Adriana Martins Fernandes

Diagnóstico molecular de osteogênese imperfeita através do sequenciamento paralelo em larga escala de 15 genes candidatos

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

Programa de Endocrinologia

Orientador: Dr. Bruno Ferraz de Souza

São Paulo

2019

Dados Internacionais de Catalogação na Publicação (CIP)

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

©reprodução autorizada pelo autor

Fernandes, Adriana Martins
Diagnóstico molecular de osteogênese imperfeita através do sequenciamento paralelo em larga escala de 15 genes candidatos / Adriana Martins Fernandes. -- São Paulo, 2019.
Tese(doutorado)--Faculdade de Medicina da Universidade de São Paulo.
Programa de Endocrinologia.
Orientador: Bruno Ferraz de Souza.

Descritores: 1.Osteogênese imperfeita 2.Diagnóstico 3.Genética 4.Sequenciamento de nucleotídeos em larga escala 5.Colágeno tipo I 6.Fraturas ósseas

USP/FM/DBD-250/19

Responsável: Erinalva da Conceição Batista, CRB-8 6755

Este trabalho foi desenvolvido na Unidade de Doenças Osteometabólicas da Divisão de Endocrinologia, no Laboratório de Carboidratos e Radioimunoensaio (LIM-18) e no Laboratório de Endocrinologia Celular e Molecular (LIM-25) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, com apoio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (Projeto Temático 2013/02162-8)

Dedicatória

Dedico esta tese aos meus amados pais, que tanto me apoiaram e incentivaram ao longo desta jornada.

A minha conquista também é de vocês.

Agradecimentos

Inicialmente quero agradecer ao meu orientador, Bruno Ferraz de Souza, por ter aceitado me guiar ao longo desta jornada, me oferecendo todo seu conhecimento, sua exímia didática e incrível acessibilidade. Agradeço a tão prazerosa convivência e também a paciência nos momentos difíceis ao longo de todos esses anos.

À Berenice Bilharinho de Mendonça, agradeço ter possibilitado a inclusão deste trabalho em um projeto temático, permitindo sua viabilidade financeira. Junto à Ana Claudia Latrônico, agradeço a paixão pela pesquisa disseminada na Disciplina de Endocrinologia e Metabologia da Faculdade de Medicina da USP.

Agradeço aos pacientes e seus familiares, que consentiram em participar desta pesquisa. À enfermagem do ambulatório da Divisão de Endocrinologia do Hospital das Clínicas da FMUSP, que realizou as coletas de sangue dos pacientes deste estudo. A toda a equipe dos laboratórios LIM-18 e LIM-25, pela convivência durante os anos deste trabalho e, em especial, ao Rodolfo Batista e à Marcia Correia, que realizaram a extração das amostras de DNA, e à Aritânia Santos e à Ana Mercedes Cavaleiro pela ajuda durante os experimentos realizados.

Agradeço ao SELA, onde foi realizado o sequenciamento de larga escala, etapa fundamental deste projeto, e à Mônica Malheiros França, responsável pela execução dos sequenciamentos deste trabalho.

Agradeço ao Antônio Marcondes Lerário pela análise bioinformática e ao Alexander Augusto de Lima Jorge pelos conselhos sobre a interpretação desta análise. À Leslie Kulikowski e à Evelin Zanardo pela colaboração nos estudos de citogenômica.

Aos que compuseram minha banca de qualificação, Telma Palomo de Oliveira, Luciani Renata Silveira de Carvalho e Alexander Augusto de Lima Jorge, agradeço o

tempo que dispuseram e as considerações que tanto contribuíram para o desenvolvimento deste trabalho e elaboração da tese.

Ao Pedro Henrique Silveira Corrêa e à Regina Matsunaga Martin, que, com vasto conhecimento, me apresentaram ao metabolismo ósseo.

Às amigas Manuela Rocha Braz e Vivian Roberta Ferreira Simões, agradeço pelo agradável convívio e ajuda durante o atendimento dos pacientes no ambulatório.

Aos amigos, agradeço o incentivo e apoio ao longo dessa jornada. À minha família, por acreditarem na minha capacidade, comemorarem minhas conquistas e admirarem o meu trabalho. Especialmente aos meus pais, que cresceram tanto frente a adversidades, me inspirando, e que me ofereceram todo o suporte para que eu conquistasse tudo o que almejei. Ao meu irmão, por estar sempre ao meu lado.

Ao meu marido, meu maior incentivador, agradeço por acreditar tanto em mim, pela certeza que tem de que tudo vai dar certo e pelo prazer que tem em me ver crescer. E, por fim, ao nosso filho, minha maior inspiração.

