parcela dos pacientes sem diagnóstico estabelecido por WES.

No nosso estudo 52,3% dos casos analisados ficaram sem diagnóstico. O avanço de conhecimento em genética clínica e o aperfeiçoamento das técnicas de identificação genômica são essenciais para diminuirmos este número de pacientes sem causa definida. Técnicas como sequenciamento completo do genoma por long-reads¹⁰³, o RNA-seq¹⁰⁴ para avaliação de expressão gênica e o sequenciamento de bissulfito em escala genômica para identificação de padrões de metilação¹⁰⁵, são técnicas que veem sendo utilizadas em pesquisa, mas ainda são muito caras e precisam de mais estudos em grandes coortes para melhor avaliação de taxas diagnósticas.

Estes aperfeiçoamentos dos estudos genômicos vão permitir ao longo do tempo um melhor entendimento da função dos genes e assim possibilitar novas associações genótipo-fenótipo, como nos casos do *SCUBE3* e possivelmente do *TBC1D32* e *CDK4*, beneficiando mais pacientes.

Acreditamos que encontrar resultados genéticos pode economizar tempo ao diminuir a ampla gama de exames e avaliações médicas pelas quais os pacientes são submetidos. Além disso, pode-se fornecer informações precisas para o aconselhamento familiar e tratamento dos pacientes.

Pela alta taxa de positividade em encontrar a causa genética, sugerimos a utilização do WES como primeira abordagem genético-molecular a ser utilizada na investigação de pacientes com baixa estatura sindrômica com suspeita de doença genética (Figura 12).

Figura 12. Sugestão de abordagem molecular para pacientes com baixa estatura sindrômica.

8 CONCLUSÃO

Pacientes com baixa estatura sindrômica, possuem alta probabilidade de apresentar uma etiologia monogênica que justifique seu quadro clínico. A técnica de sequenciamento paralelo em larga escala de todos os exons foi eficaz em estabelecer o diagnóstico etiológico do distúrbio de crescimento sindrômico em 47,7% dos pacientes estudados mesmo na ausência de um gene candidato.

Houve uma grande heterogeneidade nos genes onde encontramos variantes deletérias, porém a maioria deles se enquadra nos genes que participam de processos celulares fundamentais, mostrando a importância destas vias para o funcionamento do organismo.

O compartilhamento de genes candidatos em plataformas colaborativas permite obter evidências que reforçam ou refutam a associação destes com o fenótipo investigado, permitindo a caracterização de novas síndromes.

O sequenciamento do exoma é o exame de escolha na investigação de casos de baixa estatura sindrômica sem diagnóstico clínico, e com o aperfeiçoamento da análise de variantes de número de cópias baseado em dados gerados pelo WES, é possível também identificar CNVs com sensibilidade semelhante a análise cromossômica por microarray, como aCGH.

8 ANEXOS

Anexo A

Protocolo para avaliação inicial de paciente com baixa estatura.

Paciente: _____ RGHC: _____

Data de nascimento: ____ / ____ / ____

Sexo: M () F ()

Altura da Mãe (cm) _____

Altura da Pai (cm) _____

Idade gestacional (sem) _____

APGAR: ____/____

Peso ao nascimento (g) _____ Comprimento ao nascimento (cm) _____

Perim Cef (cm) _____

M: masculino; F: feminino

Complicações na gestação e/ou parto:

() – Alcoolismo ou uso de drogas na gestação:

() – Infecções na gestação:

Data da avaliação: ____ / ____ / ____

Altura (cm): _____ Altura sentada (cm): _____ Envergadura (cm): _____

Peso (kg): _____ Perímetro Cefálico (cm) _____

Triagem para avaliação de quadro sindrômico

() – Consanguinidade: _____

- () – Outra(s) mal formações: _____
- () – Dismorfismos faciais: _____
- () – Microcefalia: _____
- () – Deformidade esquelética: _____
- () – Desproporção corpórea: _____
- () – Atraso no DNPM: _____
- () – Retardo mental ou autismo: _____

Anexo B

Critérios de avaliação utilizados para calcular Escore-Z de acordo com as características informadas.

