UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

EDUARDA MORGANA DA SILVA

**DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA
MACROSTOMIA**

Ribeirão Preto

2015

EDUARDA MORGANA DA SILVA

**DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA
MACROSTOMIA**

**Dissertação de Mestrado apresentada à
Faculdade de Medicina de Ribeirão Preto,
Universidade de São Paulo, para obtenção do
título de mestre em Ciências.**

Área de Concentração: Genética

**Orientador: Prof. Dr. Victor Evangelista de
Faria Ferraz**

**Co-orientador: Prof. Dr. Wilson Araújo da Silva
Junior**

Ribeirão Preto

2015

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Silva, Eduarda Morgana da

Determinação da Base Molecular da Síndrome Ablefaria Macrostomia.

Ribeirão Preto, 2015.

75 f.:il; 30cm.

Dissertação de Mestrado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP –
Área de concentração: Genética.

Orientador: Victor Evangelista de Faria Ferraz

Palavras-chave: Síndrome Ablefaria Macrostomia, exoma, TWIST2, doença rara, Nova
Geração de Sequenciamento.

FOLHA DE APROVAÇÃO

SILVA, Eduarda Morgana da

Determinação da base molecular da Síndrome Ablefaria Macrostomia.

Dissertação de Mestrado apresentada à Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, para obtenção do título de mestre em Ciências.

Área de Concentração: Genética

Orientador: Prof. Dr. Victor Evangelista de Faria Ferraz

Co-orientador: Prof. Dr. Wilson Araújo da Silva

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

DEDICATÓRIA

“A todos os que se dedicam à ciência, e com
as próprias mãos a torna possível”.

AGRADECIMENTOS

Agradeço primeiramente a Deus, por ser tão afortunada de amor nessa vida.

Agradeço aos meus mestres, Prof. Dr. Victor Ferraz e Prof. Dr. Wilson Araújo, por todo apoio, ajuda e amizade durante a construção da minha carreira.

Aos mestres de formação pessoal e profissional Verônica Roncelli e André Lima, com vocês aprendi muito, serei sempre grata.

Agradeço aos meus pais, Josi e Jackson, que tanto me amam e me apoiam nas minhas aventuras e desventuras em série. Sem o apoio de vocês, eu nada seria. Agradeço à minha irmã Vitória, por quem tenho enorme apreço e que enche meu coração de alegria. Agradeço aqui ao meu namorado Juca, por tantos anos de companheirismo e amor, são tantas histórias e tantos momentos vividos, que não há como esquecer.

Às minhas avós Maria, Marilu e Odélia, e ao meu avô Gustavo. É tanto amor, conselhos e dicas de como ser feliz, que essas linhas jamais serão suficientes para expressar minha gratidão.

Aos meus tios-irmãos Meri, Rodrigo e Michele e às suas crianças amadas, que me causam um vazio no coração de saudade.

À minha família por opção: Marli, Gelso, Ricardo, Flávia e a pequena Paola, por todo o carinho e atenção que sempre tiveram comigo.

Ao Marcos e à Marcela, pois por mais imprevisível que seja a vida, ela sempre nos dá uma segunda chance. Ao meu tio Marcelo, a Tati e meus primos, por me acolher e receber sempre tão bem em sua casa.

Ao meu primo e grande amigo Eduardo, por quem tenho muito carinho e sei que posso contar.

Agradeço às minhas queridas de longa data Babi e Taci, com quem partilho todos os bons e maus momentos da minha vida, a vocês eu serei eternamente grata.

Às minhas amigas Salesianas: Pêra, Gabi, Ana, Rá, Mari, Amanda e Taisa, aos muitos anos de amizade, jogos e folias.

Ao meu hexágono: Nika, Jô, Bá e Jenny, com vocês os anos de faculdade passaram voando, tudo foi muito mais divertido.

A todos os meus amigos abraçados por essa cidade do interior, que formaram comigo uma nova família: Simone, Ayling, Natalia, Jaqueline, Isabela, Melina, Júlia, Elius, Ádamo, João, Jorge, Teva, Sindy, Clarissa, Carlos, Lucía, Juliana e Magno, muito obrigada de verdade, guardarei vocês pra sempre com muito carinho em meu coração.

Um agradecimento especial ao meu time de basquete, CCC Medicina Ribeirão.

Aos amigos e companheiros de laboratório: Kamila, Luiza, Jéssica, Dri, Ane, Greice, Thaís, Rafa, Fred, Bruninha, Anelisa, Carol, Carolina, Dani, Diana, Graziela, Aline, Ana Júlia, Marcelo, Jorge, Karina, Júlio, Dalila, Willys e Cris. Por todas as risadas, ajuda e a amizade.

Aos meus amigos, que muito me ajudaram e riram comigo: Jake, Inês e Christos.

Ao Prof. Dr. Michael Simpson da King's College London, por toda a ajuda e atenção dispensada durante toda minha estadia em seu laboratório.

A Prof. Dra. Iscia Lopes e a sua equipe, por me receberem em seu laboratório de forma tão agradável.

A toda a equipe envolvida direta ou indiretamente no atendimento dos pacientes deste estudo, especialmente Dr. Antônio Augusto Velasco e Cruz, Dra. Ester Silverira Ramos e Dr. João Monteiro de Pina Neto.

A Susie, Silvia, Meire, Carla, e todos os funcionários do departamento de genética da FMRP-USP, assim como do Hemocentro, por auxiliar nas questões burocráticas.

A Michele e ao Gerenciamento de Projetos da FMRP-USP, por toda ajuda e competência em momentos fundamentais.

Ao CNPQ e à FAPESP, pelo apoio financeiro.

A família deste estudo, que permitiu todo e qualquer conhecimento gerado.

RESUMO

SILVA, E. M. Determinação da base molecular da Síndrome Ablefaria Macrostomia. 2015. 75 f. Dissertação (Mestrado). Faculdade de Medicina de Ribeirão Preto, 2015.

A Síndrome Ablefaria Macrostomia (SAM) é uma condição rara, onde os pacientes apresentam características clínicas marcantes como o encurtamento ou ausência das pálpebras superiores e inferiores, ausência de sobrancelhas e cílios, macrostomia por defeitos na fusão dos lábios, entre outros. O padrão de herança da síndrome não está elucidado, tendo a herança autossômica dominante com expressividade variável sido sugerida. SAM possui sobreposição fenotípica com a Síndrome de Barber-Say e com a Síndrome de Fraser, porém nenhum gene já descrito apresentou mutação nos pacientes portadores da SAM. A abordagem genômica no estudo de doenças raras tem sido amplamente utilizada, devido principalmente ao surgimento da Nova Geração de Sequenciamento, que possui alto poder de discriminar as sequências nucleotídicas com grande cobertura, em um curto período de tempo. No presente estudo o sequenciamento completo do exoma foi realizado, com cinco indivíduos de uma mesma família, três membros afetados e dois não, e permitiu a análise das regiões codificantes nestes indivíduos. A base molecular da Síndrome Ablefaria Macrostomia é aqui sugerida como autossômica dominante, e decorrente da mutação nova não sinônima c.223G>A (p.E75K) no gene TWIST2. Essa mutação patogênica ocasiona a troca de um aminoácido pequeno de carga negativa, o ácido glutâmico, para um aminoácido de cadeia maior carregado positivamente, a lisina. A modelagem *in silico* da proteína Twist2 mostrou que a estrutura geral tridimensional da proteína não foi alterada, mas a troca do aminoácido ocorre na posição 75 dentro do domínio básico HLH, e pode impedir a formação de dímeros, ou a própria ligação ao DNA. Sugere-se ainda que a heterogeneidade de fenótipos associados a mutações no gene TWIST2, pode ser atribuída às interações que essa proteína é capaz de formar, e a ampla ação regulatória que ela desempenha em diversos genes do desenvolvimento.

Palavras-chave: Síndrome Ablefaria Macrostomia, exoma, TWIST2, doença rara, Nova Geração de Sequenciamento.

ABSTRACT

SILVA, E. M. Determining the molecular basis of Ablepharon Macrostomia Syndrome. 2015. 75 f. Dissertação (Mestrado). Faculdade de Medicina de Ribeirão Preto, 2015.

Ablepharon-Macrostomia Syndrome (AMS) is a rare condition characterised by absent or hypoplastic eyelids, absent eyebrows and eyelashes, macrostomia caused by fusion defects of the mouth with unfused lateral commissures, as well as other clinical features. The inheritance pattern has not been confirmed and while autosomal dominant inheritance with variable expressivity has been suggested, recessive inheritance has not been ruled out. The phenotype of AMS overlaps that of Barber-Say and Fraser Syndrome, but any reported gene for these syndromes is mutated on AMS patients. The genomic approach for rare disease studies has been widely used mainly due to the emergence of Next Generation Sequencing, which is very effective at determining nucleotide sequences with large coverage in a short period of time. The whole exome sequencing of five family members was undertaken, with three affected and two unaffected, and the coding regions of the individuals were subsequently analysed. The molecular basis of AMS is suggested here as autosomal dominant, and due to a novel non-synonymous mutation c.223G>A (p.E75K), in TWIST2 gene. This pathogenic mutation causes glutamic acid, a small negatively charged amino acid, to be substituted for a larger and positively charged lysine. The *in silico* protein modeling of Twist2 shows that the general 3D-structure of the protein is not affected, but the amino acid change is located inside the basic Helix-Loop-Helix domain which could disrupt dimerization and DNA binding. It has also been suggested that the phenotype heterogeneity associated with mutations on TWIST2 gene can be attributed to the interactions that this protein is capable of, and the role that it plays in the regulation of several developmental genes.

Key words: Ablepharon Macrostomia Syndrome, exome, TWIST2, rare disease, next generation sequencing.

LISTA DE FIGURAS

Figura 1. Paciente com as características clínicas da SAM.....	16
Figura 2. Comparação entre o Sequenciamento Completo do Genoma e do Exoma.....	21
Figura 3. Heredograma da família em estudo.....	32
Figura 4. Fluxograma de filtros utilizados para seleção de variantes.....	36
Figura 5. Cromatogramas da validação da mutação (c.223G>A) no gene TWIST2.....	40
Figura 6. Estruturas tridimensionais das proteínas selvagem e mutante, obtidas com a ferramenta <i>SWISSMODEL</i>	41
Figura 7. Representação da troca do aminoácido referência, o ácido glutâmico, pelo aminoácido lisina.....	42
Figura 8. Modelo tridimensional da proteína Twist2, gerado pela ferramenta HOPE.....	42
Figura 9. Estrutura básica da proteína Twist2.....	47

LISTA DE TABELAS

Tabela 1. Critérios Clínicos da SAM nos indivíduos afetados.....	31
Tabela 2. Dados do Sequenciamento do Exoma.....	33
Tabela 3. Dados das mutações anotadas.....	34
Tabela 4. Dados das Mutações em Sítios de Recombinação e Escore no RegulomeDB.....	35
Tabela 5. Ranqueamento das Mutações Seleccionadas, por meio da ferramenta SPRING.....	37
Tabela 6. Análise de Genes Candidatos.....	38
Tabela 7. Homologia da sequência de aminoácidos da proteína Twist2 entre espécies.....	52

LISTA DE ABREVIATURAS E SIGLAS

3D	Tridimensional
A	Adenina
aa.	Aminoácido
ADAT1	do inglês, <i>Adenosine Deaminase, tRNA-specific 1</i>
BAIAP3	do inglês, <i>BAIL-associated protein 3</i>
bHLH	do inglês, <i>basic Helix-Loop-Helix</i>
C	Citosina
C8orf37	do inglês, <i>Chromosome 8 open reading frame 37</i>
CHCHD3	do inglês, <i>Coiled-coil-helix-coiled-coil-helix domain containing 3</i>
ChIP-seq	Sequenciamento de regiões da Cromatina Imuno-precipitadas
Crom.	Cromossomo
DAS	do inglês, <i>Dense Alignment Surface method</i>
DNA	Ácido Desoxirribonucleico
DNase	do inglês, <i>A desoxirribonuclease</i>
DNase-seq	Sequenciamento de Sítios de Sensibilidade a DNase
E	Ácido Glutâmico
ENCODE	do inglês, <i>The Encyclopedia of DNA Elements</i>
FMRP	Faculdade de Medicina de Ribeirão Preto
FRAS1	do inglês, <i>Fraser syndrome 1 homolog</i>
FREM1	do inglês, <i>FRAS1 related extracellular matrix 1</i>
FREM2	do inglês, <i>FRAS1 related extracellular matrix 2</i>
FREM3	do inglês, <i>FRAS1 related extracellular matrix 3</i>
G	Guanina
GAIIX	do inglês, <i>Genome Analyzer Iix</i>
GEO	do inglês, <i>Gene Ontology</i>
GRIP1	do inglês, <i>Glutamate receptor interacting protein 1</i>
hg19	do inglês, <i>Human genome 19</i>
HLH	Domínio hélice-alça-hélice
HOPE	do inglês, <i>Have yOur Protein Explained</i>
HSSP	do inglês, <i>Homology-Derived Secondary Structure of Proteins</i>
IFT140	do inglês, <i>Intraflagellar transport 140</i>
IGV	do inglês, <i>Integrative Genome Viewer</i>
IMPC	do inglês, <i>International Mouse Phenotyping Consortium</i>
Indels	Mutações do tipo inserções e deleções
KANSL1	do inglês, <i>KAT8 regulatory NSL complex subunit 1</i>
LAG	Lipodistrofia Adquirida Generalizada
LAP	Lipodistrofia Adquirida Parcial
LCG	Lipodistrofia Congênita Generalizada
LFP	Lipodistrofia Familiar Parcial
MAGEB2	do inglês, <i>melanoma antigen family B2</i>
metil-seq	Sequenciamento de Regiões Metiladas

miRNA	Micro RNA
NCBI	do inglês, <i>National Center for Biotechnology Information</i>
NEO1	do inglês, <i>Neogenin 1</i>
NGS	Nova Geração de Sequenciamento
OMIM	do inglês, <i>Online Mendelian Inheritance in Man</i>
PCR	do inglês, <i>Polymerase Chain Reaction</i>
PKD1	do inglês, <i>polycystic kidney disease 1</i>
PolyPhen2	do inglês, <i>Polymorphism Phenotyping 2</i>
RANBP2	RAN binding protein 2
RegulomeDB	do inglês, <i>Regulome Data Base</i>
RIF1	do inglês, <i>Replication timing regulatory factor 1</i>
RNA	Ácido Ribonucleico
RNA-seq	Sequenciamento de RNA
SAM	Síndrome Ablefaria Mascrostomia
SBS	Síndrome Barber-Say
SF	Síndrome de Fraser
SIFT	do inglês, <i>The Sorting Intolerant from Tolerant</i>
SLC39A12	do inglês, <i>Solute carrier family 39 (zinc transporter), member 12</i>
SLN	Sinal de Localização Nuclear
SNVs	do inglês, <i>Single Nucleotide Variants</i>
SPRING	do inglês, <i>Snv PRioritization via the INtegration of Genomic data</i>
T	Timina
TEM	Transição Epitélio Mesênquima
TLE3	do inglês, <i>Transducin-like enhancer of split 3</i>
TWIST1	do inglês, <i>Twist family bHLH transcription factor 1</i>
TWIST2	do inglês, <i>Twist family bHLH transcription factor 2</i>
Uniprot	do inglês, <i>Universal Protein Resource</i>
USP	Universidade de São Paulo
UTR	do inglês, <i>Untranslated</i>
ZNF354A	do inglês, <i>Zinc finger protein 354^a</i>

LISTA DE SÍMBOLOS

%	Porcentagem
cm	Centímetro
Kg	Kilogramas
pb	Pares de base
pM	Picomolar
ng	Nanograma
μL	Microlitro

SUMÁRIO

1. INTRODUÇÃO	15
1.1. A SÍNDROME ABLEFARIA MACROSTOMIA (SAM)	15
1.1.1. O PADRÃO DE HERANÇA DA SAM.....	16
1.1.2. SOBREPOSIÇÃO FENOTÍPICA COM AS SÍNDROMES DE BARBER-SAY E FRASER.....	16
1.2. NOVA GERAÇÃO DE SEQUENCIAMENTO (NGS).....	18
1.2.1. SEQUENCIAMENTO COMPLETO DO GENOMA E EXOMA	19
2. OBJETIVOS.....	21
2.1. OBJETIVO GERAL	21
2.2. OBJETIVOS ESPECÍFICOS.....	21
3. METODOLOGIA	22
3.1. OBTENÇÃO DAS AMOSTRAS.....	22
3.2. ASPECTOS ÉTICOS	22
3.3. EXTRAÇÃO E QUANTIFICAÇÃO DE DNA.....	22
3.4. NOVA GERAÇÃO DE SEQUENCIAMENTO	22
3.4.1. PREPARO DA BIBLIOTECA GENÔMICA	22
3.4.2. CLUSTERIZAÇÃO.....	23
3.4.3. SEQUENCIAMENTO	24
3.5. ANÁLISE DE DADOS.....	24
3.5.1. FERRAMENTAS DE BIOINFORMÁTICA	24
3.5.2. ANOTAÇÕES, HERANÇA E PREDIÇÕES.....	25
3.6. VALIDAÇÃO POR SEQUENCIAMENTO SANGER	27
3.7. AVALIAÇÃO <i>IN SILICO</i> DO EFEITO DA MUTAÇÃO NA PROTEÍNA	27
4. RESULTADOS.....	28
4.1. REAVALIAÇÃO CLÍNICA.....	28
4.2. DADOS DO SEQUENCIAMENTO.....	30
4.3. DADOS E FILTROS DE SELEÇÃO DAS MUTAÇÕES	31
4.4. AVALIAÇÃO <i>IN SILICO</i> DO EFEITO DA MUTAÇÃO NA PROTEÍNA	39
5. DISCUSSÃO.....	41
5.1. REAVALIAÇÃO CLÍNICA	41
5.2. DADOS DO SEQUENCIAMENTO DE NOVA GERAÇÃO.....	42
5.3. SELEÇÃO DE MUTAÇÕES	43
5.4. VALIDAÇÃO DA MUTAÇÃO SELECIONADA E VARIABILIDADE INTRAFAMILIAL	45
5.5. O GENE TWIST2	46
5.6. MUTAÇÕES NO GENE TWIST2.....	48
5.7. MODELAGEM <i>IN SILICO</i> DA PROTEÍNA TWIST2.....	49
6. CONCLUSÕES	52
7. REFERÊNCIAS	54
8. ANEXOS.....	64
ANEXO A. APROVAÇÃO DO COMITÊ DE ÉTICA EM PESQUISA	64
ANEXO B. TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	68

ANEXO C. TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO.....	71
---	----

1. INTRODUÇÃO

1.1. A Síndrome Ablefaria Macrostomia (SAM)

A Síndrome Ablefaria Macrostomia (SAM) é uma condição congênita extremamente rara, primeiramente relatada por McCarthy e West (1977), onde dois meninos apresentavam sinais clínicos marcantes, como encurtamento ou ausência de pálpebras (ablefaria), ausência de sobrancelhas e cílios, alargamento da boca (macrostomia) por defeitos de fusão dos lábios, orelhas rudimentares, pele seca e redundante, ausência de mamilos, alterações na genitália, pele seca e redundante, além de expressivo atraso no desenvolvimento da fala. Outros achados clínicos foram posteriormente descritos, como hérnia ventral e ausência de lanugo (HORNBLASS; REIFLER, 1985), ausência de arco zigomático (JACKSON et al., 1988), onfalocele e camptodactilia (Figura 1) (FERRAZ et al., 2000).