Normatização adotada

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

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

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

• Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

• Abreviaturas dos genes de acordo com *HUGO Gene Nomenclature Committee*

Sumário

Lista de abreviaturas

Lista de símbolos de genes e proteínas

Lista de figuras

Lista de tabelas

Lista de quadros

Resumo

Abstract

1. Introdução 1
   1.1. Fisiopatologia da OI 3
      1.1.1. Defeitos primários da estrutura e processamento do colágeno tipo 1 4
      1.1.2. Defeitos na modificação pós-traducional do colágeno tipo 1 5
      1.1.3. Defeitos no *crosslinking* e processamento do colágeno tipo 1 5
      1.1.4. Defeitos na formação e mineralização óssea 6
      1.1.5. Defeitos na diferenciação e função dos osteoblastos 7
   1.2. Diagnóstico clínico de OI 8
   1.3. Diagnóstico molecular de OI 9
   1.4. Classificação de OI 10
   1.5. Tratamento e prognóstico 12
   1.6. Justificativa do estudo 14
2. Objetivo 15
3. Materiais e métodos 16
   3.1. Casuística 16
      3.1.1. Diagnóstico clínico de osteogênese imperfeita 16
   3.2. Avaliação clínica da coorte 18
   3.3. Análise molecular 18
      3.3.1. Obtenção de amostras de DNA 18
      3.3.2. Painel de genes candidatos 18
      3.3.3. Sequenciamento paralelo em larga escala 19
      3.3.4. Análise bioinformática dos resultados do SPLE 20
      3.3.5. Sequenciamento confirmatório pelo método de Sanger 22
      3.3.6. Análise de variantes de número de cópia 23
      3.3.7. Análise de segregação familiar 24

3.4 Análise de associação entre achado molecular e características clínicas 25
4. Resultados..... 26
4.1. Coorte..... 26
4.2. Resultados do SPLE..... 35
4.2.1. Variantes encontradas no gene *COL1A1*..... 42
4.2.2. Variantes encontradas no gene *COL1A2*..... 44
4.2.3. Variantes encontradas nos demais genes candidatos 45
4.2.3.1. Variante de número de cópias gênicas 49
4.3. Associação entre achado molecular e características clínicas 51
5. Discussão..... 56
5.1. Variantes identificadas em *COL1A1* e *COL1A2*..... 59
5.2. Contextualização dos achados moleculares..... 61
6. Conclusões..... 67
7. Referências 68

Apêndices

Lista de abreviaturas

5’ UTR	Região 5’ não traduzida
a	Ano
aa	Aminoácidos
ACMG-AMP	Colégio Americano de Genética e Genômica Médica – Associação de Patologia Molecular
AD	Autossômica dominante
AR	Autossômica recessiva
CNV	Variante de número de cópias gênicas
CONTRA	*Copy Number Analysis for Targeted Resequencing*
DGI	Dentinogênese imperfeita
DXA	Absortiometria por dupla fonte de raio-X
F	Feminino
HMZ	Homozigoto
HTZ	Heterozigoto
IU	Intra útero
IGV	*Integrative Genomics Viewer*
m	Mês
M	Masculino
ND	Não disponível
OI	Osteogênese imperfeita
pb	Pares de bases
PCR	Reação em cadeia da polimerase
Pt	Identificador dos pacientes incluídos no estudo
RN	Recém-nascido
RNAm	Ácido ribonucléico mensageiro
SPLE	Sequenciamento paralelo em larga escala
VP	Variante patogênica (classificação de acordo com ACMG)
VPP	Variante provavelmente patogênica (classificação de acordo com ACMG)
VSI	Variante de significado incerto (classificação de acordo com ACMG)
Z	Desvio padrão para idade e sexo

Lista de símbolos de genes e proteínas

COL1A1	Collagen type I alpha 1 chain
COL1A2	Collagen type I alpha 2 chain
CREB3L1	cAMP responsive element binding protein 3 like 1
CRTAP	Cartilage associated protein
FKBP10	FK506-binding protein 10
FKBP65	FK506-binding protein 65 (proteína codificada pelo gene *FKBP10*)
HSP47	Heat shock protein 47 (proteína codificada pelo gene *SERPINH1*)
IFITM5	Interferon induced transmembrane protein 5
LRP5	Low density lipoprotein receptor-related protein 5
LRP6	Low density lipoprotein receptor-related protein 6
MBTPS2	Membrane bound transcription factor peptidase, site 2
OASIS	Old astrocyte specifically induced substance (codificada por *CREB3L1*)
P3H1	Prolyl 3-hydroxylase 1, previamente denominado *LEPRE1*
PEDF	Pigment epithelium-derived factor (proteína codificada por *SERPINF1*)
PLOD2	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase 2
PPIB	Peptidylprolyl isomerase B
SERPINF1	Serpin family F member 1
SERPINH1	Serpin family H member 1
SP7	SP7 transcription factor
SPARC	Secreted protein acidic and cysteine rich
TMEM38B	Transmembrane protein 38B
WNT1	WNT family member 1

Lista de figuras

Figura 1 - Mecanismos fisiopatológicos da OI..... 3
Figura 2 - Heredogramas dos casos familiares..... 30
Figura 3 - Representação das variantes identificadas em COL1A1 e COL1A2 43
Figura 4 - Potencial defeito digênico envolvendo *P3H1* e *WNT1*..... 47
Figura 5 - Estudo por imagem da portadora da variante *IFITM5* c.-14C>T 48
Figura 6 - Deleção homozigótica dos éxons 1 e 2 do gene *TMEM38B* 50
Figura 7 - Panorama da distribuição diagnóstica de OI em diferentes contextos 63

Lista de tabelas

Tabela 1 - Genes associados a osteogênese imperfeita 2
Tabela 2 - Características gerais da coorte 27
Tabela 3 - Características clínicas dos casos isolados 28
Tabela 4 - Características clínicas dos casos familiares 29
Tabela 5 - Dados clínicos adicionais dos pacientes estudados 32
Tabela 6 - Variantes genéticas identificadas na coorte 36
Tabela 7 - Análise de segregação familiar 39
Tabela 8 - Características clínicas dos portadores de COL1A2 p.(Gly772Ser) 45
Tabela 9 - Associação entre defeitos em *COL1A1*, *COL1A2* ou nos demais genes candidatos e achados clínicos 52
Tabela 10 - Associação entre defeitos nos genes do colágeno 1 (*COL1A1* e *COL1A2*) ou nos demais genes candidatos e achados clínicos 53
Tabela 11 - Associação entre defeitos em *COL1A1* ou *COL1A2* e achados clínicos.... 54
Tabela 12 - Associação entre substituição de glicina ou outras variantes em COL1A1 e COL1A2 e achados clínicos 55
Tabela 13 - Estudos de diagnóstico molecular em OI 62

Lista de quadros

Quadro 1 - Classificação da gravidade clínica da OI por Mrosk et al. em 2018 17

Resumo

Fernandes AM. *Diagnóstico molecular de osteogênese imperfeita através do sequenciamento paralelo em larga escala de 15 genes candidatos* [tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2019.