Escore-Z	Critérios utilizados
Comprimento e peso ao nascimento	Usher and McLean, 1969 ¹
Altura, peso e índice de massa corpórea (IMC)	Curva de referência americana ² , validado para a população brasileira ³
Altura sentada	Fredriks AM et al., 2005 ⁴
Perímetro cefálico	Curva de referência americana ² , validado para a população brasileira ³

1. Usher R, McLean F. Intrauterine growth of live-born Caucasian infants at sea level: standards obtained from measurements in 7 dimensions of infants born between 25 and 44 weeks of gestation. *J Pediatr*, 1969;74(6):901-10.
2. Kuczmarski RJ, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat* 11, 2002(246):1-190.
3. Silva DA, et al. Comparison between the growth of Brazilian children and adolescents and the reference growth charts: data from a Brazilian project. *J Pediatr (Rio J)*, 2010;86(2):115-20.
4. Fredriks AM, van Buuren S, van Heel WJ, Dijkman-Neerincx RH, Verloove-Vanhorick SP, Wit JM. Nationwide age references for sitting height, leg length, and sitting height/height ratio, and their diagnostic value for disproportionate growth disorders. *Arch Dis Child*. 2005;90(8):807-12. doi:10.1136/adc.2004.050799

Anexo C

Resumo das linhas de evidência utilizadas na classificação pelo guia ACMG-AMP.

	Benigna (B)		Patogênica (P)			
	Forte (S)	Pequena (P)	Pequena (P)	Forte (S)	Moderada (M)	Muito Forte (VS)
Dados populacionais	MAF incompatível com a doença BS1/BA1			Ausente em bancos de dados PM2	Maior prevalência em afetados PS4	
Predição <i>In silico</i>		Sugerem ausência de impacto BP4	Sugerem patogenicidade PP3	Mudança de tamanho da proteína PM4 Nova mudança de aminoácido na mesma posição de uma variante já descrita PM5	Mesma mudança de aminoácido já descrita como patogênica PS5	LOF em gene onde este tipo de variante é deletéria PVS1
Dados funcionais	Estudos mostram ausência de impacto BS3		Não sinônima onde este tipo de variante é conhecida como deletéria PP2		Estudos mostram patogenicidade de PS3	
Compatibilidade com fenótipo			Compatível com fenótipo PP4			
Segregação familiar	Não-segrega BS4		Segrega com a família PP1			

Anexo D

Critérios utilizados para classificação final de patogenicidade segundo o guia ACMG-AMP

Patogênica	1 PVS1 + 1 PS <i>ou</i> 1 PVS1 + 2 PM <i>ou</i> 1 PVS1 + 1 PM + 1 PP <i>ou</i> 12 PS <i>ou</i> 1 PS + 3 PM <i>ou</i> 1 PS + 2 PM + 2 PP <i>ou</i> 1 PS + 1 PM + 4 PP
Provavelmente patogênica	1 PVS + 1 PM <i>ou</i> 1 PS + 1-2 PM <i>ou</i> 1 PS + 2 PP <i>ou</i> 3 PM <i>ou</i> 2 PM + 2 PP <i>ou</i> 1 PM + 4 PP
Benigna	1 BA1 <i>ou</i> 2 BS
Provavelmente benigna	1 BS + 1 BP <i>ou</i> 2 BP
Variante de significado Incerto	Outro critério não mencionado acima

Anexo E

Heredogramas de todas as famílias onde foram encontradas variantes Patogênicas ou Provavelmente Patogênicas