Figura 1. Paciente com as características clínicas da SAM. A: Características faciais, como ablefaria/microablefaria, macrostomia, ausência de cílios e alterações nasais. B: Pele redundante, anomalias auriculares e cicatriz decorrente da onfalocele. Adaptado de FERRAZ et al., 2000.

Desde a primeira documentação da síndrome, apenas 17 casos foram relatados, oito pacientes do sexo feminino e nove do sexo masculino (MCCARTHY; WEST, 1977; HORNBLASS; REIFLER, 1985; MARKOUIZOS et al., 1990; PRICE et al., 1991; CRUZ et al., 1995; PELEGRINO et al., 1996; FERRAZ et al., 2000; AMOR; SAVARIRAYAN, 2001; STEVENS; SARGENT, 2002; BRACANTI et al., 2004; CAVALCANTI et al., 2007; KALLISH et al., 2011; LARUMBE et al., 2011; ROHENA et al., 2011).

1.1.1. O Padrão de Herança da SAM

O padrão de herança da SAM não foi ainda esclarecido, tendo a herança autossômica recessiva primeiramente sido sugerida por McCarthy e West (1977), devido à descrição de gêmeos monozigóticos, filhos de um casal consanguíneo. Porém, esse modo de herança foi questionado pelo relato de uma família em que pai e filha foram descritos com a SAM (CRUZ et al., 1995). Ferraz e colaboradores (2000) relataram a recorrência da síndrome nessa mesma família descrita por Cruz e colaboradores, onde as duas irmãs apresentaram as características clássicas de SAM, e o pai características mais brandas da síndrome, como fissuras nas pálpebras, sobrancelhas arqueadas e rareadas, anormalidades auriculares, e cabelo esparsa. Assim, o padrão de herança autossômico dominante com expressividade variável foi sugerido. Rohena e colaboradores (2011) reforçaram a hipótese de herança autossômica dominante, ao descrever o caso de pai e filha com características clássicas de SAM.

A análise de cariótipo dos pacientes descritos com a SAM não mostrou alterações, com a exceção de um paciente, onde foram encontradas uma inversão parcial entre as bandas 18q12.3 e 18q21.2, e uma deleção de 18q21.3–18q23, o que sugeriu que o locus da SAM poderia ser mapeado no cromossomo 18 (PELLEGRINO et al., 1996). Porém, nenhum outro paciente com SAM apresentou alteração no cariótipo, e nenhum gene foi ainda definitivamente relacionado à síndrome.

1.1.2. Sobreposição fenotípica com as Síndromes de Barber-Say e Fraser

A Síndrome Ablefaria Macrostomia apresenta sobreposição fenotípica com algumas síndromes já descritas, como a Síndrome Barber-Say (SBS) e a Síndrome de Fraser (SF). A SBS é uma condição congênita rara caracterizada por alterações palpebrais (ectrópio), macrostomia, anormalidades auriculares e nasais, pele redundante, hipertricose severa, baixa linha capilar frontal e retardo no crescimento (BARBER et al., 1982; ROCHE et al., 2010).

Diversas características da SBS são sobrepostas a SAM, porém essas são consideradas entidades distintas, devido principalmente ao padrão de alteração ocular. Entretanto, Haensel e colaboradores (2009) descreveram um caso de microablefaria em um paciente com SBS, sugerindo que essas duas síndromes representem apenas uma. O padrão de herança da síndrome é caracterizado como autossômico dominante, pela descrição de herança entre pais e filhos (DINULOS; PAGON, 1999; ROCHE et al., 2010). A herança autossômica dominante, em comum com SAM, reforça a hipótese de que a SBS e a SAM possam ser causadas por mutações no mesmo gene, porém nenhum gene foi ainda descrito para nenhuma das síndromes.

A Síndrome de Fraser é principalmente caracterizada por criptoftalmo, ausência ou malformação dos ductos lacrimais, malformações auriculares, palato ogival, segmentação ao longo do plano médio das fossas nasais e língua, hipertelorismo, estenose laríngea, sindactilia, ampla separação da sínfise pública, deslocamento do umbigo e mamilos, mesentério primitivo do intestino delgado, alterações no desenvolvimento dos rins, anomalias na genitália e aparelho reprodutor (FRASER, 1962). A SF é descrita como autossômica recessiva, podendo ser causada por mutações homozigotas nos genes FRAS1, FREM2 e GRIP1 (MCGREGOR et al., 2005; CAVALCANTI et al., 2007), ou em heterozigose composta para FRAS1 (SLAVOTINEK et al., 2011). As Síndromes SF e SAM podem ser clinicamente diferenciadas pelas características faciais, principalmente pelas anomalias oculares particulares, malformação renal e laringotraqueal (SCHANZE et al., 2013)

O compartilhamento de algumas características fenotípicas nessas três síndromes gerou o questionamento sobre a possibilidade de a base molecular ser comum para as síndromes. Assim, alterações nos mesmos genes, poderiam ser responsáveis pelas síndromes (MAZZANTI et al., 1998; FERRAZ et al., 2000; CAVALCANTI et al., 2007; HAENSEL et al., 2009). Porém, apenas a SF possui genes descritos, dois genes codificantes de proteína de matriz extracelular, FRAS1 (MCGREGOR et al., 2003) e FREM2 (JADEJA et al., 2005), e um gene codificante de proteína mediadora do tráfego e da organização de proteínas transmembranas: GRIP1 (VOGEL et al., 2012). Esses três genes, em ação conjunta com

FREM1, formam um complexo proteico de matriz extracelular, que é expresso nas membranas basais durante o desenvolvimento embriogênico e contribui para a integridade epitelial-mesenquimal (SMYTH; SCAMBLER, 2005; KIYOZUMI et al., 2006).

Schanze e colaboradores (2013) com intuito de verificar a possibilidade desses mesmos genes identificados na SF serem responsáveis pela SAM, realizaram sequenciamento tradicional para os genes FRAS1, FREM1, FREM2, FREM3 e GRIP1 em 11 pacientes portadores da SAM, incluindo os casos previamente reportados por nós (FERRAZ et al., 2000), e não identificaram mutação patogênica nos pacientes. Esses achados sugerem que SAM não compartilha os genes já descritos para FS, e que constitui uma entidade separada do grupo FRAS-FREM (SCHANZE et al., 2013). Dessa forma, uma abordagem mais abrangente, que permita a investigação ampla de mutações presentes nos pacientes com SAM é necessária, para que genes possam relacionados ao fenótipo da síndrome.

1.2. Nova Geração de Sequenciamento (NGS)

O sequenciamento tradicional do tipo Sanger foi por mais de duas décadas a principal ferramenta para determinação de sequências nucleotídicas, e possibilitou que o primeiro genoma humano completo fosse sequenciado em um consórcio que durou mais de uma década (INTERNATIONAL HUMAN GENOME CONSORTIUM, 2004). Porém, a necessidade de novas técnicas que permitissem o sequenciamento de diversos genomas humanos em menores períodos, com custo menor, impulsionou o desenvolvimento da Nova Geração de Sequenciamento (NGS). A maior vantagem da NGS está no alto poder de discriminar as sequências nucleotídicas com grande cobertura, em um curto período de tempo e com custos reduzidos, tornando possível o sequenciamento de genomas completos (METZKEL, 2010).

Dentre as aplicações da NGS está o resequenciamento de regiões alvo codificantes (exoma) e genomas completos, que tem provado ser uma alternativa efetiva para pesquisas que buscam estabelecer a base genética de doenças (NG et al., 2010a; GONZAGA-JAUREGUI; LUPSKI; GIBBS, 2012; LIGHT et al., 2012; BAINBRIDGE et al., 2012). Outras aplicações da NGS para estudos em larga escala têm sido utilizadas, como o sequenciamento de transcriptomas (RNA-seq) (WANG; GERSTEIN; SNYDER, 2009), a análise de perfis de

metilação (metil-seq) (LAIRD, 2010), de padrões proteínas ligantes ao DNA (ChIP-seq) (PARK, 2009), e também de regiões regulatórias (DNase-seq) (SONG; CRAWFORD, 2010).

1.2.1. Sequenciamento Completo do Genoma e Exoma

O sequenciamento do genoma associado às plataformas de NGS permite a análise em larga escala das sequências nucleotídicas, identificando mutações em porções codificantes e não codificantes do genoma. O maior desafio da análise de genomas completos se deve ao grande número de informação gerada, com aproximadamente 3,3 milhões de variantes de nucleotídeo único (da sigla em inglês, SNVs) e 1.000 variantes no número de cópias (da sigla em inglês, CNVs) obtidos por genoma humano diploide (GONZAGA-JAUREGUI; LUPSKI; GIBBS, 2012). O genoma humano possui cerca de 3 bilhões de bases nucleotídicas, sequenciar todo o material genético com grande cobertura encarece a técnica, e também necessita do uso de ferramentas robustas de processamento dos dados por bioinformática (MAJEWSKI et al., 2011; WORTHEY, 2013).

O sequenciamento do exoma requer um passo adicional no processo de construção da biblioteca genômica em relação ao sequenciamento do genoma completo, uma vez que captura seletivamente apenas as regiões codificantes de proteínas, em genes já anotados de bancos de dados (GOH; CHOI, 2012). Esse processo diminui a quantidade de dados gerados, focando em genes codificantes de proteínas que representam aproximadamente 1% do genoma, gerando em torno de 20.000 variantes por exoma analisado, reduzindo assim os custos do sequenciamento e o tempo gasto na análise de dados (Figura 2) (GONZAGA-JAUREGUI; LUPSKI; GIBBS, 2012).

	Sequenciamento Completo do Genoma	Sequenciamento Completo do Exoma
Custo	✓Alto custo, em processo de barateamento;	✓O custo é de um décimo a um terço do genoma completo;
Técnica	✓Sem etapa de captura;	✓Com etapa de captura para regiões codificantes;
Variantes	✓Cobre todas as variantes (SNVs e CNVs) em regiões codificantes e não codificantes; ✓~3,3 milhões de variantes;	✓É limitada a regiões codificantes e sítios de recombinação já anotados (~1% do genoma); ✓~20.000 variantes;
Doenças	✓Identificação de genes em doenças Mendelianas e Complexas.	✓Identificação de genes em doenças Mendelianas.

Figura 2. Comparação entre o Sequenciamento Completo do Genoma e do Exoma. Adaptado de GONZAGA-JAUREGUI; LUPSKI; GIBBS, 2012.

Desde a primeira descrição de sua aplicação, o sequenciamento do exoma tem emergido como uma ferramenta poderosa para a determinação de variantes causais de doenças mendelianas comuns (CHOI et al., 2009; NG et al., 2010b; BAMSHAD et al., 2011; MARIAN, 2012; CHOU et al., 2012; LIN et al., 2012), e também no estudo de doenças raras (TAKATA et al., 2011; ABUZENADAH et al., 2013; MOSHOUS et al., 2013). Rabbani e colaboradores (2012) relataram que mais de 100 genes já foram identificados como causadores de doenças mendelianas, por meio do sequenciamento completo do exoma. Em um estudo mais recente sobre o uso desta técnica de sequenciamento do exoma para detecção de variantes em doenças raras, Beaulieu e colaboradores (2014) analisaram mutações causais para 264 doenças, e puderam determinar genes causais de 146 destas doenças raras. Esses dados mostram o potencial do uso da técnica de sequenciamento do exoma para a detecção de mutações causais em doenças raras.

2. OBJETIVOS

2.1. Objetivo geral

Determinar a base molecular e o padrão de herança da Síndrome Ablefaria Macrostomia, em uma família que apresenta casos da síndrome recorrentes na irmandade.

2.2. Objetivos específicos

- Reavaliar clinicamente uma família com casos da SAM;
- Identificar as mutações presentes nesta família;
- Caracterizar e filtrar as mutações encontradas quanto ao padrão de herança, patogenicidade e função gênica;
- Determinar a mutação causal da SAM e validar por sequenciamento Sanger tradicional;
- Avaliar os possíveis efeitos da mutação na proteína e suas implicações no fenótipo da síndrome.

3. METODOLOGIA

3.1. Obtenção das amostras

A família em estudo, já descrita (CRUZ et al., 1995; FERRAZ et al., 2000), é composta por cinco indivíduos: pai, mãe e três filhas. Foram coletadas amostras de sangue periférico dos cinco familiares para posterior extração do material genético.

3.2. Aspectos éticos

O projeto foi aprovado sob o parecer número 1.012.074, pelo Comitê de Ética em Pesquisa (CEP) do Hospital das Clínicas de Ribeirão Preto FMRP/USP (Anexo A). Termos de Consentimento e Assentimento Livre e Esclarecido, específicos para análise de dados em larga escala foram criados, e estão em anexo B e C, respectivamente.

3.3. Extração e quantificação de DNA

O DNA genômico foi extraído das amostras de sangue, por meio do Kit QIAmp DNA Blood Midi (QIAGEN), seguindo o protocolo estabelecido pelo fabricante. Após a extração, os DNAs das amostras foram quantificados com o equipamento QuantiFluorTM-ST Fluorimeter (Promega), utilizando o kit QuantifluorTM dsDNA System, conforme recomendações do fabricante. As amostras foram em seguida diluídas para a obtenção alíquotas com a concentração de 5ng/ μ L.

3.4. Nova geração de Sequenciamento

3.4.1. Preparo da Biblioteca genômica

Foram necessários 50ng de DNA (ácido desoxirribonucleico) de cada amostra para o preparo da biblioteca genômica com o Kit Nextera Rapid Capture Expanded Exome (Illumina). A primeira etapa deste preparo envolveu a fragmentação do DNA por meio de transposons e ligação de adaptadores e índices às extremidades dos fragmentos gerados. A amplificação por PCR (do inglês, *Polymerase Chain Reaction*) é responsável pela adição desses índices às extremidades. A avaliação dos fragmentos quanto ao tamanho, foi realizada com o equipamento Bioanalyzer 2100 (Agilent Technologies), onde a eletroforese digital em um chip High Sensitivity de DNA foi realizada, sendo observado um rastro característico de fragmentação do DNA, que possuía entre 150 e até pouco menos de 1000 pares de bases nucleotídicas. As amostras foram novamente quantificadas individualmente para determinar quanto material de cada paciente seria necessário para totalizar 500ng total de DNA na biblioteca genômica.

A próxima grande etapa envolveu a hibridação com sondas alvo para éxons, regiões 5' e 3'UTR (do inglês, *untranslated*), e regiões de miRNA (micro RNA), que efetivamente enriqueceram de forma seletiva os fragmentos de DNA. Após a hibridação com as sondas, utilizaram-se esferas de estreptavidina para a captura das sondas contendo as regiões de interesse ligadas. Foram realizadas três lavagens para a remoção das ligações não específicas, e posteriormente foi realizada a eluição das bibliotecas. Essas bibliotecas foram submetidas a uma segunda hibridação para aumentar o enriquecimento, seguido de outro ciclo de lavagens e uma última eluição. As regiões enriquecidas da biblioteca foram então amplificadas por PCR e quantificadas por fluorimetria, uma nova corrida no Bioanalyzer 2100 foi realizada para determinação do tamanho médio das bibliotecas, sendo ainda submetida à quantificação precisa por meio de PCR em tempo real com o kit Kapa Library Quantification, de acordo com as recomendações do fabricante.