Osteogênese imperfeita (OI) é um conjunto de displasias esqueléticas hereditárias, heterogêneas dos pontos de vista clínico e genético. A maioria dos casos resulta de defeitos no colágeno tipo 1, entretanto diversos outros genes candidatos vêm sendo descritos. Não há definição de critérios mínimos para o diagnóstico clínico de OI, portanto o diagnóstico molecular vem permitir precisão diagnóstica, e melhor compreensão fisiopatológica e prognóstica. O objetivo deste estudo foi identificar o diagnóstico molecular da OI através da análise de genes candidatos por sequenciamento paralelo em larga escala (SPLE), buscando-se associar características clínicas ao achado molecular. Para tanto, as regiões codificadoras e junções íntron-éxon dos 15 genes candidatos *COL1A1*, *COL1A2*, *CRTAP*, *P3H1*, *PPIB*, *SERPINH1*, *FKBP10*, *PLOD2*, *BMP1*, *IFITM5*, *SERPINF1*, *WNT1*, *TMEM38B*, *SP7* e *CREB3L1* foram capturadas e sequenciadas por SPLE na plataforma Illumina. Variantes genéticas raras (frequência alélica <0,5%) com alto impacto sobre a proteína codificada foram confirmadas por sequenciamento Sanger e submetidas a análise de segregação familiar, quando possível. Variantes do número de cópias gênicas foram buscadas utilizando-se o software CONTRA e confirmadas por *array* citogenômico. Foram incluídos no estudo 49 indivíduos com OI, correspondendo a 30 casos isolados e 8 casos familiares, em sua maioria adultos (82%), com mediana de idade de 24 anos. O acometimento esquelético era leve em 37% dos indivíduos, moderado em 30% e grave em 33%. O SPLE foi realizado em todos os 49 pacientes da coorte, com cobertura média entre 354 a 1382 leituras, e superior a 50 em 99% das regiões-alvo. Identificou-se o diagnóstico molecular em 37 dos 38 casos (97%). Em 18 casos foram encontradas variantes em *COL1A1* e em 9 casos em *COL1A2*, totalizando 71% dos casos com defeito no colágeno tipo 1. Variantes nos demais genes candidatos foram encontradas em 10 casos isolados: *SERPINF1* (n=2), *FKBP10* (n=2), *PLOD2* (n=2), *IFITM5* (n=1), *P3H1* (n=1), *TMEM38B* (n=1) e *WNT1* + *P3H1* (n=1). Apenas uma variante de número de cópias foi encontrada (deleção dos éxons 1 e 2 de *TMEM38B*). Das 42 variantes encontradas, 23 já foram descritas anteriormente em associação a OI e 19 são variantes novas. As variantes em COL1A1 e COL1A2 (56% envolvendo troca de glicina) estão dispersas ao longo das proteínas e associadas a grande

variabilidade fenotípica, como demonstrado pela variante COL1A2 p.(Gly772Ser) encontrada em três casos não relacionados com quadros clínicos diversos. Não foram encontradas variantes genéticas adicionais que pudessem explicar a variabilidade fenotípica. Variantes nos genes não colágenos estiveram associados a OI de maior gravidade (p=0,012), variantes em *COL1A1* à manifestação de esclera azulada (p=0,009), e, dentre as variantes de *COL1A1* e *COL1A2*, aquelas que resultaram em substituição de glicina estiveram associadas à dentinogênese imperfeita (p=0,003). Concluindo, este estudo demonstra que o diagnóstico molecular de OI por SPLE é eficaz. Frente à literatura, encontrou-se maior proporção de defeitos em genes não colágenos, que pode ser atribuída à gravidade dos pacientes incluídos ou a características da OI no Brasil, já que nossa população é pouco estudada do ponto de vista molecular. Espera-se que a precisão diagnóstica possibilitada pelo SPLE venha permitir tratamento e seguimento personalizados aos indivíduos com OI.

Descritores: osteogênese imperfeita; diagnóstico; genética; sequenciamento de nucleotídeos em larga escala; colágeno tipo I; fraturas ósseas.

Abstract

Fernandes AM. *Molecular diagnosis of osteogenesis imperfecta through massively parallel sequencing of 15 candidate genes* [thesis]. São Paulo: "Faculdade de Medicina, Universidade de São Paulo"; 2019.