ID17: *IGF1R*:c.2629C>T ID18: *INPP5K*:c.1088T>C ID19: *KDM5C*:c.1510G>A ID20: *KIF11*:c.2402_2403del ID21: *KMT2A*:c.9911delT ID22: *KRAS*:c.65A>G ID23: 1. *MVK*:c.803T>C
2. *MVK*:c.1129G>A ID24: 1. *PCNT*:c.8413C>T
2. *PCNT*:c.9099G>C ID25: *PDHA1*:c.976C>T ID26: *POC1A*:c.275+2T>G ID27: 1. *POLG1*:c.1760C>T
2. *POLG1*:c.752C>T ID28: *PUF60*:c.83_86dupCAGC ID29: *RPL5*:c.742C>T ID30: *RPS6KA3*:c.1764+2T>C ID31: *SCN1A*:c.2400delA ID32: *SCUBE3*:c.611G>A ID33: *SETD5*:c.1852C>T ID34: *SMARCA4*:c.2519_2542del ID35: *SRCAP*:c.7330C>T ID36: *SRCAP*:c.7330C>T ID37: *STAT5B*:c.1960C>T ID38: 1. *TERT*:c.2780T>G
2. *TERT*:c.2911C>T ID39: *chr11*:2290980-3845365x3 ID40: *chr5*:175394113-177422934x3

ID41: chr8:133108784-146294098x3
chr15:97439378-102383473x1

REFERÊNCIAS

1. Şıklar Z, Berberoğlu M. Syndromic Disorders with Short Stature. *Journal of Clin Res Pediat Endocrinol*. 2014;6(1):1-8. doi:10.4274/Jcrpe.1149
2. Łaczmańska I. Short stature in geneticsyndromes: Selectedissues. *Advances in Clinicaland Experimental Medicine*. 2018;27:409–14. doi: 10.17219/acem/67051
3. Mehlman CT, et al. Evaluation of the Child with Short Stature. *Orthop Clin North Am*. 2015. Oct;46(4):523-31. doi: 10.1016/j.ocl.2015.06.006
4. Grunauer M, Jorge AAL. Genetic short stature. *Growth Horm IGF Res*. 2018 Feb;38:29-33. doi: 10.1016/j.ghir.2017.12.003
5. Dauber A. New genetic tools in the diagnosis of growth defects. *Growth Hormone & IGF Research* 38. 2018;24-8, <https://doi.org/10.1016/j.ghir.2017.11.002>.
6. Jee YH, Baron J, Nilsson O. New developments in the genetic diagnosis of short stature. *Curr Opin Pediatr*. 2018;30(4):541-7. doi: 10.1097/MOP.0000000000000653. PMID: 29787394; PMCID: PMC7241654.
7. Caspersson T, et al. Chemical differentiation along metaphase chromosomes. *Exp Cell Res*. 1968;49(1):219-22.
8. Lejeune J, Turpin R, Gautier M. [Mongolism; a chromosomal disease (trisomy).]. *Bull Acad Natl Med*. 1959;143(11-12):256-65.
9. Ford CE, et al. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet*. 1959;1(7075):711-3.
10. Jacobs PA, Strong JA. A case of human intersexuality having a possible XXY sex-determining mechanism. *Nature*. 1959;183(4657):302-3.
11. den Dunnen JT, White SJ. MLPA and MAPH: sensitive detection of deletions and duplications. *Curr Protoc Hum Genet*. 2006;7:Unit 7 14.

12. Funari MF, et al. Cryptic intragenic deletion of the SHOX gene in a family with Leri-Weill dyschondrosteosis detected by Multiplex Ligation-Dependent Probe Amplification (MLPA). *Arq Bras Endocrinol Metabol.* 2008;52(8):1382-7.
13. Funari MF, et al. Usefulness of MLPA in the detection of SHOX deletions. *Eur J Med Genet.* 2010;53(5):234-8.
14. Trarbach EB, et al. Screening of autosomal gene deletions in patients with hypogonadotropic hypogonadism using multiplex ligation-dependent probe amplification: detection of a hemizygosis for the fibroblast growth factor receptor 1. *Clin Endocrinol (Oxf).* 2010;72(3):371-6.
15. Kallioniemi A, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science.* 1992;258(5083):818-21.
16. Canton AP, Costa SS, Rodrigues TC, et al. Genome-wide screening of copy number variants in children born small for gestational age reveals several candidate genes involved in growth pathways. *Eur J Endocrinol.* 2014;171(2):253-262. doi:10.1530/EJE-14-0232
17. Homma, T.K., et al., Recurrent Copy Number Variants Associated with Syndromic Short Stature of Unknown Cause. *Horm Res Paediatr*, 2018; 89(1):13-21.
18. Sanger F. and A.R. Coulson, A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J Mol Biol.* 1975;94(3):441-8.
19. Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. *Science.* 2003;300(5617):286-90.
20. Jorge AA, et al. [Founder effect of E180splice mutation in growth hormone receptor gene (GHR) identified in Brazilian patients with GH insensitivity]. *Arq Bras Endocrinol Metabol.* 2005;49(3):384-9.
21. Jorge AA, et al. The first homozygous mutation (S226I) in the highly-conserved WSXWS-like motif of the GH receptor causing Laron syndrome: suppression of GH secretion by GnRH analogue therapy not restored by dihydrotestosterone administration. *Clin Endocrinol (Oxf).* 2004;60(1):36-40.