3.4.2. Clusterização

As bibliotecas quantificadas foram diluídas para 10pM e submetidas a clusterização no equipamento cBot (Illumina) com o kit TruSeq PE Cluster v2, seguindo recomendações do fabricante. Durante a geração dos *clusters*, cada fragmento de DNA da biblioteca foi incorporado à superfície da *flowcell*, coberta com oligonucleotídeos, utilizados para amplificar o DNA de forma isotérmica, dando origem aos clusters. A linearização retira os

dois adaptadores e assegura que a hibridação e o sequenciamento ocorram somente em uma das fitas de cada cluster. A terceira etapa de clusterização refere-se ao bloqueio das hidroxilas presentes nas extremidades 3' dos clusters linearizados, para evitar a adição de substratos de fluoróforos a qualquer outro segmento que não seja o iniciador de sequenciamento. A desnaturação garante a formação de *templates* de fita simples, que foram hibridados aos iniciadores de sequenciamento.

3.4.3. Sequenciamento

A biblioteca clusterizada foi sequenciada com o kit TruSeq SBS kit v5, por meio de 55 ciclos em cada sentido (paired-end), no equipamento Illumina Genome Analyzer IIX (GAIIx). A tecnologia de sequenciamento baseia-se na terminação da síntese da cadeia de DNA após a detecção do sinal fluorescente. Por terminação entende-se a clivagem do nucleotídeo marcado já detectado que possibilita a adição do próximo nucleotídeo. Os eventos de incorporação e clivagem do nucleotídeo marcado e modificado ocorrem sucessivas vezes dentro da *flowcell*, permitindo a leitura base a base, ao mesmo tempo elimina possíveis erros em sequências com repetições da mesma base (regiões homopoliméricas). As amostras foram sequenciadas de acordo com instruções do fabricante. O tamanho dos fragmentos foi de 2X55 bases (total de 110 bases) podendo sofrer alterações de acordo com análise de qualidade do sequenciamento.

3.5. Análise de Dados

3.5.1. Ferramentas de Bioinformática

A primeira etapa para a análise dos dados gerados pelo sequenciamento foi a **Chamada de Bases** (do inglês, *Base-calling*), onde foram obtidas as imagens das sequências de DNA pelo sequenciador GAIIx e convertidas para arquivos *bcl*, que contém informações básicas das bases nucleotídicas. Através do software CASAVA v.1.8 (Illumina) os arquivos *bcl* foram convertidos para arquivos *fastq*, devido à compatibilidade com outros algoritmos de alinhamento e também para a visualização de dados de qualidade dessas sequências.

Os dados brutos das sequências foram pré-processados, nas seguintes etapas: mapeamento, recalibração dos escores de qualidade das bases, ordenação, marcação e remoção de duplicatas. No mapeamento as reads foram alinhadas contra o genoma de referência hg19 (<http://hgdownload.cse.ucsc.edu/>), utilizando-se o software Novoalign (<http://www.novocraft.com/products/novoalign/>) por meio do algoritmo Needleman Wunsch. Ainda por meio do software Novoalign as bases foram recalibradas. A recalibração dos escores de qualidade das bases garante que a qualidade das bases seja mais precisa, uma vez que estes podem estar super ou subestimados devido a erros empíricos do sequenciador. Esta etapa aumenta a acurácia na chamada de variantes. Por meio da ferramenta Samtools, as reads alinhadas foram armazenadas em arquivos com extensão *SAM* e convertidas ao formato *BAM*, em seguida foram ordenadas, para que então as sequências duplicadas geradas pela PCR fossem marcadas e removidas.

A próxima etapa da análise de dados foi a **Chamada de Variantes** (do inglês, *Variant calling*) onde as variantes foram detectadas também pela ferramenta Samtools (<http://sourceforge.net/projects/samtools/>), que simultaneamente identificou variantes de nucleotídeo único e inserções/deleções (Indels). As sequências foram ainda filtradas por qualidade e indexadas em arquivos no formato *vcf*.

3.5.2. Anotações, Herança e Predições

As variantes encontradas foram anotadas por meio da ferramenta Annovar (<http://www.openbioinformatics.org/annovar/>), para obter informações já descritas em bancos de dados sobre essas alterações, como a anotação do gene, anotação da troca de aminoácidos, escores de conservação da região, frequência alélica, patogenicidade, entre outros. Um banco interno de exomas sequenciados de pacientes atendidos pelo Guy's Hospital, associado à King's College London, foi também utilizado para a anotação de variantes. Essas variantes já anotadas foram então analisadas quanto ao modo de herança e função gênica.

A análise de herança foi realizada assumindo um modelo autossômico dominante. Sabidamente a incidência de SAM é muito baixa, assim somente as variantes novas que ainda não foram descritas foram analisadas. As variantes em comum nos três membros afetados foram selecionadas, e as que estivessem presentes também nos indivíduos não afetados foram

excluídas. A visualização e seleção das variantes foram realizadas por meio do programa Integrative Genome Viewer (IGV) (ROBINSON et al., 2011; THORVALDSDÓTTIR; ROBINSON; MESIROV, 2013).

Foram excluídas variantes localizadas em sítios de recombinação que não fossem essenciais ao processamento dos transcritos e que apresentassem escores altos quando analisados pela ferramenta RegulomeDB. Este programa gera escores para as variantes em regiões não codificantes do genoma, com base em elementos regulatórios já conhecidos. Esses elementos regulatórios incluem: sítios de hipersensibilidade a DNAase, sítios de ligação de fatores de transcrição, e regiões promotoras que tenham sido bioquimicamente caracterizadas para regulação da transcrição. A busca destes dados pelo RegulomeDB ocorre em bancos de dados públicos como GEO (<http://www.ncbi.nlm.nih.gov/geo/>) e Projeto Encode (<https://www.encodeproject.org/>), assim como em dados de publicações (BOYLE et al., 2012).

As variantes restantes foram submetidas a um ranqueamento com base na patogenicidade e na associação com a SAM, por meio da ferramenta SPRING. Essa ferramenta é baseada no efeito funcional de seis escores, calculados com métodos já existentes (SIFT, PolyPhen2, LRT, MutationTaster, GERP e PhyloP), e cinco escores de associação provenientes de diversos bancos de dados genômicos (Gene Ontology, Protein Protein Interactions, Protein Sequences, Protein Domain Annotations e Gene Pathway Annotations). O SPRING calcula o valor de q por meio do método de Fisher com correção de dependência, visando integrar os seis valores de p para os efeitos funcionais, e os cinco valores de p de associação. Com base no valor de q gerado, as variáveis são priorizadas (WU; LI; JIANG, 2014).

Além da priorização realizada pelo software SPRING, as funções gênicas foram também analisadas manualmente com base em bancos de dados como OMIM (<http://www.omim.org/>), NCBI (<http://www.ncbi.nlm.nih.gov/guide>), Orphanet (<http://www.orpha.net/consor/cgi-bin/index.php>) e Genecards (<http://www.genecards.org/>). Essa última análise teve como intuito verificar a relação dos genes selecionados com o fenótipo da síndrome em estudo, e baseadas nessa análise as variantes foram mantidas ou descartadas, permitindo a seleção da mutação responsável pela SAM.

3.6. Validação por Sequenciamento Sanger

A validação da mutação selecionada como potencialmente causadora da doença, foi realizada pela técnica de sequenciamento bi-direcional padrão do tipo Sanger, por meio do Kit de sequenciamento Big Dye Terminator Cycle Sequencing Ready Reaction, em um sequenciador automático ABI 3500XL (Applied Biosystems, CA, USA).

3.7. Avaliação *in silico* do efeito da Mutação na Proteína

A mutação já validada, e sua implicação a nível funcional na proteína foi analisada por meio das ferramentas SWISSMODEL e HOPE. A sequência de aminoácidos da proteína Twist2, correspondente ao registro Q8WVJ9, foi extraída do banco de proteínas Uniprot, e então submetida à modelagem pelas duas ferramentas.

SWISSMODEL fornece anotações da estrutura quaternária, de ligantes essenciais e cofatores que permitem a construção de modelos estruturais completos, incluindo suas estruturas oligoméricas (BIASINI et al., 2014). A ferramenta gera um modelo que estima a qualidade baseada no potencial global QMEAN, que é fornecido como um Escore-Z que relaciona os valores obtidos com os escores calculados, baseado em estruturas de alta resolução por raio-X (BENKERT et al., 2011).

A ferramenta HOPE também analisa os efeitos estruturais e funcionais de mutações pontuais, por meio da coleta de informações em diversas fontes que incluem: cálculos de coordenadas 3D da proteína utilizando os serviços online WHAT IF, escores de conservação utilizando HSSP, anotações de sequências do banco de dados UniProt, e predições utilizando DAS. Todos os dados são armazenados e combinados com as propriedades conhecidas dos aminoácidos, para então serem usados em um esquema de decisões para identificar os efeitos da mutação na função e na estrutura 3D da proteína (VENSELAAR et al., 2010).

4. RESULTADOS

4.1. Reavaliação Clínica

Os indivíduos da família em estudo foram avaliados clinicamente e características típicas da síndrome, anteriormente descritas (CRUZ et al., 1995; FERRAZ et al., 2000) foram reavaliadas. A avaliação, de forma resumida, está apresentada na Tabela 1. Novas características foram também relatadas, como a escassez de tecido subcutâneo e a formação de coxim gorduroso na região lombar.

Tabela1. Critérios Clínicos da SAM nos indivíduos afetados.

Indivíduos Portadores da SAM	I-1	II-1	II-3
Características gerais			
Sexo	Masculino	Feminino	Feminino
Idade da última consulta	41 anos	21 anos	16 anos
Altura	162,0cm (↓P3)	153,5cm (P3)	148,0cm (↓P3)
Peso	60,6kg (P10-P25)	46,2kg (P5-P10)	46,8kg (P10-P25)
Perímetro cefálico	57,0cm (P90)	53,6cm (P10-P25)	51,7cm (P3)
Índice cefálico	0,87cm	0,8cm	0,8cm
Distância intercantar externa	10,0cm (↑P97)	8,5cm (P25-P50)	8,2cm (P25-P50)
Distância intercantar interna	3,6cm (P90-P97)	3,5cm (P90-P97)	3,6cm (P90-P97)
Comprimento da mão direita	17,5cm (P3-P25)	15,8cm (P3-P25)	15,6cm (P3-P25)
Dedo médio direito	8,0cm (P50)	5,6cm (↓P3)	6,5cm (↓P3)
Menarca	.	sim(regular)	sim(regular)
Telarca	.	Não	Não
Pubarca	Não	Não	Não
Características Faciais			
Orelha externa rudimentar	Sim	Sim	Sim
Microblefaria	Não	Sim	Sim
Sobrancelhas rarefadas	Sim	Não	Não
Sobrancelhas arqueadas	Sim	Sim	Sim
Cílios usentes/irregulares	Sim	Sim	Sim
Coloboma palpebral	Sim	Não	Não
Ponte nasal baixa	Sim	Sim	Sim
Narinas hipoplásicas e antevertidas	Sim	Sim	Sim
Macrostomia	Não	Sim	Sim

Tabela1. Critérios Clínicos da SAM nos indivíduos afetados (Continuação).

Indivíduos Portadores da SAM	I-1	II-1	II-3
Pele, subcutâneo, cabelo e unhas			
Pele grossa	Não	Sim	Sim
Pele seca	Não	Sim	Sim
Pele redundante	Não	Sim	Sim
Ausência de Lânugo ao nascimento	Não	Sim	Sim
Cabelo esperso	Não	Sim	Sim
Unhas hipoplásicas	Não	Sim	Sim
Escassez de tecido subcutâneo	Não	Sim	Sim
Coxim gorduroso em região lombar	Não	Sim	Sim
Região torácica			
Mamilos ausentes ou rudimentares	Não	Sim	Sim
Onfalocele	Não	Não	Sim
Sistema nervoso central			
Atraso no desenvolvimento	Não	Não	Sim
Atraso de linguagem	Não	Não	Sim
Outras características			
Baixa estatura	Sim	Sim	Sim
Hipoplasia dos lábios maiores	.	Sim	Sim
Ânus anteriormente localizado	Não	Sim	Sim
Atenuação distal de falanges	Não	Sim	Sim
Camptodactilia/Pseudocamptodactilia	Não	Sim	Sim

As características típicas da síndrome estão claramente presentes na filha mais velha (II-1) e também na filha mais nova (II-3), tendo o pai (I-1) um fenótipo mais leve da síndrome. A mãe (I-2) e a segunda filha (II-2) do casal não apresentam as características clínicas da doença (Figura 3). Os cariótipos das duas meninas portadoras da síndrome e dos pais haviam sido previamente realizados, visando à detecção de anormalidades citogenéticas, mas nenhuma alteração foi encontrada (FERRAZ et al., 2000).

Figura 3. Heredograma da família em estudo.

4.2. Dados do Sequenciamento

O sequenciamento do exoma de todos os familiares obteve cobertura média geral semelhante entre os indivíduos, de 23,68 vezes para cada base sequenciada, ou seja, cada região do exoma foi coberta por aproximadamente 23 reads (Tabela 2). É possível verificar o número de reads mapeadas em regiões únicas em relação ao genoma de referência, e também as reads mapeadas em posições únicas dentro das regiões alvo do exoma. O kit utilizado nesse sequenciamento permitiu a detecção de variantes localizadas em regiões adjacentes aos éxons. Assim, também pode ser observado o número de reads mapeadas nas regiões alvo acrescidos de 150 bases nucleotídicas antes e depois do sítio alvo. Cada exoma obteve a cobertura aproximada de 94% pelo menos uma vez. Quando contabilizamos 10 vezes de cobertura esse número cai para 67% do exoma, e, com 20 vezes, 44% do exoma é coberto (Tabela 2).

Tabela 2. Dados do Sequenciamento do Exoma.

Exomas Sequenciados	I-1	I-2	II-1	II-2	II-3
Número de reads únicas mapeadas	34.011.832	38.939.861	35.054.495	36.567.476	30.179.178
Número de reads únicas pro alvo +/-150pb*	21.074.221	23.416.143	21.619.479	22.380.849	18.941.469
Número de reads únicas pro alvo	18.856.939	20.815.589	19.283.119	20.002.037	16.928.718
Bases mapeadas com a cobertura >1x	93,74%	94,43%	93,97%	94,04%	93,47%
Bases mapeadas com a cobertura >5x	80,95%	84,16%	82,58%	82,69%	79,90%
Bases mapeadas com a cobertura >10x	65,61%	70,14%	67,74%	68,17%	63,42%
Bases mapeadas com a cobertura >20x	43,50%	47,92%	45,24%	46,43%	40,10%
Cobertura media	23,38X	25,58X	23,77X	24,75X	20,93X

* Número de reads únicas pras regiões alvo (éxons) acrescido de 150pb nas regiões posições flaqueadoras à esquerda (-150) e à direita (+150).

4.3. Dados e Filtros de Seleção das Mutações

O número total de variantes anotadas, geradas pelo sequenciamento, foi de aproximadamente 22.239 variantes por indivíduo. Essas alterações foram filtradas para a seleção de variantes novas, resultando em média 217 mutações por indivíduo (Tabela 3).

Tabela 3. Dados das mutações anotadas.

Exomas Sequenciados	I-1		I-2		II-1		II-2		II-3	
Tipos de Mutações	Descritas	Novas	Descritas	Novas	Descritas	Novas	Descritas	Novas	Descritas	Novas
Todas as mutações	21.497	242	22.761	206	22.233	230	22.458	204	21.160	206
Heterozigotas	13.394	233	14.570	203	14.164	227	14.447	199	13.411	202
Homozigotas	8.103	9	8.191	3	8.169	3	8.011	5	7.749	4
Indels <i>frameshift</i>	75	14	85	6	83	12	79	13	74	9
Heterozigotas	39	14	41	6	40	12	44	12	42	9
Homozigotas	36	0	44	0	43	0	35	1	32	0
Indels não <i>frameshifts</i>	131	5	143	14	118	10	135	11	126	9
Heterozigotas	71	5	94	14	65	10	70	11	82	8
Homozigotas	60	0	49	0	53	0	65	0	44	1
SNVs sinônimas	10.133	62	10.606	55	10.409	68	10.501	48	9.920	60
Heterozigotas	6.417	60	6.940	54	6.759	66	6.904	47	6.347	58
Homozigotas	3.716	2	3.666	1	3.650	2	3.597	1	3.573	2
SNVs não sinônimas	8.621	123	9.172	98	8.913	100	9.061	99	8.603	97
Heterozigotas	5.351	119	5.804	98	5.671	99	5.781	98	5.467	97
Homozigotas	3.270	4	3.368	0	3.242	1	3.280	1	3.136	0
SNVs stoploss	7	0	6	0	9	0	9	0	5	0
Heterozigotas	2	0	3	0	6	0	6	0	2	0
Homozigotas	5	0	3	0	3	0	3	0	3	0
SNVs stopgain	74	5	71	1	74	1	66	1	69	1
Heterozigotas	70	5	56	1	64	1	58	1	61	1
Homozigotas	4	0	15	0	10	0	8	0	8	0
Sítios de Recombinação	2.084	30	2.300	32	2.244	37	2.229	31	1.997	29
Heterozigotas	1.223	27	1.414	30	1.330	37	1.369	29	1.190	28
Homozigotas	861	3	886	2	914	0	860	2	807	1
Variantes desconhecidas	372	3	378	0	383	2	378	1	366	1

*Indels = Inserções e/ou deleções; Frameshift = troca do quadro de leitura; SNVs= Variantes de Nucleotídeo Único; Stopgain = ganho de códon de parada; Stoploss = perda de códon de parada.