Osteogenesis imperfecta (OI) is a clinically and genetically heterogeneous group of hereditary bone dysplasias. Even though most cases result from defects in type 1 collagen, several other candidate genes are arising. There is no definition of clinical criteria for the diagnosis of OI, hence the advent of molecular diagnosis may allow diagnostic precision, and better understanding of pathophysiology and prognosis. The objective of this study was to identify the molecular diagnosis of OI through the molecular analysis of candidate genes using massively parallel sequencing (MPS), and to associate clinical manifestations to the molecular findings. For these purposes, the coding regions and exon-intron boundaries of the 15 candidate genes *COL1A1*, *COL1A2*, *CRTAP*, *P3H1*, *PPIB*, *SERPINH1*, *FKBP10*, *PLOD2*, *BMP1*, *IFITM5*, *SERPINF1*, *WNT1*, *TMEM38B*, *SP7* and *CREB3L1* were captured and sequenced in the Illumina platform. Rare genetic variants (allelic frequency < 0.5%) impacting on the codified protein were confirmed by Sanger sequencing and submitted to segregation analysis when possible. Copy number variants (CNV) were sought using CONTRA and confirmed by SNP array. Forty-nine individuals with OI were included, corresponding to 30 sporadic and 8 familial cases, being mostly adults (82%) with a median of 24 years of age. OI was mild in 37% of individuals, moderate in 30% and severe in 33%. All 49 patients were analyzed by MPS, with mean coverage ranging from 354 to 1382 reads, and >50 depth in 99% of target regions. A molecular diagnosis was obtained in 37 out of 38 cases (97%). *COL1A1* variants were found in 18 cases, and *COL1A2* variants in 9 cases, meaning that 71% of cases had type 1 collagen-related OI. Variants in the other candidate genes were found in 10 sporadic cases: *SERPINF1* (n=2), *FKBP10* (n=2), *PLOD2* (n=2), *IFITM5* (n=1), *P3H1* (n=1), *TMEM38B* (n=1) and *WNT1* + *P3H1* (n=1). Only one CNV was found (deletion of exons 1 and 2 of *TMEM38B*). Among the 42 identified variants, 23 had already been reported in association to OI and 19 are novel variants. Identified COL1A1 and COL1A2 variants (56% glycine substitution) are scattered throughout the proteins, and associated to wide phenotypic variability, as demonstrated by the COL1A2 p.(Gly772Ser) variant found in three unrelated cases with different clinical presentations. Additional genetic variants that could help explain the phenotypic variability were not found. Variants in non-collagen

genes were associated to severe OI (p=0.012), variants in *COL1A1* to blue sclerae (p=0.009), and, amongst *COL1A1* and *COL1A2* variants, those resulting in glycine substitution were associated to dentinogenesis imperfecta (p=0.003). In conclusion, this study shows that the molecular diagnosis of OI through MPS is effective. In comparison to the literature, a higher proportion of non-collagen defects was found, which can be due to the OI severity of included patients or to a peculiarity of OI in Brazil, given that our population is underrepresented in molecular studies of OI. Hopefully, the diagnostic precision enabled by MPS will allow personalized treatment and follow-up of individuals with OI.

Descriptors: osteogenesis imperfecta; diagnosis; genetics; high-throughput nucleotide sequencing; type I collagen; fractures, bone.

1. Introdução

Osteogênese imperfeita (OI) é um conjunto de displasias esqueléticas hereditárias, heterogêneas dos pontos de vista clínico e genético (Marini et al., 2017; Rauch & Glorieux, 2004). Clinicamente se caracteriza por fraturas de fragilidade, baixa massa óssea, deformidades ósseas e prejuízo no crescimento, com gravidade variando desde casos letais, com fraturas intrauterinas, a casos leves com algumas fraturas sem deformidades ou perda de funcionalidade. Manifestações típicas extraesqueléticas podem estar presentes como a esclera azulada, dentinogênese imperfeita (DGI), hiperelasticidade de ligamentos e pele, perda auditiva e prejuízo de função cardio-pulmonar. A incidência é estimada em aproximadamente 1 a cada 10.000 nascimentos (Forlino et al., 2011).

A causa da maioria dos casos de osteogênese imperfeita reside direta ou indiretamente em alterações do colágeno tipo 1, proteína mais abundante na matriz óssea, que vão desde anormalidades em sua estrutura primária, alterações quantitativas da proteína, modificações pós-traducionais em sua estrutura terciária, transporte intracelular e incorporação à matriz (Forlino et al., 2011). Mais recentemente, vias fisiopatológicas independentes do colágeno tipo 1, como defeitos em fatores locais da mineralização óssea ou na diferenciação de osteoblastos, foram descritas, expandindo as bases moleculares da OI (Tabela 1). Mais de 1500 variantes patogênicas nos genes *COL1A1* e *COL1A2*, que codificam as cadeias α1 e α2 do colágeno tipo 1, já foram identificadas, além de variantes patogênicas em outros genes relacionados a OI.