22. Scalco RC, et al. Effectiveness of the combined recombinant human growth hormone and gonadotropin-releasing hormone analog therapy in pubertal patients with short stature due to SHOX deficiency. *J Clin Endocrinol Metab.* 2010;95(1):328-32.
23. Jorge AA, et al. SHOX mutations in idiopathic short stature and Leri-Weill dyschondrosteosis: frequency and phenotypic variability. *Clin Endocrinol (Oxf).* 2007;66(1):130-5.
24. Pugliese-Pires PN, et al. A novel STAT5b mutation causing GH insensitivity syndrome associated with hyperprolactinemia and immune dysfunction in two male siblings. *Europ J Endocrinol.* 2010;163:349-55.
25. Ferreira LV, et al., PTPN11 (protein tyrosine phosphatase, nonreceptor type 11) mutations and response to growth hormone therapy in children with Noonan syndrome. *J Clin Endocrinol Metab.* 2005;90(9):5156-60.
26. Mardis ER, Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet.* 2008;9:387-402.
27. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet.* 2010;11(1):31-46.
28. Ng SB, Turner EH, Robertson PD, et al. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature.* 2009;461(7261):272-6. doi:10.1038/nature08250
29. Molecular Findings Among Patients Referred for Clinical Whole-Exome Sequencing *JAMA.* 2014;312(18):1870-9. doi: 10.1001/jama.2014.14601
30. Kamil G, Yoon JY, Yoo S, et al. Clinical relevance of targeted exome sequencing in patients with rare syndromic short stature. *Orphanet J Rare Dis.* 2021;16:297. <https://doi.org/10.1186/s13023-021-01937-8>
31. Eggermann T, Elbracht M, Kurth I, et al. Genetic testing in inherited endocrine disorders: joint position paper of the European reference network on rare endocrine conditions (Endo-ERN). *Orphanet J Rare Dis.* 2020;15(1):144. Published 2020 Jun 8. doi:10.1186/s13023-020-01420-w
32. Liu P, Meng L, Normand EA, et al. Reanalysis of Clinical Exome Sequencing Data. *N Engl J Med.* 2019;380(25):2478-80. doi:10.1056/NEJMc1812033

33. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of Mendelian disorders. *N Engl J Med* 2013;369:1502–11.
34. Lee H, Deignan JL, Dorrani N, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014;312:1880–7.
35. Stark Z, et al. A prospective evaluation of whole-exome sequencing as a first-tier molecular test in infants with suspected monogenic disorders. *Genet Med*. 2016 Nov;18(11):1090-6. doi: 10.1038/gim.2016.1
36. Dauber A, Rosenfeld RG, Hirschhorn JN. Genetic evaluation of short stature. *J Clin Endocrinol Metab*. 2014;99(9):3080-92. doi:10.1210/jc.2014-1506
37. Wit JM, et al. MECHANISMS IN ENDOCRINOLOGY: Novel genetic causes of short stature. *Eur J Endocrinol*. 2016;174(4):R145-73.
38. Grunauer M, Jorge AAL. Genetic short stature. *Growth Horm IGF Res*. 2018;38:29-33. doi:10.1016/j.ghir.2017.12.003
39. Yengo L, Sidorenko J, Kempner KE, et al. Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum Mol Genet*. 2018;27(20):3641-49. doi:10.1093/hmg/ddy271
40. David A, Hwa V, Metherell LA, et al. Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. *Endocr Rev*. 2011;32(4):472-97. doi:10.1210/er.2010-0023
41. J.A. Phillips, B.L. Hjelle, P.H. Seeburg, M. Zachmann, Molecular basis for familial isolated growth hormone deficiency, *Proc. Natl. Acad. Sci. U. S. A*. 1981;78: 6372–5.
42. Villela TR, Freire BL, Braga NTP, Arantes RR, Funari MFA, Alexander JAL, Silva IN. Growth Hormone insensitivity (Laron syndrome): Report of a new family and review of Brazilian patients. *Genet Mol Biol*. 2020; 42(4):e20180197. doi: 10.1590/1678-4685-GMB-2018-0197. PMID: 31429861; PMCID: PMC7197995.
43. Walenkamp MJ, Wit JM. Single gene mutations causing SGA. *Best Pract Res Clin Endocrinol Metab*. 2008;22(3):433-46.
44. Begemann M, et al. Paternally Inherited IGF2 Mutation and Growth Restriction. *N Engl J Med*. 2015;373(4):349-56.

45. Masunaga Y, Inoue T, Yamoto K, Fujisawa Y, Sato Y, Kawashima-Sonoyama Y, Morisada N, Iijima K, Ohata Y, Namba N, Suzumura H, Kuribayashi R, Yamaguchi Y, Yoshihashi H, Fukami M, Saitsu H, Kagami M, Ogata T. IGF2 Mutations. *J Clin Endocrinol Metab.* 2020;105(1):dgz034. doi: 10.1210/clinem/dgz034. PMID: 31544945.
46. Leal Ade C, et al. Mutations in insulin-like growth factor receptor 1 gene (IGF1R) resulting in intrauterine and postnatal growth retardation. *Arq Bras Endocrinol Metabol.* 2011;55(8):541-9.
47. Ester WA, et al. Two short children born small for gestational age with insulin-like growth factor 1 receptor haploinsufficiency illustrate the heterogeneity of its phenotype. *J Clin Endocrinol Metab.* 2009;94(12):4717-27.
48. Gkourogianni A, et al., Clinical Characterization of Patients With Autosomal Dominant Short Stature due to Aggreacan Mutations. *J Clin Endocrinol Metab.* 2017;102(2):460-9.
49. Spranger, J., Bone dysplasia 'families'. *Pathol Immunopathol Res*, 1988; 7(1-2):76-80.
50. Pejchalova K, Krejci P, Wilcox WR. C-natriuretic peptide: an important regulator of cartilage. *Mol Genet Metab.* 2007;92(3):210-5.
51. Tidyman WE, Rauen KA. The RASopathies: developmental syndromes of Ras/MAPK pathway dysregulation. *Curr Opin Genet Dev.* 2009;19(3):230-6.
52. Malaquias AC, et al. Growth standards of patients with Noonan and Noonan-like syndromes with mutations in the RAS/MAPK pathway. *Am J Med Genet A.* 2012;158A(11):2700-6.
53. Santen GW, et al. Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat Genet.* 2012;44(4):379-80.
54. Ye B, et al. Suppression of SRCAP chromatin remodelling complex and restriction of lymphoid lineage commitment by Pcid2. *Nat Commun*, 2017; 8(1):1518.
55. Tonkin ET, et al. NIPBL, encoding a homolog of fungal Scc2-type sister chromatid cohesion proteins and fly Nipped-B, is mutated in Cornelia de Lange syndrome. *Nat Genet.* 2004;36(6):636-41.