As variantes novas e não sinônimas, de cada um dos indivíduos afetados, foram comparadas entre si visando à seleção de alterações em comum entre os três familiares afetados, uma vez que a investigação foi realizada assumindo um modelo de herança autossômica dominante, como sugerido anteriormente (FERRAZ et al., 2000; ROHENA et al., 2011). Essa seleção reduziu para 25 o número de possíveis variantes ligadas à síndrome. As mutações presentes nos membros não afetados da família foram descartadas, o que permitiu a eliminação de nove variantes, resultando em 16 variantes possíveis.

A análise das mutações localizadas em sítios de recombinação permitiu a eliminação de quatro variantes. Essas variantes foram descartadas por apresentarem baixos escores no RegulomeDB (Tabela 4), resultando em 12 variantes (Figura 4). Essas 12 mutações selecionadas foram ainda ranqueadas de acordo com a análise realizada no software SPRING, que leva em consideração escores de patogenicidade e informações de bancos de dados sobre os genes em que essas mutações se encontram (Tabela 5).

As doze mutações selecionadas são correspondentes a 11 genes, uma vez que duas mutações encontram-se no mesmo gene. A função e as doenças relacionadas a estes genes foram descritos por meio de informações já conhecidas em bancos de dados como OMIM, NCBI, Orphanet e Genecards (Tabela 6).

Tabela 4. Dados das Mutações em Sítios de Recombinação e Escore no RegulomeDB.

Gene	Crom.	Posição	Referência	Alterado	Mutação	Escore RegulomeDB
KANSL1	17	44302737	G	A	Sítio de Recombinação	2a
CHCHD3	7	132766826	G	A	Sítio de Recombinação	2b
TLE3	15	70356813	T	C	Sítio de Recombinação	4
MAGEB2	X	30236683	G	C	Sítio de Recombinação	*

* Não apresentou escore.

Figura 4. Fluxograma de filtros utilizados para seleção de variantes.

Tabela 5. Ranqueamento das Mutações Seleccionadas, por meio da ferramenta SPRING.

Ranque	Gene	Crom.	Posição	Referência	Alterado	Troca aa.	Tipo de Mutação	Valor de P	Valor de Q
1º	TWIST2	2	239757079	G	A	E75K	Não sinônima	6,55E-05	4,17E-04
2º	ADAT1	16	75652495	C	A	R94S	Não sinônima	6,94E-05	4,17E-04
3º	BAIAP3	16	1396459	G	C	R872P	Não sinônima	4,05E-04	1,62E-03
4º	IFT140	16	1575955	G	A	R901C	Não sinônima	1,36E-03	3,22E-03
5º	PKD1	16	2158345	C	T	V2275M	Não sinônima	1,41E-03	3,22E-03
6º	BAIAP3	16	1396312	G	A	D850N	Não sinônima	1,70E-03	3,22E-03
7º	NEO1	15	73542053	T	G	L629V	Não sinônima	1,88E-03	3,22E-03
8º	SLC39A12	10	18292247	T	G	V636G	Não sinônima	5,53E-03	8,30E-03
9º	ZNF354A	5	178139873	T	G	S336R	Não sinônima	4,06E-02	5,41E-02
10º	C8orf37	8	96281399	C	A	E7X	SNV Stopgain	9,44E-02	1,13E-01
11º	RIF1	2	152320093	C	G	D1353E	Não sinônima	1,23E-01	1,34E-01
12º	RANBP2	2	109367807	G	A	R454Q	Não sinônima	4,07E-01	4,07E-01

Valor de p e q : calculado pelo software SPRING, com base no teste de Fisher. **Legendas.** Crom.: Cromossomo; Troca de aa.: Troca de aminoácido; Stopgain : ganho do códon de parada; Stoploss: perda do códon de parada.

Tabela 6. Análise de Genes Candidatos.

Gene	Função Gênica	Doença Associada
TWIST2	Codifica um fator de transcrição envolvido na regulação do desenvolvimento mesodérmico durante a embriogênese. Também atua na diferenciação de várias linhagens celulares, como: musculares, cartilaginosas, osteogênicas, adipogênicas e células mielóides. Na pele, é precocemente expresso nas camadas mesenquimais indiferenciadas da epiderme que irão se desenvolver em derme (LI; CSERJESI; OLSON, 1995; SOSIC et al., 2003).	Displasia Focal Facial Dermal Tipo III e alguns tipos de câncer (FRANCO et al., 2011a; FRANCO et al., 2011b).
ADAT1	Codifica um RNA transportador específico de deaminação da adenosina, o que converte a adenosina na posição 37 do RNA transportador em inosina (MAAS; GERBER; RICH, 1999; MAAS; KIM; RICH, 2000; MAAS; KIM; RICH, 2001; MACBETH et al, 2005).	Sem registro.
BAIAP3	Codifica uma proteína que interage com o domínio citoplasmático do inibidor da angiogênese cérebro específico (BAI1), que é um membro da família de receptores da secretina. É altamente expresso no cérebro e pouco expresso em testículos (NAGASE et al, 1998; SHIRATSUCHI et al, 1998) .	Sem registro.
IFT140	Codifica uma subunidade do complexo de transporte intraflagelar, que está envolvida na gênese, reabsorção e sinalização dos cílios primários. Esses cílios são responsáveis por receber sinais do ambiente e transduzir para o núcleo (MUKHOPADHYAY et al., 2010; VINAYAGAM et al, 2011).	Ciliopatias (doenças renais), que podem ter manifestações extrarrenais, como: degeneração da retina, alterações cardíacas, anomalias do esqueleto e alterações do sistema nervoso central e periférico (PERRAUT et al., 2012; SCHMIDTS et al., 2012; JONASSEN et al., 2012; KHAN; BOLZ; BERGMANN, 2014).
PKD1	Codifica uma proteína integral de membrana que regula múltiplas vias de manutenção do desenvolvimento normal da estrutura renal tubular. Está envolvido na permeabilidade de cálcio, e na homeostase intracelular de cálcio (KEITH et al, 1987; JEFFERY et al, 1998; HANAOKA et al, 2000)	Doença Renal Policística, tipo I (SONG et al, 2009).

Tabela 6. Análise de Genes Candidatos (continuação).

Gene	Função Gênica	Doença Associada
NEO1	Codifica uma proteína de superfície celular e é um membro da superfamília das imunoglobulinas. Está envolvido no desenvolvimento normal do sistema nervoso. Esta proteína pode ainda estar envolvida no crescimento celular, diferenciação e adesão célula-célula em diversos tipos de câncer (MATSUNAGA et al, 2004; WILSON; KEY, 2007).	Está relacionado tipos de câncer como: mama (LEE et al., 2005) e meduloblastoma (MILLA et al, 2014).
SLC39A12	Codifica uma proteína com estrutura característica de transportador de zinco. Essa proteína possui um importante papel no transporte de zinco durante o desenvolvimento normal do sistema nervoso de vertebrados (CHOWANADISAI, 2013).	Sem registro.
ZNF354A	Codifica uma proteína do tipo zinc finger, expressa principalmente em tecido renal e durante a espermatogênese (WITZGALL et al, 1993; OMORI et al, 1997; PRZYBORSKI et al, 1998).	Tumor testicular (PRZYBORSKI et al, 1998).
C8orf37	Codifica uma proteína com função desconhecida, mas que está localizada na base de cílios primários em células epiteliais de pigmento da retina. É altamente expressa no cérebro, coração e na retina (ESTRADA-CUZCANO et al, 2012).	Distrofia de cones e bastonetes e Retinose Pigmentar (ESTRADA-CUZCANO et al, 2012).
RIF1	Codifica uma proteína de reparo, normalmente encontrada em telômeros aberrantes. Está também relacionada ao controle do tempo de replicação, embora deleções deste gene não inviabilizem a transcrição (XU; BLACKBURN, 2004; CORNACCHIA et al, 2012; YAMAZAKI et al, 2012; DI VIRGILIO et al, 2013).	Sem registro.
RANBP2	Codifica uma proteína componente do complexo de poro nuclear, que possui papel principal na facilitação da importação e exportação de proteínas, sumoilação de proteínas, tráfego intracelular e manutenção de energia (VETTER et al, 1999; PICHLER et al, 2002; SAITOH et al., 2012).	Encefalopatia Necrotizante Aguda e Tumor Miofibroblástico Inflamatório (NEILSON et al., 2009; LI et al, 2013).

Dentre os genes analisados, o gene TWIST2 se destacou na análise realizada pelo software SPRING e também na relação genótipo-fenótipo. A mutação no gene TWIST2 ocorreu no primeiro éxon, na posição nucleotídica 223, com a troca de uma guanina por uma adenina. O aminoácido codificado pela trinca de nucleotídeos correspondente também foi alterado (Figura 3), com a troca de um ácido glutâmico na posição 75 da proteína, por uma lisina (c.223G>A p.E75K). Essa mutação foi coberta no sequenciamento de nova geração por 39 reads para I-1, 16 reads para II-1 e 20 reads para II-3. Visando ainda confirmar a autenticidade da mutação na família, foi realizado o sequenciamento tradicional (Sanger) e foi confirmada a presença da mutação em dois membros afetados (II-1 e II-3) e a ausência nos membros não afetados (I-2 e II-2). Porém, o membro I-1 cuja mutação foi detectada no sequenciamento de nova geração, não foi amplificado pelo sequenciamento Sanger, e pode ser observado o nucleotídeo referência G (guanina), em destaque (Figura 3).

Figura 5. Cromatogramas da validação da mutação (c.223G>A) no gene TWIST2. Destaque para mutação em heterozigose para os pacientes, II-1 e II-3, e o nucleotídeo referência para os pacientes I-1, I-2 e II-2.

4.4. Avaliação *in silico* do efeito da Mutação na Proteína

A modelagem *in silico* da proteína Twist2, com ambas as ferramentas HOPE e SWISSMODEL, sugerem que a mutação (c.223G>A) afeta o resíduo E75 e pode alterar a atividade de ligação de Twist2 ao DNA, mas não altera a estrutura global da proteína (Figura 6). O modelo da proteína selvagem pelo SWISSMODEL obteve um escore QMEAN maior que o da proteína mutante, -0.15 e -1.59, respectivamente. Os dados de similaridade mostraram que a sequência selvagem da proteína possui 52,56% de homologia com a proteína molde utilizada na construção da estrutura 3D, e que a sequência da proteína mutante possui 29,41% de similaridade.

Figura 6. Estruturas tridimensionais das proteínas selvagem e mutante, obtidas com a ferramenta SWISSMODEL. A: Proteína Tipo selvagem Twist2 (Q8WVJ9, extraída do Uniprot). B: Proteína mutante (p.E75K).

Dados obtidos com o HOPE mostram que o aminoácido mutante Lisina é maior que o original, e é carregado positivamente, enquanto o aminoácido referência Ácido Glutâmico é carregado negativamente (Figura 7). Outra representação da estrutura 3D da proteína, e o enfoque da troca de aminoácidos são apresentados (Figura 8).

Figura 7. Representação da troca do aminoácido referência, o ácido glutâmico, pelo aminoácido lisina. Troca decorrente da mutação c.223G>A p.E75K encontrada nos três indivíduos com fenótipo da SAM.

Figura 8. Modelo tridimensional da proteína Twist2, gerado pela ferramenta HOPE. Enfoque para a região onde ocorre a troca de aminoácidos na posição E75, com destaque para representação do aminoácido selvagem (ácido glutâmico) em verde, e o mutante (lisina) em vermelho.

5. DISCUSSÃO

5.1. Reavaliação clínica

As características clínicas da SAM apresentadas pelos pacientes da família em estudo são típicas do quadro clássico da doença, e já foram descritos por nosso grupo (CRUZ et al., 1995; FERRAZ et al., 2000), os detalhes são relatados na Tabela 1. Durante a reavaliação clínica foi observada como característica nova, a perda de tecido adiposo e o acúmulo de gordura na região lombar, características comuns em quadros de lipodistrofia. Lipodistrofia é um termo genérico utilizado para descrever um grupo heterogêneo de desordens metabólicas (FIORENZA et al., 2011), cuja característica principal envolve o padrão de distribuição da perda de tecido adiposo subcutâneo, podendo ser parcial ou generalizada. Outras características podem também estar associadas à lipodistrofia, como resistência à insulina, diabetes mellitus, esteatose hepática e dislipidemia (CHAN; ORAL, 2010; GARG, 2011).

A classificação da lipodistrofia em subtipos leva em consideração o padrão parcial ou generalizado da perda de gordura, e o quadro genético ou adquirido. De forma simplificada, existem quatro subtipos principais de lipodistrofia: lipodistrofia congênita generalizada (LCG), lipodistrofia adquirida generalizada (LAG), lipodistrofia familiar parcial (LFP) e lipodistrofia adquirida parcial (LAP) (HANDELSMAN et al., 2013). Neste estudo, as duas filhas com SAM apresentam quadros similares de perda de gordura subcutânea e de acúmulo de gordura na região lombar, o que sugere um quadro de lipodistrofia parcial associado a SAM. Os últimos exames de glicemia em jejum, e de enzimas hepáticas foram normais para as duas filhas, não apresentando assim outros distúrbios metabólicos associados à lipodistrofia. O pai não possui esses exames, sendo necessária a verificação dos mesmos.

Informações presentes no IMPC (<http://www.mousephenotype.org/>), mostram modelos de camundongos para doenças associadas a genes, com base no algoritmo PhenoDigm, que calcula um percentual de similaridade, onde o melhor modelo de camundongo existente que esteja relacionado à doença tem um escore de 100% (BROWN; MOORE, 2012). O gene TWIST2, sugerido neste estudo como causador da SAM, foi submetido à busca dentro deste banco de fenótipos e mostrou associação com algumas síndromes. Dentre essas, a Lipodistrofia Familiar Parcial (escore 79,06%) e a Displasia Focal

Facial Dermal tipo III, também chamada Síndrome de Setleis (escore 67,07%). A Síndrome de Setleis possui algumas características sobrepostas a SAM, como ausência de cílios em ambas as pálpebras, múltiplos defeitos na pálpebra superior, sobrancelhas arqueadas, ponte nasal rasa, nariz bulboso (SETLEIS et al., 1963). Essas informações sugerem que o quadro de lipodistrofia apresentado pelos pacientes neste estudo, pode estar relacionado à mutação encontrada no gene TWIST2, e consequentemente a vias associadas a SAM. Porém, o aprofundamento na investigação de como essa característica está realmente relacionada a SAM é necessária.

5.2. Dados do Sequenciamento de Nova Geração

Meynert e colaboradores (2013) afirmam que para identificar pelo menos 90% das variantes de nucleotídeo único heterozigotas, se faz necessária de 17 a 37 vezes de cobertura média. Contudo, não há um número fixo de vezes que todo o exoma precisa ser sequenciado para detectar de forma eficiente as variantes nucleotídicas. Nossos dados mostram que com a cobertura média do exoma de 23 vezes foi possível detectar até mesmo variantes raras, e que, apesar dessa cobertura não garantir que 100% do exoma tenha sido lido, pode ser considerado um bom parâmetro para detecção de variantes.

Clark e colaboradores (2011) inferem que para alcançar pelo menos 10 vezes de cobertura em cada base sequenciada, em 90% ou mais do exoma, é necessário 80 vezes de cobertura média de todo exoma. Nossos dados permitem verificar que com 10 vezes de cobertura para cada base, 67% do exoma foi coberto, mostrando que provavelmente são necessários menos de 80 vezes de cobertura média para alcançar 90% do exoma. Outro estudo mais recente de Meynert e colaboradores (2014) afirmam ser necessárias 40 vezes de cobertura média para atingir a mesma sensibilidade dos dados de Clark et al. Esse decréscimo no número de vezes que o exoma precisa ser sequenciado para garantir uma boa sensibilidade na detecção das variantes, se deve em parte ao melhoramento na tecnologia de sequenciamento.

Meynert et al. (2014) afirmam ainda que para a detecção eficiente de variantes, o número de reads cobrindo uma mesma região deve ser de 2 a 3 para variantes homozigotas, e de 9 a 10 para variantes heterozigotas, e que esta cobertura garante aproximadamente 95% de

sensibilidade na detecção das variantes. Dessa forma, a maior parte das variantes encontradas está dentro desse valor ótimo de sensibilidade.

5.3. Seleção de Mutações

O número de mutações encontradas se manteve semelhante entre os pacientes, apresentando maior número as mutações sinônimas, não sinônimas e em sítios de recombinação. O número de mutações novas é sempre muito inferior ao número de mutações já conhecidas, representando em torno de 1% do total de mutações (Tabela 3). Nos Estados Unidos da América, uma doença é considerada rara quando existe menos de 200.000 indivíduos afetados (US FOODDRUG ADMINISTRATION, 1983) e na Europa quando afeta 1 a cada 2000 indivíduos (EUROPEAN COMISSION, 2008). A incidência da SAM é muito inferior, com 17 casos descritos no mundo, 14 deles bem documentados e três outros casos com diagnóstico pouco claro (SCHANZE et al., 2013). Assim, a seleção de variantes novas, implica na detecção dessas alterações com incidência muito baixa na população.