Tabela 1 - Genes associados a osteogênese imperfeita

Gene	ID transcrito	Éxons (n)	Transcrito (pb)	Proteína (aa)	Herança	Fenótipo	Defeito fisiopatológico
COL1A1	ENST00000225964	51	6727	1464	AD	Leve a letal	Colágeno tipo 1
COL1A2	ENST00000297268	52	5411	1366	AD		
CRTAP	ENST00000320954	7	6630	401	AR	Grave a letal	3α-hidroxilação do colágeno tipo 1
P3H1	ENST00000236040	14	2993	804	AR		
PPIB	ENST00000300026	5	1081	216	AR	Moderado a letal	
SERPINH1	ENST00000524558	5	3391	418	AR	Grave	Chaperonas do colágeno tipo 1
FKBP10	ENST00000321562	10	2658	582	AR	Deformidades progressivas, Sd. de Bruck	
PLOD2	ENST00000282903	20	3732	758	AR		
BMP1	ENST00000306385	20	4229	986	AR	Moderado a grave, associado a hérnia umbilical	Clivagem do propeptídeo C-terminal do colágeno
IFITM5	ENST00000382614	2	736	132	AD	Histologia em malha, calcificação de membrana interóssea e calo hipertrófico	Controle local da mineralização
SERPINF1	ENST00000254722	8	1548	418	AR	Histologia em escamas de peixe	
WNT1	ENST00000293549	4	1185	370	AR	Moderado, deformidades progressivas	Via de sinalização WNT
TMEM38B	ENST00000374692	6	3525	291	AR	Moderado a grave	Diferenciação de osteoblastos
SP7	ENST00000536324	3	3148	418	AR	Moderado	
CREB3L1	ENST00000529193	12	2687	519	AR	Grave a letal	Regulação da expressão de *COL1A1*

Atualizado a partir de (Marini et al., 2017) e Online Mendelian Inheritance in Man (OMIM). ID transcrito, identificador do transcrito principal na base de dados Ensembl; pb, pares de bases; aa, aminoácidos; AD, autossômica dominante; AR, autossômica recessiva.

1.1. Fisiopatologia da OI

A busca pelo diagnóstico molecular em indivíduos com OI ampliou marcadamente o conhecimento sobre a sua fisiopatologia. Inicialmente, OI foi atrelada a defeitos da composição da matriz óssea, diretamente ou indiretamente relacionados ao colágeno tipo 1. Entretanto, com o maior conhecimento das bases moleculares da OI, sabe-se hoje que a doença também pode decorrer de alterações da mineralização da matriz óssea ou da função dos osteoblastos. Em 2017, Marini e colaboradores propuseram cinco vias fisiopatológicas principais que levam ao quadro clínico de OI, envolvendo os genes associados à doença (Figura 1) (Marini et al., 2017).

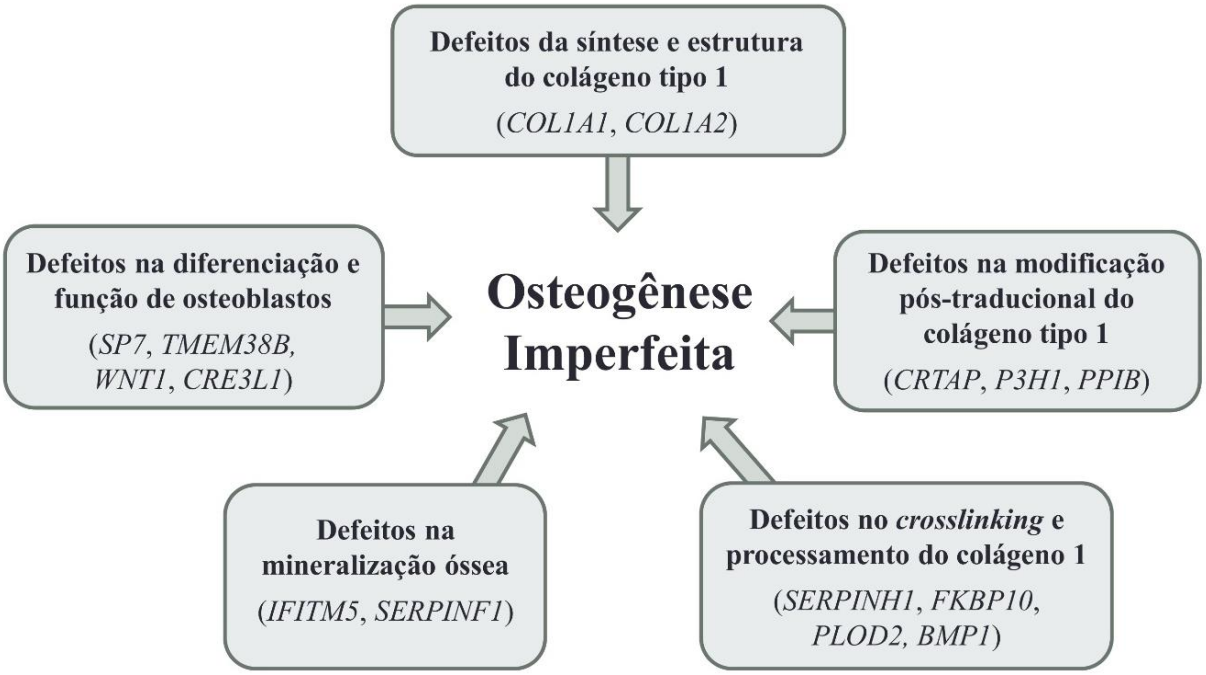

Figura 1 - Mecanismos fisiopatológicos da OI
Esquema ilustrativo baseado nos mecanismos fisiopatológicos propostos por Marini e coautores em 2017 (Marini et al., 2017).

1.1.1. Defeitos primários da estrutura e processamento do colágeno tipo 1

O colágeno tipo 1 é formado a partir do pró-colágeno, um heterodímero composto por duas cadeias pró-α1 e uma cadeia pró-α2, codificadas pelos genes *COL1A1* e *COL1A2* respectivamente. O posicionamento do aminoácido glicina a cada três resíduos permite a estrutura espacial em tripla-hélice (Myllyharju & Kivirikko, 2004). O pró-colágeno sofre diversas modificações pós-traducionais até ser incorporado à matriz óssea.