56. Hatkevich T, et al. Bloom Syndrome Helicase Promotes Meiotic Crossover Patterning and Homolog Disjunction. *Curr Biol*. 2017;27(1):96-102.
57. Maldjian P, Sanders AE. 22q11 Deletion Syndrome with Vascular Anomalies. *J Clin Imaging Sci*. 2018;8:1.
58. Binder G, et al. The endocrine phenotype in silver-russell syndrome is defined by the underlying epigenetic alteration. *J Clin Endocrinol Metab*. 2008;93(4):1402-7.
59. Man O-OMli. MIM® - Online Mendelian Inheritance in Man. An Online Catalog of Human Genes and Genetic Disorders. Johns Hopkins University, 2017.
60. Hasegawa K, Tanaka H. Children with short-limbed short stature in pediatric endocrinological services in Japan. *PediatrInt*. 2014;56:809-12.
61. Guo MH, Shen Y, Walvoord EC, et al. Whole exome sequencing to identify genetic causes of short stature. *Horm Res Paediatr* 2014; 82: 44-52.10
62. Kaur A, Phadke SR. Analysis of short stature cases referred for genetic evaluation. *Indian J Pediatr*. 2012;79:1597-600.
63. Renes JS, et al. Bloom syndrome in short children born small for gestational age: a challenging diagnosis. *J Clin Endocrinol Metab*. 2013;98(10):3932-8.
64. Valencia CA, et al. Clinical impact and cost-effectiveness of whole exome sequencing as a diagnostic tool: a pediatric center's experience. *Front. Pediatr*. 2015.
65. Seaver LH, Irons M, Committee ACoMGAPPaG. ACMG practice guideline: genetic evaluation of short stature. *Genet Med* 2009; 11: 465-470
66. Hauer NN, Popp B, Schoeller E, et al. Clinical relevance of systematic phenotyping and exome sequencing in patients with short stature. *Genet Med*. 2018;20(6):630-638. doi:10.1038/gim.2017.159
67. Homma TK, Freire BL, Honjo Kawahira RS, Dauber A, Funari MFA, Lerario AM, et al. Genetic Disorders in Prenatal Onset Syndromic Short Stature Identified by Exome Sequencing. *J Pediatr*. 2019;215:192-8. Epub 2019/10/17. doi: 10.1016/j.jpeds.2019.08.024. PubMed PMID: 31630891.

68. Freire BL, Homma TK, Funari MFA, Lerario AM, Vasques GA, Malaquias AC, et al. Multigene Sequencing Analysis of Children Born Small for Gestational Age With Isolated Short Stature. *J Clin Endocrinol Metab*. 2019;104(6):2023-30. doi: 10.1210/jc.2018-01971. PubMed PMID: 30602027.
69. Kuczmarski RJ, et al. 2000 CDC Growth Charts for the United States: methods and development. *Vital Health Stat* 11. 2002(246):1-190.
70. Chomczynski P, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal Biochem*. 1987;162(1):156-9. doi:10.1006/abio.1987.9999
71. Naslavsky MS, et al. Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum Mutat*. 2017Jul;38(7):751-63. doi: 10.1002/humu.23220. Epub 2017 May 3. PMID: 28332257.
72. Lerario AM, et al. Onuchic LF, Marie SKN, Mendonca BB. SELAdb: A database of exonic variants in a Brazilian population referred to a quaternary medical center in São Paulo. *Clinics (Sao Paulo)*. 2020;75:e1913. doi: 10.6061/clinics/2020/e1913. Epub 2020 Aug 10. PMID: 32785571; PMCID: PMC7410354.
73. Demidov G, Ossowski S. ClinCNV: novel method for allele-specific somatic copy-number alterations detection, bioRxiv 837971; doi: <https://doi.org/10.1101/837971>.
74. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. *Bioinformatics*. 2012Sep;28(18):i333-9. doi: 10.1093/bioinformatics/bts378. PMID: 22962449; PMCID: PMC3436805.
75. Richards S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-24.
76. Jarvik GP, Browning BL. Consideration of Cosegregation in the Pathogenicity Classification of Genomic Variants. *Am J Hum Genet*, 2016; 98(6):1077-81.