A busca por mutações presentes nos três membros afetados da família, que não fossem comuns aos membros não afetados, foi realizada com base em evidências de um padrão de herança autossômico dominante (FERRAZ et al., 2000; ROHENA et al., 2011). Essas duas etapas de seleção possibilitaram o descarte de um grande número de mutações, facilitando a busca pela mutação responsável pela SAM.

Outro passo importante durante a seleção das variantes foi a exclusão das mutações que estivessem presentes em sítios de recombinação. As mutações nesses sítios foram avaliadas por meio da ferramenta RegulomeDB, que atribuiu escores às mutações, com valores de intermediário a alto (2a, 2b, 4), tendo ainda uma mutação que não gerou escore. A amplitude dos escores varia de 1 a 6, tendo os escores mais baixos (1a, 1b, 1c, 1d, 1e, e 1f) maior probabilidade de atuar em algum processo importante de regulação da expressão gênica, uma vez que existem dados de expressão já relacionados às mutações. As classes seguintes (de 2 a 6) representam anotações análogas as da categoria 1, porém como não possuem dados de expressão, não há efeitos conhecidos diretos para suas regiões de ligação (BOYLE et al., 2012).

A mutação em sítio de recombinação que não gerou escore no RegulomeDB, pode ser decorrente do fato de não existirem dados de regulação de expressão gênica para a região, ou por falta de descrição de variantes de nucleotídeo único nessa posição. Porém, essa alteração

foi ainda visualizada no IGV, e foi verificado que a mutação não está nos nucleotídeos conservados, de cada lado do íntron (GT/AG), sítios importantes para o processo de recombinação alternativa (SNUSTAD; SIMMONS, 2008). Outro fator relevante para a exclusão dessas mutações em sítios de recombinação está relacionado ao fenótipo da síndrome, que sugere um sítio de regulação com grande importância para a formação da proteína, consequentemente com escore mais impactante do que os observados nas mutações avaliadas.

As mutações restantes desses processos de seleção foram ranqueadas pelo software SPRING, visando estimar a significância de que a mutação seja patogênica e causal para a doença em estudo, a eficiência da ferramenta já foi comprovada na identificação de variantes novas para casos de autismo, epilepsia, encefalopatias e deficiência intelectual (WU; LI; JIANG, 2014). Um estudo recente de Yang e colaboradores (2015) também fez uso da ferramenta SPRING para ranquear as mutações encontradas em uma família com a cardiopatia “não compactação do ventrículo esquerdo”, e puderam determinar a provável variante causal da doença.

Os dados do ranqueamento (Tabela 5), obtidos pelo SPRING, sugerem que o gene TWIST2 possui a maior probabilidade de ser a mutação causal da SAM, justamente por apresentar um baixo valor de q , e consequentemente uma menor taxa falso-positiva, um indicativo de maior significância da relação com SAM. O gene ADAT1 apresentou o mesmo valor de q que o gene TWIST2, mas foi ranqueado em segundo lugar por apresentar maior valor de p . Os genes ZNF354A, C8orf37, RIF1 e RANBP2, ranqueados de 9º a 12º, não apresentaram valor de p e q significativo, um forte indicativo de não estarem realmente relacionados a SAM.

Além do ranqueamento, todas as doze variantes foram analisadas quanto à função (Tabela 6). Os genes ranqueados da 9ª a 12ª posição, anteriormente citados por não possuir relevância significativa na análise com o SPRING, foram descartados, justamente por não haver relação de mutações destes genes com características fenotípicas encontradas na síndrome em estudo. Os seis genes com ranque intermediário (3ª à 8ª posição), que apresentaram um valor de p e q ainda significativos, também não possuíam função que pudesse relacioná-los a aspectos fenotípicos da síndrome. O gene ADAT1, que havia obtido o mesmo valor de q do gene TWIST2, foi também descartado, uma vez que esse gene codifica um RNA (ácido ribonucleico) transportador específico de deaminação da adenosina, e não há alterações fenotípicas ou doenças associadas a mutações nesse gene (MAAS; KIM; RICH, 2001; MACBETH et al, 2005).

TWIST2 é o único gene segregando nos indivíduos afetados da família que codifica uma proteína envolvida na regulação do desenvolvimento da derme (LI; CSERJESI; OLSON, 1995; SOSIC et al., 2003), tecido fundamentalmente alterado na SAM. O gene TWIST2 foi descrito com mutações em homozigose, como o responsável pela Síndrome de Setleis, que possui características sobrepostas a SAM, como a ausência de cílios, sobrancelhas arqueadas, alterações nasais, entre outras. Mutações em heterozigose em familiares de membros com a Síndrome de Setleis também mostraram alterações de fenótipo como ausência parcial de cílios nas pálpebras inferiores e superiores e distiquíase (CERVANTES-BARRAGÁN et al., 2011). A associação do fenótipo da Síndrome de Setleis com SAM sugere a importância desse gene para a determinação do fenótipo na síndrome em estudo.

5.4. Validação da Mutação Seleccionada e Variabilidade Intrafamiliar

A mutação (c.223G>A) no gene TWIST2 detectada pelo sequenciamento de nova geração nos três membros afetados apresentou boa cobertura, o que garante alta sensibilidade de detecção e um menor risco de gerar variantes falso-positivas. Apesar da alta cobertura da variante nos pacientes afetados, a mesma foi ainda verificada por Sequenciamento Sanger, e foram validadas as mutações nas duas filhas afetadas, comprovando a veracidade da mutação. Porém, no pai com as características mais leves da síndrome não foi confirmada a mutação pelo Sequenciamento Sanger, sugerindo a possibilidade de mosaicismos, o que justificaria ainda a variabilidade intrafamiliar da SAM na família em estudo.

O mosaicismos é caracterizado pela ocorrência de duas linhagens celulares distintas em um mesmo indivíduo, e é uma das causas da variabilidade clínica entre indivíduos que apresentam a mesma mutação (STRACHAN; READ, 1999). O mosaicismos pode ser detectado por meio da análise quantitativa de alelos mutantes em diversos tecidos, como sangue, fibroblastos e saliva nos pacientes afetados. Estudos recentes como de Donkervoolt e colaboradores (2015) demonstraram o mosaicismos como causa da variabilidade intrafamiliar, e sugerem que o mosaicismos esporádico e o herdado possam ser mais comuns do que se espera. Assim, sugere-se aqui a importância da investigação de mosaicismos na família deste estudo como potencial causa da variabilidade intrafamiliar.

5.5. O Gene TWIST2

O gene TWIST2 possui dois éxons separados por um longo íntron. Apenas o primeiro éxon é responsável pela codificação de toda a proteína (CHEN; BEHRINGER, 1995; PERRIN-SCHMITT et al., 1997). Tukel e colaboradores (2010) mapearam esse locus gênico no braço longo do cromossomo 2, posição 2q37.3. São descritos dois transcritos codificadores dessa mesma proteína, um de 1.434 pares de bases e outro de 1.186 pares de bases, a diferença entre os transcritos está na região 3'UTR.

A estrutura da proteína Twist2 contém 160 aminoácidos (Figura 4), que incluem dois sinais de localização nuclear (SLN) na porção amino-terminal, um do aminoácido 29 a 32, e outro do aminoácido 52 a 56. Um domínio básico, da posição 66 a 77, um domínio hélice-alça-hélice (HLH) da posição 78 a 118, e um domínio repressor TWIST BOX, do aminoácido 141 a 160 na porção carboxi-terminal (LEE et al., 2000; GONG; LI, 2002).

Figura 9. Estrutura básica da proteína Twist2. Representação dos sinais de localização nuclear em vermelho (SLN1 e SLN2), do domínio básico de ligação em amarelo, do domínio HLH, em verde, e do domínio TWIST BOX em azul. (Adaptado de GONG; LI, 2002).

O gene TWIST2 codifica um fator de transcrição pertencente à classe B da família das proteínas bHLH (do inglês, *basic Helix-Loop-Helix*), que é altamente conservada entre espécies. Essa família de proteínas engloba vários fatores de transcrição, que em geral apresentam um complexo e refinado modo de ação, controlando o destino celular nos primeiros estágios da embriogênese (FRANCO et al., 2011a).

As proteínas bHLH se ligam ao DNA por meio da formação de complexos homodímeros ou heterodímeros de suas regiões HLH, permitindo que os aminoácidos da

região básica se liguem diretamente ao DNA. O domínio de ligação ao DNA é capaz de reconhecer elementos *cis* regulatórios, contendo a sequência consenso 5'-NCANNTGN-3', denominada região E-box (MURRE et al., 1989). O domínio bHLH inclui resíduos altamente conservados de treonina e serina, cuja fosforilação tem demonstrado afetar a formação de dímeros e a afinidade de ligação ao DNA em outros membros da família das proteínas Twist (FIRULLI et al., 2005; FIRULLI; CONWAY, 2008).

A proteína Twist2 foi primeiramente identificada na análise de interação proteína-proteína, utilizando como sonda a proteína E12, pertencente à classe A dessa mesma família bHLH (STAUDINGER et al., 1993). Twist2 foi inicialmente denominada Dermo-1, devido ao seu padrão de expressão na derme de embriões de camundongos (LI et al., 1995). Porém, passou a se chamar Twist2 por sua grande homologia e sobreposição no padrão de expressão com a proteína Twist1 (LEE et al., 2000; SOSIC et al., 2003). Twist2 possui 66% de identidade com a proteína Twist1, essa similaridade chega a 98% quando comparadas as regiões bHLH (TUKEL et al., 2010).

Franco e colaboradores (2011a) em um trabalho de revisão sintetizam os modos de ação da proteína Twist2 em: ligação direta ao DNA, por meio das sequências E-box; recrutamento de coativadores e repressores; sequestro de proteínas E moduladoras; e interrupção da própria ação de repressão/ativação por meio de interações proteína-proteína. Sendo que a ativação e repressão de diversos genes controlados pelo gene TWIST2, é dependente de modificações pós-traducionais, interação com outras moléculas e contexto celular.

O padrão de expressão do gene TWIST2 sugere que ele atue como regulador da expressão gênica em um subconjunto de linhagens celulares mesenquimais, incluindo no desenvolvimento da derme. Transcritos de TWIST2 foram detectados no dermatomo, esclerotomo, brotos dos membros e arcos branquiais em camundongos. Há aumento na expressão da proteína Twist2 no desenvolvimento mesodérmico, durante a embriogênese. É precocemente expresso nas camadas mesenquimais indiferenciadas da epiderme que irão se desenvolver em derme (LI; CSERJESI; OLSON, 1995; SOSIC et al., 2003).

Sharabi e colaboradores (2008) descreveram o papel fundamental de Twist2 no desenvolvimento de linhagens mielóides, assim como na função de resposta inflamatória de células mielóides maduras. A ação conjunta de Twist2 e Twist1 inibe a diferenciação muscular (SPICER et al., 1998; GONG; LI, 2002), e regula a diferenciação de osteoblastos,

sendo a participação de Twist2 nessa regulação temporalmente após a do gene Twist1 (LEE et al., 2000).

Os genes TWIST1 e TWIST2 estão também frequentemente ativados em cânceres humanos, e tem demonstrado anular a senescência induzida por oncogenes, por meio da inibição das vias supressoras de tumor como p53 e Rb, enquanto ao mesmo tempo promovem a Transição Epitélio Mesênquima (TEM) (ANSIEU et al., 2008). Estes achados desafiam a visão clássica de que as vias de senescência e TEM são reguladas independentemente, já que a superexpressão de TWIST1 e TWIST2 em neoplasias possibilita a regulação simultânea dessas vias (SMIT; PEEPER, 2008). A expressão aumentada das proteínas Twist facilita a TEM, e a expressão dessas proteínas tem sido indicadores de pior prognóstico em pacientes com câncer (KWOK et al.; 2005, KYO et al., 2006).

5.6. Mutações no Gene TWIST2

Sosic e colaboradores (2003) realizaram estudos com camundongos homozigotos para o alelo nulo Twist2 e com duplo heterozigotos para alelos Twist1 e Twist2. Ambos apresentaram atrofia de múltiplos tecidos e alta expressão de citocinas pró-inflamatórias, resultando em morte perinatal por caquexia. Esses dados sugerem que a inibição de citocinas seja sensível à dosagem das proteínas Twist. Camundongos *knockout* para TWIST2 possuem anormalidades no desenvolvimento de vértebras, na cicatrização de feridas e na regeneração dos pelos.

Tukel et al., (2010) descreveram mutações homozigotas para o gene TWIST2 em pacientes com a Displasia Focal Facial Dermal Tipo III (Síndrome de Setleis), e visualizaram características similares dessa síndrome com camundongos *knockout* para o TWIST2. Este estudo mostra que embora Twist1 e Twist2 apresentem grande homologia, as mutações heterozigotas no gene TWIST1 causam outra síndrome, a Síndrome de Saethre-Chotzen, que cursa com craniossinostose, sugerindo que a função destes genes seja independente no desenvolvimento da pele e dos ossos.

A Síndrome de Setleis (OMIM 227260) é de herança autossômica recessiva, e compreende a terceira classe das displasias focais faciais dermais, foi inicialmente descrita em pacientes de famílias consanguíneas de Porto Rico (SETLEIS et al., 1963). Essa síndrome é caracterizada por marcas temporais bilaterais, aparência leonina, ausência de

cílios em ambas as pálpebras, múltiplos defeitos na pálpebra superior, ausência de glândulas meibomianas, sobrelanceas arqueadas, queixo em fenda, aumento da mobilidade da pele facial, ponte nasal rasa, nariz bulboso, entre outras (MARION et al., 1987; CLARK et al., 1989; TSUKAHARA et al., 1995).

As mutações c.324C>T e c.486C>T foram primeiramente descritas em homozigose no gene TWIST2, para Síndrome de Setleis em uma família de descendência árabe e outra de descendência porto riquenha, respectivamente. As duas mutações causam instabilidade e truncamento da proteína nas posições 65 e 119 da cadeia de aminoácidos (p.Q65X e p.Q119X) (TUKEL et al., 2010).

Cervantes-Barragán e colaboradores (2011) descreveram uma nova mutação (c.168delC) que causa troca de quadro de leitura (p.S57AfsX45), em uma família de ancestralidade mexicana-nahua composta por seis indivíduos, onde dois irmãos com as características clássicas da Síndrome de Setleis apresentaram a mutação em homozigose. Os outros dois irmãos e os pais apresentaram a mutação em heterozigose, sendo possível perceber um fenótipo mais leve nesses indivíduos, como a distiquíase nas pálpebras superiores e ausência parcial de cílios nas pálpebras inferiores.

Recentemente Rosti e colaboradores (2014) caracterizaram outra mutação nova em homozigose c.326T>C no gene TWIST2 (p.Leu109Pro), em três irmãos afetados com características da Síndrome de Setleis de uma família turca. Por meio de predição *in silico* da proteína, foi possível determinar que a troca da Leucina por uma Prolina na posição 109 da cadeia de aminoácidos alterou a estrutura tridimensional da proteína, tornando-a incapaz de formar dímeros e impedindo assim a ligação de Twist2 aos seus genes alvos do desenvolvimento facial.

5.7. Modelagem *in silico* da Proteína Twist2

No presente estudo, a mutação c.223G>A no gene TWIST2, foi detectada em heterozigose no pai e nas duas filhas afetadas pela SAM. Essa mutação causa a troca do aminoácido ácido glutâmico na posição 75 da proteína, pelo aminoácido lisina (p.E75K) (Figura 3), as consequências dessas alterações na proteína foram analisadas por meio de modelagem *in silico*.

A modelagem *in silico* realizada com a ferramenta SWISSMODEL construiu um possível modelo de estrutura 3D da proteína, baseado em um molde similar à sequência de aminoácido submetida para análise. Foi observado que a estrutura global da proteína (Figura 6) não sofreu alteração, mas a mutação causa a alteração no resíduo E75, que está localizado dentro do sítio básico de ligação ao DNA, o que sugere alteração na atividade de ligação de Twist2 ao DNA.

Na ferramenta SWISSMODEL a qualidade do modelo é baseada no escore QMEAN. Esse escore é construído por meio de características geométricas do modelo: ângulos de torção dos aminoácidos, pareamento de distâncias atômicas e acessibilidade do solvente (BIASINI et al., 2014). O modelo da proteína selvagem obteve um escore QMEAN maior que o da proteína mutante, -0.15 e -1.59, respectivamente. O que mostra que o modelo criado para a proteína selvagem possui maior confiabilidade, que o da proteína mutante. Porém ambos os valores são negativos, representativos de baixa qualidade para o modelo criado, dados que podem ser justificados pela falta de descrição da estrutura tridimensional da proteína, e uso de proteínas homólogas para a construção da estrutura tridimensional (3D) da proteína em estudo.

Sabidamente o domínio básico HLH é altamente conservado entre as espécies (Tabela 7), e essa diferença apresentada na similaridade (de 52,56% para 29,41%) com a troca de apenas um aminoácido reforça a importância do resíduo E75 na estrutura da proteína.

Tabela 7. Homologia da sequência de aminoácidos da proteína Twist2 entre espécies. A alta conservação do resíduo E (ácido glutâmico), onde a mutação foi encontrada está evidenciada.