Variantes com ganho de parada nos genes *COL1A1* e *COL1A2*, deflagrando mecanismo de controle de qualidade de RNAm e levando à produção de menor quantidade de cadeias α, resultam em defeito quantitativo de colágeno tipo 1. O quadro clínico resultante é leve, no qual as fraturas se iniciam com a deambulação e diminuem após a puberdade. Esclera azulada e perda auditiva são comuns (Willing et al., 1996).

Alterações qualitativas ou estruturais do colágeno tipo 1, frequentemente por substituição da glicina (80%) ou alterações de *splicing* (20%) prejudicam a formação da tripla-hélice, alterando a estrutura da proteína e consequentemente prejudicando a secreção e processamento do colágeno (Marini et al., 2007). Acredita-se que a formação de cadeias defeituosas de colágeno no retículo endoplasmático ative respostas de controle de qualidade que resultem em estresse local e até mesmo apoptose (Forlino et al., 2011). Ainda, a alteração da composição proteica da matriz extracelular resultante do defeito estrutural do colágeno interfere na interação entre as fibrilas, na mineralização e na sinalização entre osteócitos, osteoblastos, osteoclastos e matriz. A gravidade da repercussão celular e tecidual depende do defeito molecular encontrado justificando a variabilidade fenotípica de quadros com poucas fraturas e sem deformidades a quadros letais (Ben Amor et al., 2011; Rauch et al., 2010).

Variantes patogênicas nos genes do colágeno tipo 1 apresentam padrão de herança autossômica dominante. A melhor compreensão das modificações pós-traducionais do

colágeno permitiu esclarecer a fisiopatologia de casos raros de OI não relacionados a variantes patogênicas nos genes *COL1A1* e *COL1A2*, com herança autossômica recessiva.

1.1.2. Defeitos na modificação pós-traducional do colágeno tipo 1

A colágeno 3-propil-hidroxilase é um complexo proteico composto por 3 proteínas codificadas pelos genes *CRTAP*, *P3H1* (*LEPRE1*) e *PPIB*, presente no retículo endoplasmático, responsável por hidroxilar certos resíduos prolina da molécula de colágeno permitindo a formação da estrutura em tripla-hélice da proteína. Variantes patogênicas nesses genes levam a quadro clínico grave, com fragilidade óssea, fraturas desde o nascimento, comprometimento da estatura e alto índice de mortalidade (Morello et al., 2006; Cabral et al., 2007; van Dijk et al., 2009).

1.1.3. Defeitos no *crosslinking* e processamento do colágeno tipo 1

Ainda no retículo endoplasmático encontram-se chaperonas responsáveis pela isomerização da prolina, também necessária para a formação da estrutura em tripla-hélice do colágeno tipo 1. Defeitos nas proteínas HSP47 e FKBP65, codificadas respectivamente pelos genes *SERPINH1* e *FKBP10*, levam a quadros clínicos moderados a graves, com deformidades progressivas, baixa estatura e esclera e dentição normais (Christiansen et al., 2010; Alanay et al., 2010). A Síndrome de Bruck, doença recessiva caracterizada por contraturas congênitas e quadro grave de fragilidade óssea também está relacionada a variantes patogênicas nos genes *FKBP10* e *PLOD2*. Como a gravidade das contraturas é bastante variável e a fragilidade óssea é sempre significativa, os diagnósticos de OI e Síndrome de Bruck se sobrepõem (Puig-Hervas et al., 2012; Leal et al., 2018).

O gene *BMP1* codifica uma protease responsável pela clivagem do propeptídeo C-terminal do pró-colágeno tipo 1 para sua maturação em colágeno tipo 1 após secretado, e

está envolvido na fisiopatologia de casos raros de OI. Variantes patogênicas em homozigose deste gene acarretam quadros graves da doença com deformidades e hiperflexibilidade articular (Martinez-Glez et al., 2012).

1.1.4. Defeitos na formação e mineralização óssea

Além das variantes patogênicas relacionadas às alterações pós-traducionais do colágeno, defeitos no controle local da mineralização da matriz óssea podem resultar em OI. Uma variante patogênica na região 5' não traduzida do gene *IFITM5* gerando um novo códon de início e adicionando 5 novos resíduos a proteína, leva a ganho de função e se relaciona a uma apresentação clínica específica de OI de herança autossômica dominante, podendo corresponder a cerca de 5% dos casos em algumas casuísticas (Cho et al., 2012; Semler et al., 2012). A fragilidade óssea de moderada a grave é acompanhada de calcificação da membrana interóssea do antebraço, luxação da cabeça do rádio e banda radiodensa metafisária. Cerca de metade dos pacientes com essa variante patogênica apresenta calos hiperplásicos na consolidação de fraturas e a histologia típica em malha (*mesh-like*) do osso lamelar está presente em 100% dos casos (Glorieux et al., 2000). O gene *IFITM5* codifica uma proteína de membrana presente nos osteoblastos que postula-se estar associada ao controle de mineralização.

Variantes patogênicas no gene *SERPINF1*, com padrão de herança autossômico recessivo, levam a deficiência de mineralização óssea, com quadro de fragilidade óssea moderada a grave, deformidades progressivas e histologia típica do osso lamelar em escamas de peixe. Tal gene codifica a proteína PEDF, que exerce papel na homeostase óssea incluindo o controle da mineralização da matriz osteoide (Becker et al., 2011; Homan et al., 2011).