77. Homma TK, Freire BL, Honjo R, et al. Growth and Clinical Characteristics of Children with Floating-Harbor Syndrome: Analysis of Current Original Data and a Review of the Literature. *Horm Res Paediatr*. 2019;92(2):115-23. doi:10.1159/000503782
78. Lin et al., SCUBE3 loss-of-function causes a recognizable recessive developmental disorder due to defective bone morphogenetic protein signaling, *The American Journal of Human Genetics* (2021), <https://doi.org/10.1016/j.ajhg.2020.11.015>
79. Lessel D, Gehbauer C, Bramswig NC, et al. BCL11B mutations in patients affected by a neurodevelopmental disorder with reduced type 2 innate lymphoid cells. *Brain*. 2018;141(8):2299–311. doi:10.1093/brain/awy173
80. Freire BL, Homma TK, Funari MFA et al. Homozygous loss of function BRCA1 variant causing a fanconi-anemia-like phenotype, a clinical report and review of previous patients. *Eur J MedGenet* 2017
81. de Bruin C, Finlayson C, Funari MF, Vasques GA, Lucheze Freire B, Lerario AM, Andrew M, Hwa V, Dauber A, Jorge AA. Two Patients with Severe Short Stature due to a FBN1 Mutation (p.Ala1728Val) with a Mild Form of Acromicric Dysplasia. *Horm Res Paediatr*. 2016;86(5):342-8. doi: 10.1159/000446476. Epub 2016 Jun 2. PMID: 27245183; PMCID: PMC5135661.
82. de Bruijn DR, dos Santos NR, Kater-Baats E, et al. The cancer-related protein SSX2 interacts with the human homologue of a Ras-like GTPase interactor, RAB3IP, and a novel nuclear protein, SSX2IP. *Genes Chromosomes Cancer*. 2002;34(3):285-98. doi:10.1002/gcc.10073
83. Martínez-Carreres L, et al. CDK4 Regulates Lysosomal Function and mTORC1 Activation to Promote Cancer Cell Survival. *Cancer Res*. 2019 Oct 15;79(20):5245-59. doi: 10.1158/0008-5472.CAN-19-0708. Epub 2019 Aug 8. PMID: 31395606.
84. Baple EL, Chambers H, Cross HE, et al. Hypomorphic PCNA mutation underlies a human DNA repair disorder. *J Clin Invest*. 2014;124(7):3137-46. doi:10.1172/JCI74593

85. Kim YM, Lee YJ, Park JH, Lee HD, Cheon CK, Kim SY, et al. High diagnostic yield of clinically unidentifiable syndromic growth disorders by targeted exome sequencing. *Clin Genet*. 2017;92:594-605.
86. Huang Z, Sun Y, Fan Y, Wang L, Liu H, Gong Z, et al. Genetic evaluation of 114 Chinese short stature children in the next generation era: a single center study. *Cell Physiol Biochem*. 2018;49:295-305.
87. Vissers L, van Nimwegen KJM, Schieving JH, Kamsteeg EJ, Kleefstra T, Yntema HG, et al. A clinical utility study of exome sequencing versus conventional genetic testing in pediatric neurology. *Genet Med*. 2017;19:1055-63.
88. Raible SE, Mehta D, Bettale C, et al. Clinical and molecular spectrum of CHOPS syndrome. *Am J Med Genet A*. 2019;179(7):1126-38. doi:10.1002/ajmg.a.61174
89. Domchek SM, Tang J, Stopfer J, Lilli DR, Hamel N, Tischkowitz M, et al. Biallelic deleterious BRCA1 mutations in a woman with early-onset ovarian cancer. *Cancer Discov*. 2013;3(4):399–405.
90. Sawyer SL, Tian L, Kähkönen M, Schwartzentruber J, Kircher M, Majewski, J, et al. Biallelic mutations in BRCA1 cause a new Fanconi anemia subtype. *Cancer Discov*. 2015.5(2):135-42.
91. Galli-Tsinopoulou A, Kyrgios I, Emmanouilidou E, Maggana I, Kotanidou E, Kokka P, et al. Growth hormone deficiency: an unusual presentation of floating harbor syndrome. *Hormones (Athens)*. 2011Jul-Sep;10(3):236-40.
92. Zhang S, Chen S, Qin H, Yuan H, Pi Y, Yang Y, et al. Novel genotypes and phenotypes among Chinese patients with Floating-Harbor syndrome. *Orphanet J Rare Dis*. 2019Jun;14(1):144.
93. Cannavò S, Bartolone L, Lapa D, Venturino M, Almoto B, Violi A, et al. Abnormalities of GH secretion in a young girl with Floating- Harbor syndrome. *J Endocrinol Invest*. 2002Jan;25(1):58–64.
94. Walenkamp MJ, Robers JM, Wit JM, Zandwijken GR, van Duyvenvoorde HA, Oostdijk W, et al. Phenotypic Features and Response to GH Treatment of Patients with a Molecular Defect of the IGF-1 Receptor. *J Clin Endocrinol Metab*. 2019 Aug; 104(8): 3157-71.