Espécie	Resíduos conservados											
<i>Homo sapiens</i> NP_476527.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Pan troglodytes</i> XP_003309597.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Macaca mulatta</i> XP_001087543.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Canis lupus familiaris</i> XP_543311.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Mus musculus</i> NP_031881.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Rattus norvegicus</i> NP_067723.1	Q	S	Q	R	I	L	A	N	V	R	E	R
<i>Xenopus tropicalis</i> NP_001096679.1	H	T	Q	R	I	I	A	N	V	R	E	R
<i>Danio rerio</i> NP_001005956.1	Q	N	Q	R	V	L	A	N	V	R	E	R

A estrutura molecular 3D obtida por meio da ferramenta HOPE, mostra que a diferença de tamanho entre o resíduo selvagem e o mutante faz com que o novo resíduo não fique na posição correta para a formação da ponte de hidrogênio, já que é maior que o resíduo selvagem. A proteína Twist2 está envolvida na formação de complexos multiméricos com outras proteínas, dessa forma a alteração no tamanho do resíduo pode perturbar essas interações. O resíduo E75 está também envolvido em interações iônicas, importantes para multimerização. A mutação introduz um resíduo de carga oposta, o que pode não apenas causar perda de interação iônica, mas também causar repulsão com o parceiro de interação (Venselaar et al., 2010).

Como já descrito, a proteína Twist2 se liga aos motivos E-box (5'CANNTG3') do DNA, por meio da formação de complexos homodímeros ou heterodímeros de suas regiões HLH, permitindo que os aminoácidos da região básica se liguem diretamente ao DNA (MURRE et al., 1989). Mudanças dentro da região básica/HLH podem impedir a formação de heterodímeros ou a própria ligação ao DNA, impedindo que a proteína desempenhe sua função, e gerando alterações no fenótipo. A heterogeneidade de fenótipos associados a mutações no gene TWIST2, como a Síndrome de Setleis, pode ser atribuída às interações que essa proteína é capaz de formar, e a ampla ação regulatória que ela desempenha em diversos genes do desenvolvimento (LI; CSERJESI; OLSON, 1995; SPICER et al., 1998; GONG; LEE et al., 2000; LI, 2002; SOSIC et al., 2003; SHARABI et al., 2008).

Outros pacientes com a SAM devem ser analisados para mutações no gene TWIST2, a fim de validar que a mutação nesse gene é realmente responsável pela síndrome. O ensaio funcional da mutação encontrada também é necessário, para que a contribuição no fenótipo da síndrome possa ser determinada. Sugere-se ainda que as síndromes com sobreposição fenotípica a SAM, como a Síndrome de Barber-Say e a Síndrome de Fraser, sejam avaliadas para o gene TWIST2, pois o gene envolvido nessas síndromes pode ser o mesmo, com possíveis diferenças na posição onde a alteração se encontra.

6. CONCLUSÕES

A reavaliação clínica possibilitou a visualização de uma nova característica não descrita, a perda de tecido adiposo e o acúmulo de gordura na região lombar em duas pacientes, características comuns em quadros de lipodistrofia, o que sugere uma possível relação entre este quadro e a SAM.

O sequenciamento completo do exoma de cinco membros da família com SAM permitiu a detecção de 25 mutações novas não sinônimas em comum entre os membros afetados. O padrão de herança autossômico dominante, anteriormente sugerido, foi validado, e as 25 variantes foram filtradas, sendo descartadas as que estivessem presentes também nos membros não portadores da SAM, permanecendo 16 variantes. As mutações em sítios de recombinação, que não fossem essenciais para o processo de recombinação alternativa também foram descartadas, restando 12 mutações em potencial.

As 12 mutações selecionadas foram submetidas à análise de patogenicidade por meio de ferramentas *in silico*, que possibilitou o ranqueamento dessas em relação a SAM. A mutação no gene TWIST2 foi classificada com a maior probabilidade de estar associada à síndrome. As 12 mutações estavam presentes em 11 genes diferentes, e todos os genes foram revisados para avaliação de suas funções já descritas que pudessem relacioná-los a SAM. Novamente o gene TWIST2 se mostrou o melhor gene candidato, pela associação de sua função gênica com o fenótipo da SAM.

A mutação c.223G>A (p.E75K) identificada no gene TWIST2 foi validada em dois pacientes com SAM, e a troca de aminoácidos decorrente dessa mutação foi avaliada por meio da modelagem *in silico* da proteína. Foi possível verificar que a mutação causa a troca do resíduo ácido glutâmico pela lisina, localizada dentro do domínio conservado básico HLH, o que pode interferir na formação de heterodímeros, ou na própria ligação ao DNA, impedindo que a proteína desempenhe sua função, gerando alterações no fenótipo.

A heterogeneidade de fenótipos associados a mutações no gene TWIST2, pode ser atribuída às interações que essa proteína é capaz de formar, e a ampla ação regulatória que ela desempenha em diversos genes do desenvolvimento. Contudo, para que seja determinada a contribuição da mutação para a determinação da síndrome não somente na família em estudo, se faz necessária à análise do TWIST2 em outros pacientes com a SAM, assim como o estudo funcional da mutação durante o desenvolvimento embrionário.

A causa da variabilidade do quadro clínico dentro da família em estudo não foi determinada. Porém, sugere-se que o mosaicismo seja o responsável por essa variabilidade intrafamiliar, e que seja realizada uma avaliação quantitativa de alelos mutantes em diversos tecidos, como sangue, fibroblastos e saliva nos pacientes afetados.

Em suma, infere-se que a mutação c.223G>A p.E75K, com modo de herança autossômico dominante, constitua a base molecular da Síndrome Ablefaria Macrostomia.

7. REFERÊNCIAS

- ABUZENADAH, A.M.; ZAHER, G. F.; DALLOL, A.; et al. Identification of a novel SBF2 missense mutation associated with a rare case of thrombocytopenia using whole-exome sequencing. **Journal of Thrombosis and Thrombolysis**, v. 36, p. 501–506, 2013.
- AMOR, D.J.; SAVARIRAYAN, R. Intermediate form of ablepharon–macrostomia syndrome with CNS abnormalities. **American Journal of Medical Genetics Part A**, v. 103, p. 252–254, 2001.
- AZEVEDO, E.S.; BIONDI, J.; RAMALHO, L.M. Cryptophthalmos in two families from Bahia, Brazil. **Journal of Medical Genetics**, v. 10, p. 389–392, 1973.
- BAINBRIDGE, M.N.; WISZNIEWSKI, W.; MURDOCK, D.R.; et al. Whole-genome sequencing for optimized patient management. **Science Translational Medicine**, v.3, p. 87, re. 3, 2011.
- BAMSHAD, M. J.; NG, S.B.; BIGHAM, A.W.; et al. Exome sequencing as a tool for Mendelian disease gene discovery. **Nature Reviews Genetics**, v. 12, p. 745–755, 2011.
- BARBER, N.; SAY, B.; BELL, R.F.; MERVEILLE, O.C. Macrostomia, ectropion, atrophic skin, hypertrichosis, and growth retardation. **Syndrome Identification**, v. 8, p. 6-9, 1982.
- BEAULIEU, C.L.; MAJEWSKI, J.; SCHWARTZENTRUBER, J.; et al. FORGE Canada Consortium: Outcomes of a 2-Year National Rare-Disease Gene-Discovery Project. **American Journal of Human Genetics**, v. 94, n.6, p. 809-817, 2014.
- BIASINI, M.; BIENERT, S.; WATERHOUSE, A.; et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. **Nucleic Acids Research**, 42, 2014.
- BOYLE, A.P.; HONG, E. L.; HARIHARAN, M.; et al. Annotation of functional variation in personal genomes using RegulomeDB. **Genome Research**, v. 22, p. 1790–1797, 2012.
- BRANCATI, F.; MINGARELLI, R.; SARKOZY, A.; DALLAPICCOLA, B. Ablepharon–macrostomia syndrome in a 46-year-old woman. **American Journal of Medical Genetics Part A**, v. 127A, p. 96–98, 2004.
- BROWN, S.D.; MOORE, M.W. The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping. **Mammalian Genome**, v. 23, n.9-10, p. 632-40, 2012.
- CAVALCANTI, D. P., MATEJAS, V., LUQUETTI, D., MELLO, M. F., ZENKER, M. Fraser and ablepharon macrostomia phenotypes: concurrence in one family and association with mutated FRAS1. **American Journal of Medical Genetics Part A**, v., 143^a, p. 241-247, 2007.
- CERVANTES-BARRAGAN, D. E.; VILLARROEL, C. R.; MEDRANO-HERNANDEZ, A.; et al. Setleis syndrome in Mexican-Nahua sibs due to a homozygous TWIST2 frameshift mutation and partial expression in heterozygotes: review of the focal facial dermal dysplasias and subtype reclassification. **Journal of Medical Genetics**, v. 48, p. 716-720, 2011.

CHAN, J.L.; ORAL, E.A. Clinical classification and treatment of congenital and acquired lipodystrophy. **Endocrine Practice**, v. 16, p. 310–323, 2010.

CHEN, Z.; WANG, J.L.; TANG, B.S.; et al. Using next-generation sequencing as a genetic diagnostic tool in rare autosomal recessive neurologic Mendelian disorders. **Neurobiology Aging**, v. 13, p.197-8, 2013.

CHOI, M; SCHOLL, UI; JI, W; et al. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. **Proceedings of the National Academy of Sciences of the United States of America**, v. 106, n. 45, p. 19096–19101, 2009.

CHOU, J.; OHSUMI T. K.; GEHA, R. S. Use of whole exome and genome sequencing in the identification of genetic causes of primary immunodeficiencies. **Current Opinion in Allergy and Clinical Immunology**, v. 12, p. 623-628, 2012.

CHOWANADISAIA, W.; GRAHAM, D. M.; KEENA, C. L.; et al. Neurulation and neurite extension require the zinc transporter ZIP12 (slc39a12). **Proceedings of the National Academy of Sciences**, v. 110, n. 24, p. 9903–9908, 2013.

CLARK, M. J.; CHEN, R.; LAM, H. Y. K.; et al. Performance comparison of exome DNA sequencing technologies. **Nature Biotechnology**, v. 29, p. 908–914, 2011.

CLARK, R. D.; GOLABI, M.; LACASSIE, Y.; et al. Expanded phenotype and ethnicity in Setleis syndrome. **American Journal of Medical Genetics**, v. 34, p. 354-357, 1989.

CORDELL HJ. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. **Human Molecular Genetics**, v. 11, n.20, p. 2463–2468, 2002.

CORNACCHIA, D.; DILEEP, V.; QUIVY, J. P.; et al. Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. **EMBO Journal**, v. 31, p. 3678-3690, 2012.

CRUZ, A.A.V.; GUIMARAES F.C.; OBEID, H.N.; et al. Congenital shortening of the anterior lamella of all eyelids: The so-called ablepharon macrostomia syndrome. **Ophthalmic Plastic Reconstruction Surgery**, v. 11, p. 284–287, 1995.

DI VIRGILIO, M.; CALLEN, E.; YAMANE, A.; et al. Rif1 prevents resection of DNA breaks and promotes immunoglobulin class switching. **Science**, v. 339, p. 711-715, 2013.

DINULOS, M.B.; PAGON, R.A. Autosomal dominant inheritance of Barber-Say syndrome. **American Journal of Medical Genetics**, v. 86, p. 54-56. 1999.

DONKERVOLT, S.; HU, Y.; STOJKOVIC, T.; et al. Mosaicism for Dominant Collagen 6 Mutations as a Cause for Intrafamilial Phenotypic Variability. **Human Mutation**, v. 36, n. 1, p. 48–56, 2015.

ESTRADA-CUZCANO, A.; NEVELING, K.; KOHL, S.; et al. Mutations in C8orf37, Encoding a Ciliary Protein, are Associated with Autosomal-Recessive Retinal Dystrophies with Early Macular Involvement. **The American Journal of Human Genetics**, v. 90, p. 102–109, 2012.

EUROPEAN COMMISSION. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee

of the Regions on Rare Diseases: Europe's challenges. Brussels: European Commission; 2008.

FERRAZ, V.; MELO, D.G.; HANSING, S.E.; CRUZ, A.; PINA-NETO, J.M. Ablepharon–macrostomia syndrome: First report of familial occurrence. **American Journal of Medical Genetics**, v. 94, p. 281–283, 2000.

FIORENZA, C.G.; CHOU, S.H.; MANTZOROS, C.S. Lipodystrophy: pathophysiology and advances in treatment. **Nature Reviews Endocrinology**, v. 7, p. 137–150, 2011.

FIRULLI AB, CONWAY SJ. Phosphoregulation of Twist1 provides a mechanism of cell fate control. **Current Medicinal Chemistry**, v. 15, n. 25, p. 2641–7, 2008.

FIRULLI B. A.; KRAWCHUK, D.; CENTONZE, V. E.; et al. Altered Twist1 and Hand2 dimerization is associated with Saethre-Chotzen syndrome and limb abnormalities. **Nature Genetics**, v. 37, n. 4, p. 373–81, 2005.

FRANCO, H.L.; CASASNOVAS, J. J.; LEON, R.G.; et al. Nonsense mutations of the bHLH transcription factor TWIST2 found in Setleis Syndrome patients cause dysregulation of periostin. **The International Journal of Biochemistry & Cell Biology**, v. 43, n. 10, p. 1523–31, 2011a.

FRANCO, H.L.; CASASNOVAS, J.; RODRIGUEZ-MEDINA, J.R.; CADILLA, C.L. Redundant or separate entities?--roles of Twist1 and Twist2 as molecular switches during gene transcription. **Nucleic Acids Research**, v. 39, n. 4, p. 1177–86, 2011b.

FRASER, G.R. Our genetical 'load:' a review of some aspects of genetical variation. **Annual Human Genetics**, v. 25, p. 387–415, 1962.

GALLATI, S. Disease-modifying genes and monogenic disorders: experience in cystic fibrosis **The Application of Clinical Genetics.**, v. 7, p. 133–146, 2014.

GARG, A. Lipodystrophies: genetic and acquired body fat disorders. **The Journal of Clinical Endocrinology & Metabolism**, v. 96, p. 3313–3325, 2011.

GARG, A.; AGARWAL A.K. Lipodystrophies: Disorders of adipose tissue biology. **Biochimica et Biophysica Acta**, v. 1791, n. 6, 507–513, 2009.

GOH, G.; CHOI, M. Application of Whole Exome Sequencing to Identify Disease-Causing Variants in Inherited Human Diseases. **Genomics & informatics**, v. 10, n. 4, p. 214–219, 2012.

GONG, X. Q.; LI, L. Dermo-1, a multifunctional basic helix-loop-helix protein, represses MyoD transactivation via the HLH domain, MEF2 interaction, and chromatin deacetylation. **The Journal of Biological Chemistry**, v. 277, n. 14, p. 12310–7, 2002.

GONZAGA-JAUREGUI, C.; LUPSKI, J. R.; GIBBS, R.A. Human genome sequencing in health and disease. **Annual Review of Medicine**, v. 63, p. 35–61, 2012.

GROPMAN, A. L.; ADAMS, D.R. Atypical patterns of inheritance. **Seminars in Pediatric Neurology**, v. 14, n. 1, p. 34–45, 2007.

HAENSEL, J.; KOHLSCHMIDT, N.; PITZ, S.; et al. Case report supporting that the Barber-Say and ablepharon macrostomia syndromes could represent one disorder. **American Journal of Medical Genetics Part A**, v. 149^a, p. 2236–2240, 2009.

HANAOKA, K.; QIAN, F.; BOLETTA, A.; et al. Co-assembly of polycystin-1 and -2 produces unique cation-permeable currents. **Nature**, v. 408, p. 990-994, 2000.

HANDELSMAN, Y.; ORAL, E. A.; BLOOMGARDEN, Z. T.; et al. The clinical approach to the detection of lipodystrophy – an aace consensus statement. *Endocrine Practice*, v. 19, n. 1, p. 107–116, 2013.

HORNBLASS, A.; REIFLER, D.M. Ablepharon macrostomia syndrome. **American Journal of Ophthalmology**, v. 99, p. 552–556, 1985.

INTERNATIONAL HUMAN GENOME CONSORTIUM. Finishing the euchromatic sequence of the human genome. **Nature**, v. 431, p. 931–945, 2004.

JACKSON, IT; SHAW, KE; DEL PINAL MATORRAS, F. A new feature of the ablepharon macrostomia syndrome: Zygomatic arch absence. **British Journal of Plastic Surgery**, v. 41, p. 410–416, 1988.

JADEJA, S.; SMYTH, I.; PITERA, J.E.; et al. Identification of a new gene mutated in Fraser syndrome and mouse myelencephalic blebs. **Nature Genetics**, v. 37, p. 520–525, 2005.

JEFFERY, S.; SAGGAR-MALIK, A. K.; ECONOMIDES, D. L.; et al. Apparent normalisation of fetal renal size in autosomal dominant polycystic kidney disease (PKD1). *Clinical Genetics*, v. 53, p. 303-307, 1998.

JONASSEN, J. A.; SANAGUSTIN, J.; BAKER, S. P.; PAZOUR, G. J. Disruption of IFT complex A causes cystic kidneys without mitotic spindle misorientation. **Journal of the American Society of Nephrology**, v. 23, n. 4, p. 641-51, 2012.

KALLISH, S.; MCDONALD-MCGINN, D.M.; VAN HAELEST, M.M.; et al. Ablepharon–Macrostomia syndrome - Extension of the phenotype. **American Journal of Medical Genetics Part A**, v.155, p. 3060–3062, 2011.