1.1.5. Defeitos na diferenciação e função dos osteoblastos

Alguns genes estão relacionados a OI por participarem da diferenciação dos osteoblastos. O mecanismo mais bem compreendido é o do gene *WNT1*, que sintetiza uma glicoproteína que interage com receptores de membrana LRP5 e LRP6, desencadeando a via de sinalização da β-catenina que tem como resultado final a ativação da expressão gênica para a formação óssea. Variantes patogênicas em heterozigose no *WNT1* foram relacionadas a quadros de osteoporose, enquanto que variantes patogênicas em homozigose relacionam-se a quadros graves de OI, com múltiplas fraturas e baixa estatura (Keupp et al., 2013; Laine et al., 2013; Palomo et al., 2014).

Variantes patogênicas no gene *TMEM38B*, que codifica um canal catiônico presente na membrana do retículo endoplasmático relacionado ao efluxo de cálcio e está envolvido na diferenciação de osteoblastos, também se relacionam a casos de OI. Nesses casos o quadro clínico é moderado, com deformidades ósseas progressivas. Esclera azulada pode estar presente (Shaheen et al., 2012). Outro gene relacionado a diferenciação de osteoblastos é o *SP7*, que codifica um fator de transcrição do tipo dedos de zinco da linhagem osteoblástica, levando a um quadro de OI moderado a grave (Lapunzina et al., 2010).

Já o gene *CREB3L1* codifica a proteína OASIS que, após sofrer proteólise, libera seu domínio N-terminal que se transloca até o núcleo ativando a expressão do gene *COL1A1*. Em raros casos, deleções em homozigose foram descritas em quadros graves de OI com alta taxa de mortalidade (Symoens et al., 2013).

1.2. Diagnóstico clínico de OI

As características clínicas mais marcantes para o diagnóstico de OI são a ocorrência de múltiplas fraturas de fragilidade desde a infância, presença de deformidades esqueléticas e déficit de crescimento (Forlino & Marini, 2016). Em radiografias pode-se observar osteopenia difusa e deformidades não necessariamente observadas ao exame físico, como encurvamento de ossos longos, escoliose e cifose; além de fraturas vertebrais assintomáticas. A densitometria óssea pode auxiliar o diagnóstico mostrando uma baixa densidade mineral óssea para a idade (Forlino & Marini, 2016).

Quando a esclera azulada ou dentinogênese imperfeita (descoloração e translucidez dos dentes com predisposição a cáries e quebra, DGI) estão presentes, o diagnóstico clínico de OI é mais fácil. Entretanto, em muitos casos estes comemorativos podem estar ausentes. Por outro lado, esclera azulada ou mais escura é comum em crianças saudáveis, e nos pacientes com OI ela tende a amenizar ou até desaparecer com o passar da idade. Já a DGI é mais comum na primeira dentição, sendo rara na dentição permanente (Rauch & Glorieux, 2004). A perda auditiva progressiva é um elemento que corrobora o diagnóstico, mas comumente ocorre tardiamente, a partir da 2ª a 4ª décadas de vida.

O diagnóstico de OI se torna simples em indivíduos com história familiar positiva ou com diversas características clínicas, mas pode ser difícil se não houver familiares afetados e a história de fragilidade óssea não estiver associada às manifestações extraesqueléticas típicas. Nos quadros mais leves de OI o quadro clínico pode ser facilmente confundido com osteoporose de início precoce em adultos ou abuso físico em crianças. Não existe consenso quanto a critérios clínicos mínimos para o diagnóstico de OI, o que torna difícil o diagnóstico em alguns casos.

1.3. Diagnóstico molecular de OI

O diagnóstico molecular de OI foi inicialmente buscado através de culturas de fibroblastos da derme obtidos de biopsia cutânea. Estudos mostraram que em culturas celulares de pacientes com OI os fibroblastos não apresentavam o formato fusiforme dos fibroblastos controles, a densidade celular era menor quando se atingia uma confluência celular nas culturas e a proliferação mais lenta (Boright et al., 1984). Posteriormente, analisou-se a produção de pró-colágeno tipo 1 e colágeno tipo 1 pelos fibroblastos em cultura através da separação eletroforética das cadeias da proteína em gel de poliacrilamida. Em um grupo notou-se a produção de metade da quantidade esperada de pró-colágeno tipo 1, sem alterações na estrutura da proteína e redução na proporção de pró-colágeno tipo 1 e pró-colágeno tipo 3 produzidos pelos fibroblastos. Em outro grupo evidenciou-se a produção de proteínas com maior peso molecular devido a produção de cadeias α1 ou α2 alteradas o que resulta em alterações da estrutura do trímero formado por essas cadeias para dar origem ao pró-colágeno tipo 1. E, ainda, parte dos indivíduos com diagnóstico clínico de OI não apresentavam alteração na quantidade ou estrutura do pró-colágeno produzido pelos fibroblastos da cultura (Wenstrup et al., 1990).

Assim, o diagnóstico molecular por cultura de fibroblastos, além de ser um método diagnóstico invasivo exigindo a realização de uma biopsia cutânea, demorado pois o crescimento destas células em cultura é lento (8 a 10 semanas), e exigir uma capacitação especial para cultura das células e interpretação do resultado, só esclarece diagnóstico se houver alteração da quantidade ou estrutura do pró-colágeno tipo 1. Casos relacionados a variantes patogênicas em outros genes que não o *COL1A1* e *COL1A2* poderão apresentar resultados normais nessa avaliação.