95. García RJ, Kant SG, Wit JM, Mericq V. Clinical and genetic characteristics and effects of long-term growth hormone therapy in a girl with Floating-Harbor syndrome. *J Pediatr Endocrinol Metab.* 2012; 25(1-2): 207–12.
96. Ho SS, Urban AE, Mills RE. Structural variation in the sequencing era. *Nat Rev Genet.* 2020;21(3):171-189. doi:10.1038/s41576-019-0180-9
97. Suntharalingham JP, et al. Analysis of CDKN1C in fetal growth restriction and pregnancy loss. *F1000Res.* 2019Jan23;8:90. doi: 10.12688/f1000research.15016.2. PMID: 31497289; PMCID: PMC6713069.
98. Wakeling EL, et al. Diagnosis and management of Silver-Russell syndrome: first international consensus statement. *Nat Rev Endocrinol.* 2017 Feb;13(2):105-24. doi: 10.1038/nrendo.2016.138. Epub 2016 Sep 2. PMID: 27585961.
99. Dikow N, Maas B, et al. The phenotypic spectrum of duplication 5q35.2-q35.3 encompassing NSD1: is it really a reversed Sotos syndrome? *Am J Med Genet A.* 2013Sep;161A(9):2158-66. doi: 10.1002/ajmg.a.36046. Epub 2013 Aug 2. PMID: 23913520.
100. Rosenfeld JA, et al. Further Evidence of Contrasting Phenotypes Caused by Reciprocal Deletions and Duplications: Duplication of NSD1 Causes Growth Retardation and Microcephaly. *Mol Syndromol.* 2013Jan;3(6):247-54. doi: 10.1159/000345578. Epub 2013 Jan 5. PMID: 23599694; PMCID: PMC3569101.
101. Giabicani E, et al. Increasing knowledge in IGF1R defects: lessons from 35 new patients. *J Med Genet.* 2020 Mar;57(3):160-8. doi: 10.1136/jmedgenet-2019-106328. Epub 2019 Oct 5. PMID: 31586944.
102. Clark MM, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med.* 2018Jul9;3:16. doi: 10.1038/s41525-018-0053-8. PMID: 30002876; PMCID: PMC6037748.

103. Nakano K, Shiroma A, Shimoji M, Tamotsu H, Ashimine N, Ohki S, Shinzato M, Minami M, Nakanishi T, Teruya K, Satou K, Hirano T. Advantages of genome sequencing by long-read sequencer using SMRT technology in medical area. *Hum Cell*. 2017Jul;30(3):149-61. doi: 10.1007/s13577-017-0168-8. Epub 2017 Mar 31. PMID: 28364362; PMCID: PMC5486853.
104. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. *Nat Rev Genet*. 2019Nov;20(11):631-56. doi: 10.1038/s41576-019-0150-2. Epub 2019 Jul 24. PMID: 31341269.
105. Olova N, et al. Comparison of whole-genome bisulfite sequencing library preparation strategies identifies sources of biases affecting DNA methylation data. *Genome Biol*. 2018Mar15;19(1):33. doi: 10.1186/s13059-018-1408-2. Erratum in: *Genome Biol*. 2019 Feb 22;20(1):43. PMID: 29544553; PMCID: PMC585637.