KEITH, T.; REEDERS, S. T.; BROWN, V. A.; et al. **Linkage studies of autosomal dominant polycystic kidney disease (PKD1) and a genetic map of chromosome 16 based on 40 RFLP markers. (Abstract)** *American Journal of Human Genetics*, v. 41, p. A172, 1987.

KHAN, A. O.; BOLZ, H. J.; BERGMANN, C. Early-onset severe retinal dystrophy as the initial presentation of IFT140-related skeletal ciliopathy. **Journal of AAPOS**, v. 18, n. 2, p. 203-5, 2014.

KIYOZUMI, D.; SUGIMOTO, N.; SEKIGUCHI, K. Breakdown of the reciprocal stabilization of QBRICK/Frem1, Fras1, and Frem2 at the basement membrane provokes Fraser syndrome-like defects. **Proceedings of the National Academy of Sciences of the United States of America** v. 103, p. 11981–11986, 2006.

KWOK, W. K.; LING, M. T.; LEE, T. W.; et al. Up-Regulation of TWIST in Prostate Cancer and Its Implication as a Therapeutic Target. **Cancer Research**, v. 65, p. 5153–5162, 2005.

KYO, S.; SAKAGUCHI, J.; OHNO, S.; et al. High Twist expression is involved in infiltrative endometrial cancer and affects patient survival. **Human Pathology**, v. 37, p. 431–438, 2006.

LAIRD, P. Principles and challenges of genome-wide DNA methylation analysis. **Nature Reviews Genetics**, 11:191–203, 2010.

LANGMEAD, B.; SALZBERG, S. Fast gapped-read alignment with Bowtie 2. **Nature**, v. 9, p. 357-359, 2012.

LARUMBE, J.; VILLALTA, P.; VELEZ, I. Clinical Variant of Ablepharon Macrostomia Syndrome. **Case Reports in Dermatological Medicine**, v. 2011, Article ID 593045, 4 pages, 2011.

LEE, J. E.; KIM, H. J.; BAE, J. Y.; et al. Neogenin expression may be inversely correlated to the tumorigenicity of human breast cancer. **BMC Cancer**, v. 5, p. 154, 2005.

LEE, M.S.; LOWE, G.; FLANAGAN, S.; KUCHLER, K; GLACKIN, C. A. Human Dermo-1 has attributes similar to twist in early bone development. **Bone**, v. 27, n. 5, p. 591-602, 2000.

LI, H.; HANDSAKER, B.; WYSOKER, A.; FENNELL, T.; RUAN, J. The Sequence Alignment and Map format and SAMtools. **Bioinformatics**, v. 25, p. 2078–2079, 2009.

LI, J.; YIN, W. H.; TAKEUCHI, K.; et al. Inflammatory myofibroblastic tumor with RANBP2 and ALK gene rearrangement: a report of two cases and literature review. **Diagnostic Pathology**, v. 13, n.8, p. 147, 2013.

LI, L.; CSERJESI, P.; OLSON, E. N. Dermo-1: a novel twist-related bHLH protein expressed in the developing dermis. **Developmental Biology**, v. 172, n. 1, p. 280-92, 1995.

LIGHT, J.; WILLEMSSEN, M.H.; VAN BON, B.W.; et al. Diagnostic exome sequencing in persons with severe intellectual disability. **New England Journal of Medicine**, v. 367, p. 1921–9, 2012.

LIN, X.; TANG, W.; AHMAD, S.; LU, J.; COLBY, C.C.; ZHU, J.; YU, Q. Applications of targeted gene capture and next-generation sequencing technologies in studies of human deafness and other genetic disabilities. **Hearing Research**, v. 288, p. 67–76, 2012.

MAAS, S.; GERBER, A. P.; RICH, A. Identification and characterization of a human tRNA-specific adenosine deaminase related to the ADAR family of pre-mRNA editing enzymes. **Proceedings of the National Academy of Sciences of the United States of America**, v. 96, p. 8895-8900, 1999.

MAAS, S.; KIM, Y. G.; RICH, A. Genomic clustering of tRNA-specific adenosine deaminase ADAT1 and two tRNA synthetases. **Mammalian Genome**, v. 12, n. 5, p. 387-93, 2001.

MAAS, S.; KIM, Y. G.; RICH, A. Sequence, genomic organization and functional expression of the murine tRNA-specific adenosine deaminase ADAT1. **Gene**, v. 243, n. 1-2, p. 59-66, 2000.

MACBETH, M. R.; SCHUBERT, H. L.; VANDEMARK, A. P.; et al. Inositol hexakisphosphate is bound in the ADAR2 core and required for RNA editing. **Science**, v. 309, n. 5740, p. 1534-9, 2005.

MAJEWSKI, J.; SCHWARTZENTRUBER, J.; LALONDE, E.; MONTPETIT, A.; JABADO, N. What can exome sequencing do for you?. *Journal of Medical Genetics*, v. 48, n. 9, p. 580-589, 2011.

MARIAN, A. J. Challenges in medical applications of whole exome/genome sequencing discoveries. *Trends Cardiovascular Medicine*, v. 22, p. 219–223, 2012.

MARION, R. W., CHITAYAT, D., HUTCHEON, G., GOLDBERG, R., SHPRINTZEN, R. J., COHEN, M. M., JR. Autosomal recessive inheritance in the Setleis bitemporal 'forceps marks' syndrome. *American Journal of Diseases of Children*, v. 141, p. 895-897, 1987.

MARKOUIZOS, D.; SIDDIQI, U.; SIDDIQI, S.; RAZIUDDIN, K.; NANGIA, B. Ablepharon macrostomia syndrome: Report of a case and clinical delineation. *American Journal of Human Genetics*, v. 47A, p. 66, 1990.

MARTINEZ, S. R.; GAY, M. S.; ZHANG, L. Epigenetic mechanisms in heart development and disease. *Drug Discovery Today*, 2015.

MATSUNAGA, E.; TAUSZIG-DELAMASURE, S.; MONNIER, P. P.; et al. RGM and its receptor neogenin regulate neuronal survival. *Nature Cell Biology*, v. 6, p. 749-755, 2004.

MAZZANTI, L.; BERGAMASCHI, R.; NERI, I.; et al. Barber-Say syndrome: Report of a new case. *American Journal of Medical Genetics*, v. 78, p. 188-191, 1998.

MCCARTHY, G.T.; WEST, C.M. Ablepharon macrostomia syndrome. *Developmental Medicine & Child Neurology*, v. 19, n. 5, p. 659–663, 1977.

MCGREGOR, L.; MAKELA, V.; DARLING, S. M.; et al. Fraser syndrome and mouse blebbed phenotype caused by mutations in FRAS1/Fras1 encoding a putative extracellular matrix protein. *Nature Genetics*, v. 34, p. 203-208, 2003.

METZKER, M.L. Sequencing technologies - the next generation. *Nature Reviews Genetics*, v. 11, p. 31–46, 2010.

MEYNERT, A.; ANSARI, M.; FITZPATRICK, D.; TAYLOR, M. Variant detection sensitivity and biases in whole genome and exome sequencing. *BMC Bioinformatics*, v. 15, p. 247, 2014.

MEYNERT, A.; BICKNELL, L. S.; HURLES, M. E.; JACKSON, A. P.; TAYLOR, M. S. Quantifying single nucleotide variant detection sensitivity in exome sequencing. *BMC Bioinformatics*, v. 14, p. 195, 2013.

MILLA, L. A.; ARROS, A.; ESPINOZA, N.; et al. Neogenin1 is a Sonic Hedgehog target in medulloblastoma and is necessary for cell cycle progression. *International Journal of Cancer*, v. 134, p. 21–31, 2014.

MOSHOUS, D.; MARTIN, E.; CARPENTIER, W.; LIM, A.; CALLEBAUT, I.; CANIONI, D. Whole-exome sequencing identifies Coronin-1A deficiency in 3 siblings with immunodeficiency and EBV-associated B-cell lymphoproliferation. *Journal of Allergy and Clinical Immunology*, v. 131, p. 1594–1603, 2013.

MUKHOPADHYAY, S.; WEN, X.; CHIH, B.; et al. TULP3 bridges the IFT-A complex and membrane phosphoinositides to promote trafficking of G protein-coupled receptors into primary cilia. *Genes & Development*, v. 24, n. 19, p. 2180-93, 2010.

MURRE, C.; MCCAWE, O. S.; VAESSIN, H.; et al. Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. **Cell**, v. 58, n. 3, p. 537-44, 1989.

NAGASE, T.; ISHIKAWA, K.; SUYAMA, M.; et al. Prediction of the coding sequences of unidentified human genes. XI. The complete sequences of 100 new cDNA clones from brain which code for large proteins in vitro. **DNA Research**, v. 5, n. 5, p. 277-86, 1998.

NEILSON, D. E.; ADAMS, M. D.; ORR, C. M. D.; et al. Infection-triggered familial or recurrent cases of acute necrotizing encephalopathy caused by mutations in a component of the nuclear pore, RANBP2. **The American Journal of Human Genetics**, v. 84, p. 44-51, 2008.

NG, S.B.; BUCKINGHAM, K.J.; LEE, C.; et al. Exome sequencing identifies the cause of a mendelian disorder. **Nature Genetics**, v. 42, p. 30-35, 2010b.

NG, S.B.; NICKERSON, D.A.; BAMSHAD, M.J.; SHENDURE, J. Massively parallel sequencing and rare disease. **Human Molecular Genetics**, v. 19, p. 119-124, 2010a.

OMORI, Y.; KYUSHIKI, H.; TAKEDA, S.; et al. Cloning, expression and mapping of a novel human zinc-finger gene TCF17 homologous to rodent Kid1. **Cytogenetics and Cell Genetics**, v. 78, p. 285-288, 1998.

PARK, P. J. ChIP-seq: advantages and challenges of a maturing technology. **Nature Reviews Genetics**, v. 10, p. 669-680, 2009.

PELLEGRINO, J.E.; SCHNUR, R.E.; BOGHOSIAN-SELL, L.; et al. Ablepharon macrostomia syndrome with associated cutis laxa: Possible localization to 18q. **Human Genetics**, v. 97, p. 532-536, 1996.

PERRAULT, I.; SAUNIER, S.; HANEIN, S.; et al. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. **American Journal of Medical Genetics**, v. 90, n. 5, p. 864-70, 2012.

PERRIN-SCHMITT, F.; BOLCATO-BELLEMIN, A.L.; BOURGEOIS, P.; et al. The locations of the H-twist and H-dermo-1 genes are distinct on the human genome. **Biochimica et Biophysica Acta**, v. 1360, n. 1, p. 1-2, 1997.

PICHLER, A.; GAST, A.; SEELER, J. S.; et al. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. **Cell**, v. 108, p. 109-120, 2001.

PRICE, N.J.; PUGH, R.E.; FARNDON, P.A.; WILLSHAW, H.E. Ablepharon macrostomia syndrome. **British Journal of Ophthalmology**, v. 75, p. 317-319, 1991.

PRUITT, K. D.; BROWN, G. R.; HIATT, S. M.; et al. RefSeq: an update on mammalian reference sequences. **Nucleic Acids Research**, 2013.

PRZYBORSKI, S.A.; KNOWLES, B.B.; HANDEL, M. A.; et al. Differential expression of the zinc finger gene Zfp105 during spermatogenesis. **Mammalian Genome**, v. 9, p. 758-762, 1998.

RABBANI, B; MAHDIEH, N; HOSOMICHI, K; NAKAOKA, H; INOUE, I. “Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders”. **Journal of Human Genetics** v. 57, n. 10, p. 621–632, 2012.

ROBINSON, JT; THORVALDSDÓTTIR, H; WINCKLER, W; et al. Integrative Genomics Viewer. **Nature Biotechnology**, v. 29, p. 24–26, 2011.

ROCHE, N; HOUTMEYERS, P; JANSSENS, S; BLONDEEL, P. Barber-Say syndrome in a father and daughter. **American Journal of Medical Genetics Part A**, v. 152A, p. 2563-8, 2010.

ROHENA, L.; KUEHN, D.; MARCHEGIANI, S.; HIGGINSON, J. D. Evidence for autosomal dominant inheritance of ablepharon–macrostomia syndrome. **American Journal of Medical Genetics**, v. 155, p. 850–854, 2011.

SAITOH, N.; SAKAMOTO, C.; HAGIWARA, M.; et al. The distribution of phosphorylated SR proteins and alternative splicing are regulated by RANBP2. **Molecular Biology of the Cell**, v. 23, n. 6, p. 1115-1128, 2012.

SCHANZE, D.; HARAKALOVA, M.; STEVENS, C.A.; et al. Ablepharon macrostomia syndrome: a distinct genetic entity clinically related to the group of FRAS-FREM complex disorders. **American Journal of Medical Genetics Part A**, v. 161A, p. 3012–3017, 2013.

SCHMIDTS, M.; FRANK, V.; EISENBERGER, T.; et al. Combined NGS approaches identify mutations in the intraflagellar transport gene IFT140 in skeletal ciliopathies with early progressive kidney Disease. **Human Mutation**, v. 34, n. 5, p. 714-24, 2013.

SETLEIS, H.; KRAMER, B.; VALCARCEL, M.; EINHORN, A. H. Congenital ectodermal dysplasia of the face. **Pediatrics**, v. 32, p. 540–548, 1963.

SHIRATSUCHI T.; ODA, K.; NISHIMORI, H.; et al. Cloning and characterization of BAP3 (BAI-associated protein 3), a C2 domain-containing protein that interacts with BAI1. **Biochemical and Biophysical Research Communications**, v. 251, n. 1, p. 158-65, 1998.

SLAVOTINEK, A.; BARANZINI, S. E.; SCHANZE, D.; et al. Manitoba-oculo-trichoanal (MOTA) syndrome is caused by mutations in FREM1. **Journal of Medical Genetics**, v. 48, p. 375–382, 2011.

SLAVOTINEK, A.; BIESECKER, L. G. Genetic modifiers in human development and malformation syndromes, including chaperone proteins. **Human Molecular Genetics**, v. 12, n. 1, p. 45–R50, 2003.

SMYTH, I.; SCAMBLER, P. The genetics of Fraser syndrome and the blebs mouse mutants. **Human Molecular Genetics**, v. 14, n. 2, p. 269–274, 2005.

SONG L, CRAWFORD GE. DNase-seq: a high-resolution technique for mapping active gene regulatory elements across the genome from mammalian cells. **Cold Spring Harbor Protocols**, 5384, 2010.

SONG, X.; DI GIOVANNI, V.; HE, N.; et al. Systems biology of autosomal dominant polycystic kidney disease (ADPKD): computational identification of gene expression pathways and integrated regulatory networks. **Human Molecular Genetics**, v. 18, p. 2328 – 2343, 2009.

SOSIC, D.; RICHARDSON, J. A.; YU, K.; et al. Twist regulates cytokine gene expression through a negative feedback loop that represses NF-kappaB activity. **Cell**, v. 112, n. 2, p. 169-80, 2003.

STAUDINGER, J.; PERRY, M.; ELLEDGE, S.J.; OLSON, E.N. Interactions among vertebrate helix-loop-helix proteins in yeast using the two-hybrid system. **The Journal of Biological Chemistry**, v. 268, p. 4608–4611, 1993.

STEVENS, C.A.; SARGENT, L.A. Ablepharon–macrostomia syndrome. **American Journal of Medical Genetics**, v. 107, p. 30–37, 2002.

STRACHAN T, READ AP. Human Molecular Genetics. 2nd edition. Chapter 16, Molecular pathology. New York: Wiley-Liss; 1999.

TAKATA, A.; KATO, M.; NAKAMURA, M.; et al. Exome sequencing identifies a novel missense variant in RRM2B associated with autosomal recessive progressive external ophthalmoplegia. **Genome Biology**, v. 12, p. 92, 2011.

TENG, Y.; LI, X. The roles of HLH transcription factors in epithelial mesenchymal transition and multiple molecular mechanisms. **Clinical and Experimental Metastasis**; v. 31, n. 3, p. 367-77, 2014.

THORVALDSDÓTTIR, H; ROBINSON, JT; MESIROV, JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. **Briefings in Bioinformatics**, v. 14, p. 178-192, 2013.

TSUKAHARA, M., OKABE, T., OHTSUKA, M., FURUKAWA, S. Follow-up study in a patient with Setleis syndrome. **American Journal of Medical Genetics**, v. 57, p. 444-446, 1995.

TUKEL, T.; ŠOŠIĆ, D.; AL-GAZALI, L. I.; et al. Homozygous Nonsense Mutations in TWIST2 Cause Setleis Syndrome. **American Journal of Human Genetics**, v. 87, n. 2, p. 289-296, 2010.

US FOODDRUG ADMINISTRATION. Orphan Drug Act, Pub L. No.97-414. Washington, DC: Public Law, 1983.

VENSELAAR, H.; BEEK, T. A H. TE; KUIPERS, R. K. P.; HEKKELMAN, M. L.; VRIEND, G. Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. **BMC bioinformatics**, v. 11, n. 1, p. 548, 2010.

VETTER, I. R.; NOWAK, C.; NISHIMOTO, T.; et al. A. Structure of a Ran-binding domain complexed with Ran bound to a GTP analogue: implications for nuclear transport. **Nature**, v. 398, p. 39-46, 1998.

VINAYAGAM, A.; STELZL, U.; FOULLE, R.; et al. A directed protein interaction network for investigating intracellular signal transduction. **Science Signaling**, v. 4, n.189, 2011.

VOGEL, M. J.; VAN ZON, P.; BRUETON, L.; et al. Mutations in GRIP1 cause Fraser syndrome. **Journal of Medical Genetics**, v. 49, p. 303–306, 2012.

WANG, Z; GERSTEIN, M; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics. **Nature Reviews Genetics**, v. 10, n. 1, p. 57-63, 2009.

WILSON, N. H.; KEY, B. Neogenin: one receptor, many functions. **The International Journal of Biochemistry & Cell Biology**, v. 39, p. 874–878, 2007.

WITZGALL, R.; O'LEARY, E.; GESSNER, R.; et al. Kid-1, a putative renal transcription factor: regulation during ontogeny and in response to ischemia and toxic injury. **Molecular and Cellular Biology**, v. 13, p. 1933-1942, 1993.

WORTHEY, E.A. Analysis and Annotation of Whole-Genome or Whole-Exome Sequencing–Derived Variants for Clinical Diagnosis. **Current Protocols in Human Genetics**, v. 79, p. 9.24.1-9.24.24, 2013.

WU J, LI Y, JIANG R. Integrating Multiple Genomic Data to Predict Disease-Causing Nonsynonymous Single Nucleotide Variants in Exome Studies. **PLoS Genetics**, v. 10, n. 3, e1004237, 2014.

YANG, J.; ZHUC,M.; WANGA, Y.; et al. Whole-exome sequencing identify a newmutation of MYH7 in a Chinese family with left ventricular noncompaction. **Gene**, v. 558, p. 138–142, 2015.

8. ANEXOS

Anexo A. Aprovação do Comitê de Ética em Pesquisa

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA-MACROSTOMIA

Pesquisador: Victor Evangelista de Faria Ferraz

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP);

Versão: 2

CAAE: 40985214.4.0000.5440

Instituição Proponente: Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da USP -

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.012.074

Data da Relatoria: 20/03/2015

Apresentação do Projeto:

A abordagem genômica no estudo de doenças raras tem sido amplamente utilizada, devido principalmente ao surgimento da nova geração de sequenciamento, que possui alto poder de discriminar as sequências nucleotídicas com grande cobertura, em um curto período de tempo. O sequenciamento completo do exoma, associado às novas plataformas de sequenciamento, permite a análise de regiões funcionais do genoma humano, com eficiência sem precedentes. A Síndrome Ablefaria-Macrostomia (AMS) é uma condição rara, onde os pacientes apresentam características clínicas marcantes como o encurtamento ou ausência das pálpebras superiores e inferiores, ausência de sobrancelhas e cílios, macrostomia por defeitos na fusão dos lábios, entre outros. O padrão de herança da síndrome não está elucidado, tendo a herança autossômica dominante com expressividade variável sido sugerida, porém o padrão de herança recessivo não foi descartado. O presente estudo busca verificar a determinação da base molecular da doença, pela detecção de variantes genômicas responsáveis pelas características fenotípicas da síndrome, utilizando o sequenciamento completo do exoma. O estudo será conduzido pelo Serviço de Genética Médica do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo (HC-FMRP).

Endereço: CAMPUS UNIVERSITÁRIO
Bairro: MONTE ALEGRE **CEP:** 14.048-900
UF: SP **Município:** RIBEIRÃO PRETO
Telefone: (16)3602-2228 **Fax:** (16)3633-1144 **E-mail:** cep@hcrp.usp.br

Continuação do Parecer: 1.012.074

USP) em parceria com o Laboratório de Genética Molecular e Bioinformática. Serão participantes uma família que possui três membros afetados pela síndrome (pai- com quadro leve e duas filhas- com quadro clássico) e dois membros não afetados (mãe e uma filha). A coleta de dados constitui-se na coleta de sangue periférico para extração de DNA da família. Serão realizadas então análises das amostras: 1. Amostras de sangue periférico serão coletadas e a extração de DNA será realizada com kit padrão de extração Qiagen QIAmp DNA Blood de acordo com recomendações do fabricante; 2) O DNA genômico obtido das amostras de sangue dos cinco pacientes será utilizado para o sequenciamento do exoma paired-end, por meio do sequenciador GA IIx (plataforma Illumina); 3) O processamento dos dados obtidos do sequenciamento será realizado em hardware com cluster de 452 núcleos de processamento contendo 1.3 terabytes de memória RAM e sistema operacional Fedora 11 (Leonidas) kernel 2.6.30 e arquitetura 64 bits, disponível no Laboratório de Genética Molecular e Bioinformática do Centro Regional de Hemoterapia e Hematologia de Ribeirão Preto. As etapas de processamento são: Validação da qualidade do sequenciamento -> alinhamento contra genoma de referência hg19 utilizando algoritmo "bowtie2" -> detecção de deleções, inserções, mutações e polimorfismos pelas ferramentas GATK -> anotação e análise quanto ao efeito funcional dos SNPs pelo programa SIFT e Polyphen. 4) Validação das variantes por sequenciamento Sanger. Espera-se que esta abordagem possa determinar a base genética da síndrome e, assim, facilitar o diagnóstico molecular, auxiliar no aconselhamento genético, elucidar a variabilidade fenotípica intra-familiar e possivelmente evidenciar genes envolvidos com o desenvolvimento normal dos tecidos afetados.

Objetivo da Pesquisa:

Tem-se como objetivo geral identificar a base genética da Síndrome de Ablefaro Macrostomia por meio da utilização do sequenciamento completo do exoma da família em estudo. Quanto a objetivo secundário espera-se caracterizar clinicamente a família portadora da Síndrome de Ablefaro Macrostomia; Identificar as alterações genotípicas associadas com o fenótipo nesta família (caracterizar as alterações encontradas quanto a patogenicidade, identificar alterações genéticas que possivelmente determinem a variabilidade intra-familiar do quadro clínico) e; validar as variantes encontradas por meio de sequenciamento (Sanger).

Endereço: CAMPUS UNIVERSITÁRIO
Bairro: MONTE ALEGRE **CEP:** 14.048-900
UF: SP **Município:** RIBEIRÃO PRETO
Telefone: (16)3602-2228 **Fax:** (16)3633-1144 **E-mail:** cep@hcrp.usp.br

Continuação do Parecer: 1.012.074

Avaliação dos Riscos e Benefícios:

Segundo os pesquisadores os riscos são mínimos estando associado a um desconforto mínimo durante a coleta de sangue, a possibilidade de surgir uma mancha roxa, bem como inchaço e em raros casos infecção no local. No que tange aos benefícios trará informações aos participantes da herança genética e da possibilidade do problema acontecer novamente na família.

Comentários e Considerações sobre a Pesquisa:

Sinaliza relevância científica e social podendo contribuir com a melhor compreensão das variáveis estudadas, e orientação genética às famílias.

Considerações sobre os Termos de apresentação obrigatória:

Apresenta toda documentação necessária.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

Diante do exposto e à luz da Resolução CNS 466/2012, o projeto de pesquisa versão 2 de 18/03/2015, assim como os Termos de Consentimento Livre e Esclarecido e Termo de Assentimento versão 2 de 18/03/2015, podem ser enquadrados na categoria APROVADO.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

Projeto Aprovado: Tendo em vista a legislação vigente, devem ser encaminhados ao CEP, relatórios parciais anuais referentes ao andamento da pesquisa e relatório final ao término do trabalho. Qualquer modificação do projeto original deve ser apresentada a este CEP em nova versão, de forma objetiva e com justificativas, para nova apreciação.

Endereço: CAMPUS UNIVERSITÁRIO
 Bairro: MONTE ALEGRE CEP: 14.048-900
 UF: SP Município: RIBEIRÃO PRETO
 Telefone: (16)3602-2228 Fax: (16)3633-1144 E-mail: cep@hcrp.usp.br

Continuação do Parecer: 1.012.074

RIBEIRÃO PRETO, 06 de Abril de 2015

Assinado por:
MARCIA GUIMARÃES VILLANOVA
(Coordenador)

Endereço: CAMPUS UNIVERSITÁRIO
Bairro: MONTE ALEGRE **CEP:** 14.048-900
UF: SP **Município:** RIBEIRÃO PRETO
Telefone: (16)3602-2228 **Fax:** (16)3633-1144 **E-mail:** cep@hcrp.usp.br

Anexo B. Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA

1. NOME:
 DOCUMENTO DE IDENTIDADE: SEXO : .MASC ☐ FEM ☐
 DATA NASCIMENTO:/...../.....
 ENDEREÇO Nº APTO:
 BAIRRO: CIDADE
 CEP:..... TELEFONE: DDD (.....)

NOME DA PESQUISA: “DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA-MACROSTOMIA (SAM)”.

PESQUISADOR RESPONSÁVEL: Victor Evangelista de Faria Ferraz

CARGO/FUNÇÃO: médico (CRM-SP 71347), professor doutor do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo.

1 – Desenho do estudo e objetivo(s)

Você está sendo convidado (a) a participar, como voluntário (a) de uma pesquisa, por meio da qual pretendemos identificar possíveis causas para a Síndrome Ablefaria Macrostomia (SAM). Esta síndrome muito rara se caracteriza pelo encurtamento ou ausência das pálpebras superiores e inferiores, ausência de sobrancelhas e cílios, pouco crescimento do cabelo, boca grande, alterações no formato do nariz e orelhas, pele seca, entre outros.

A SAM é considerada de causa genética, mas a alteração exata ainda não foi totalmente esclarecida, assim como sua forma de transmissão dentro da família. O objetivo deste projeto de pesquisa é investigar possíveis causas genéticas desta síndrome a partir do estudo do DNA (material genético humano) em pessoas afetadas e seus familiares.

2 – Descrição dos procedimentos que serão realizados, com seus propósitos e identificação dos que forem experimentais e não rotineiros;

Caso aceite participar, você realizará uma consulta médica com geneticista, que incluirá exame físico rotineiro de uma consulta médica, onde mediremos a altura, peso, tamanho da cabeça e outras medidas que forem necessárias, com régua e fita métrica, sem desconforto. Após a consulta uma amostra de 10 ml do

seu sangue será colhida no Hospital das Clínicas de Ribeirão Preto e será enviada ao laboratório para realização da análise genética. Você pode sentir um desconforto mínimo durante a coleta de sangue. No local da coleta de sangue pode surgir uma mancha roxa, bem como inchaço e, em raros casos, infecção no local.

O material coletado será guardado em tubos rotulados de modo que o doador não possa ser identificado, mantendo, assim, o caráter confidencial da informação. O exame que será realizado envolve a determinação da constituição da maioria dos genes humanos já descritos, com a finalidade de avaliar a causa desta doença no nível de DNA. Queremos deixar claro, que este tipo de exame gera informações sobre todos os genes que nós possuímos, no entanto, só analisaremos os genes que possam estar envolvidos com a SAM.

O seu prontuário médico no HCRP será revisado e, informações importantes, como resultados de outros exames, serão analisadas. Todos os dados utilizados na pesquisa preservarão sua identidade, mantendo seu anonimato.

3 – Benefícios para o participante

Não há benefício direto para você com o resultado da análise. Por outro lado, a sua análise genética poderá permitir a determinação da alteração genética que causou a síndrome na sua família. Isto servirá de modelo para procurarmos se seus parentes herdaram ou não este mesmo traço genético, o que pode auxiliar na determinação do risco do problema acontecer novamente na família.

Esta análise nos auxiliará a entender melhor a sua doença. Este melhor entendimento pode, no futuro, ajudar a prever a gravidade da doença e até mesmo influenciar na conduta médica e no tratamento.

4 – Garantia de acesso, retirada do teste e confidencialidade

Em qualquer etapa do estudo, você terá acesso aos profissionais responsáveis pela pesquisa para esclarecimento de eventuais dúvidas. Você terá o direito de ser mantido atualizado sobre os resultados parciais das pesquisas, bem como ao final do estudo. Os pesquisadores responsáveis podem ser encontrados no Departamento de Genética da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, na Av. Bandeirantes 3900, em Ribeirão Preto – SP (CEP 14049-900). Telefone: (16) 3602-4500 ou (16) 3602-3080 ou ainda (16) 3602-3102.

Você tem a liberdade de retirar o seu consentimento a qualquer momento e deixar de participar do estudo, sem qualquer prejuízo à continuidade de seu tratamento no Hospital das Clínicas. As informações obtidas do seu exame genético serão analisadas em conjunto com outros pacientes, não sendo divulgada a identificação de nenhum paciente.

5 – Despesas e compensações

Não há despesas pessoais para o participante em qualquer fase do estudo, incluindo exames e consultas. Também não há compensação financeira relacionada à sua participação. Se existir qualquer despesa adicional, ela será sanada pelo orçamento da pesquisa.

6 - Compromisso do pesquisador de utilizar os dados e o material coletado somente para esta pesquisa.

Eu, Victor Evangelista de Faria Ferraz (pesquisador responsável), me comprometo a utilizar o seu material coletado (sangue) e seus dados somente para esta pesquisa.

Prof. Dr. Victor E. de Faria Ferraz
Pesquisador Responsável

Eu, _____, fui esclarecido a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo “DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA-MACROSTOMIA (SAM)”.

Fui informado que no processo de análise genética podem ser encontrados diversos tipos de alterações e que tenho direito a decidir sobre qual informação desejo receber. Abaixo assinalo as informações que gostaria de receber:

- ☐ Alterações no DNA úteis para o diagnóstico e manejo da SAM;
- ☐ Alterações encontradas na análise relacionadas a outras doenças que afetam você e sua família e que possuem tratamento;
- ☐ Alterações encontradas na análise relacionadas ao risco de desenvolver doenças preveníveis ou tratáveis;
- ☐ Alterações encontradas na análise, relacionadas ao risco de ter filhos com doença genética, apesar de você aparentar não ter nenhum sinal ou sintoma;

Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter adquirido, ou no meu atendimento neste Serviço.

Assinatura do Participante

DATA ____/____/____. Local: _____.

Anexo C. Termo de Assentimento Livre e Esclarecido

TERMO DE ASSENTIMENTO LIVRE E ESCLARECIDO

DADOS DE IDENTIFICAÇÃO DO SUJEITO DA PESQUISA

1. NOME:
 DOCUMENTO DE IDENTIDADE: SEXO : .MASC ☐ FEM ☐
 DATA NASCIMENTO:/...../.....
 ENDEREÇO Nº Comp.....
 BAIRRO: CIDADE
 CEP:..... TELEFONE: DDD (.....)

NOME DA PESQUISA: "DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA-MACROSTOMIA (SAM)".

PESQUISADOR RESPONSÁVEL: Victor Evangelista de Faria Ferraz

CARGO/FUNÇÃO: médico (CRM-SP 71347), professor doutor do Departamento de Genética da Faculdade de Medicina de Ribeirão Preto – Universidade de São Paulo.

Você está sendo convidado (a) a participar, como voluntário (a) de uma pesquisa, para identificar possíveis causas para a Síndrome Ablefaria Macrostomia (SAM). Esta síndrome muito rara se caracteriza pelo encurtamento ou ausência das pálpebras, ausência de sobrancelhas e cílios, pouco crescimento do cabelo, boca grande, alterações no formato do nariz e orelhas, pele seca, entre outros.

A SAM é considerada de causa genética, mas a alteração exata ainda não foi totalmente esclarecida, assim como sua forma de transmissão dentro da família. O objetivo deste projeto de pesquisa é investigar possíveis causas genéticas desta síndrome a partir do estudo do DNA (material genético humano) em pessoas afetadas e seus familiares.

Caso aceite participar, você realizará uma consulta médica com geneticista, que incluirá exame físico normal de uma consulta médica, onde mediremos a altura, peso, tamanho da cabeça e outras medidas que forem necessárias, com régua e fita métrica, sem desconforto. Após a consulta uma amostra de sangue será colhida no Hospital das Clínicas de Ribeirão Preto e será enviada ao laboratório para realização da análise genética. Você pode sentir um desconforto pequeno durante a coleta de sangue. No local da coleta de sangue pode surgir uma mancha roxa, bem como inchaço e, em raros casos, infecção no local.

O material coletado será guardado em tubos rotulados de modo que o doador não possa ser identificado. O exame que será realizado envolve a determinação da constituição da maioria dos genes humanos já descritos, com a finalidade de avaliar a causa desta doença no nível de DNA. Queremos deixar claro, que este tipo de

exame gera informações sobre todos os genes que nós possuímos, no entanto, só analisaremos os genes que possam estar envolvidos com a SAM.

O seu prontuário médico no HCRP será revisado e, informações importantes, como resultados de outros exames, serão analisadas. Todos os dados utilizados na pesquisa preservarão sua identidade, mantendo seu anonimato.

Não há benefício direto para você com o resultado da análise. Por outro lado, a sua análise genética poderá permitir a determinação da alteração genética que causou a síndrome na sua família. Isto servirá de modelo para procurarmos se seus parentes herdaram ou não este mesmo traço genético, o que pode auxiliar na determinação do risco do problema acontecer novamente na família.

Esta análise nos auxiliará a entender melhor a doença. Este melhor entendimento pode, no futuro, ajudar a prever a gravidade da doença e até mesmo influenciar na conduta médica e no tratamento.

Eu, _____, fui esclarecido a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo “DETERMINAÇÃO DA BASE MOLECULAR DA SÍNDROME ABLEFARIA-MACROSTOMIA (SAM)”.

Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é isenta de despesas e que tenho garantia do acesso a tratamento hospitalar quando necessário. Concordo voluntariamente em participar deste.

Assinatura do Participante