Buscou-se então diagnóstico molecular através de sequenciamento gênico pela técnica de Sanger. Nesse caso, encontra-se dificuldade no fato de as regiões codificadoras

dos genes *COL1A1* e *COL1A2* serem muito extensas (Tabela 1) e não apresentarem *hot spots*, ou seja, não existem regiões onde as variantes patogênicas ocorram com maior frequência; de fato, as centenas de variantes patogênicas já descritas relacionadas a OI se distribuem ao longo das regiões codificadoras. Além disso, há pelo menos outros 13 genes relacionados a OI na literatura que devem ser investigados na ausência de alterações nos genes do colágeno tipo 1 (Marini & Blissett, 2013).

Assim buscam-se alternativas de diagnóstico molecular capaz de avaliar de maneira mais eficaz os diversos genes que podem estar envolvidos na fisiopatologia da doença.

1.4. Classificação de OI

A classificação dos pacientes em tipos distintos de OI foi considerada útil para avaliação de prognóstico, indicação de intervenções terapêuticas e predição de resposta a estas intervenções (Rauch & Glorieux, 2004). A primeira classificação utilizada foi a de Sillence e colaboradores, que divide os pacientes em 4 tipos clínicos distintos, com padrão de herança autossômica dominante (Sillence et al., 1979). Historicamente, variantes nos genes *COL1A1* e *COL1A2* foram associadas a estes quatro tipos de apresentação clínica. Nesta classificação, OI tipo 1 inclui casos leves da doença, quando fraturas não costumam resultar em deformidades ósseas e tendem a iniciar com a deambulação, diminuindo após a puberdade. OI tipo 2 é letal no período intrauterino ou perinatal. OI tipo 3, forma mais grave dentre as crianças, é caracterizada por baixa estatura extrema e graves deformidades ósseas decorrentes de múltiplas fraturas. Pacientes com deformidades ósseas leves a moderadas e acometimento variável da estatura são classificados como tipo 4. No tipo 4 inclui-se todo paciente que não se encaixou nos 3 tipos anteriormente descritos (Sillence et al., 1979).

Posteriormente, outros tipos surgiram e foram numericamente classificados conforme eram descritos de acordo com o gene associado, com características clínicas e histológicas variadas e padrão de herança predominantemente recessivo (à exceção do tipo 5). Tendo como base o defeito genético, atingiu-se um total de 12 diferentes tipos de OI e mais 4 tipos numericamente não classificados, tornando a classificação extensa e com grande sobreposição de apresentação clínica, dificultando seu uso na prática (Marini & Blissett, 2013).

Em 2010 começaram a surgir propostas de novas classificações para OI, tomando como base a classificação de Sillence, considerando as manifestações clínicas e aspectos radiológicos, já que muitos dos mais recentes tipos de OI não se diferenciavam dos tipos 1 a 4 da classificação inicial (Van Dijk et al., 2010; Warman et al., 2011). O Grupo de Nosologia da Sociedade Internacional das Displasias Esqueléticas passou a recomendar a classificação de OI em 5 tipos, adicionando aos 4 tipos de Sillence o tipo 5 por suas manifestações radiológicas específicas. Adota-se uma classificação fenotípica onde os grupos apresentam heterogeneidade genotípica considerável, na intenção de facilitar a avaliação da evolução e prognóstico, indicação e resposta terapêutica (Van Dijk & Sillence, 2014).

Em 2017, Marini e colaboradores separam os casos de OI conforme o mecanismo fisiopatológico da doença em 5 grupos, onde no primeiro grupo são incluídos casos relacionados a defeitos na síntese e estrutura do colágeno (*COL1A1* e *COL1A2*), no segundo, casos relacionados a defeitos nas modificações pós-traducionais do colágeno (*CRTAP*, *P3H1*, *PPIB*), no terceiro, casos relacionados a defeitos no processamento e *crosslinking* do colágeno (*SERPINH1*, *FKBP10*, *PLOD2* e *BMP1*), no quarto, casos relacionados a defeitos na formação e mineralização óssea (*IFITM5* e *SERPINF1*) e no quinto grupo, casos relacionados a defeitos na diferenciação e função dos osteoblastos

(*WNT1*, *CREB3L1* e *SP7*) (Marini et al., 2017). Tal classificação auxilia no entendimento do mecanismo de cada caso, e potencialmente, em abordagens terapêuticas direcionadas, porém não se mostra útil na prática clínica na ausência de diagnóstico molecular.

Mais recentemente, vem prevalecendo na literatura a classificação clínica de casos de OI de acordo com a gravidade da doença esquelética: leve, moderada ou grave. Elementos objetivos desta classificação ainda vem sendo aprimorados (Mrosk et al., 2018).

1.5. Tratamento e prognóstico

Até o momento, não existe cura para OI; o manejo da doença é baseado nos sintomas de acordo com a gravidade do acometimento esquelético e comorbidades (Marini et al., 2017). O objetivo do tratamento é garantir a mobilidade e capacidade funcional, embasando-se em tratamento ortopédico, fisioterápico e de reabilitação para correção e prevenção de deformidades (Forlino et al., 2011; Rauch & Glorieux, 2004).

Entretanto, estes procedimentos não melhoram a resistência óssea e, portanto, há bastante interesse em terapia medicamentosa capaz de reduzir fraturas. Neste sentido, desde o final da década de 1980, bisfosfonatos vem sendo utilizados na terapia medicamentosa da OI (Rauch & Glorieux, 2004; Palomo et al., 2017). Poucas são as evidências provenientes de estudos clínicos randomizados controlados por placebo demonstrando efeito da terapia antirreabsortiva com bisfosfonatos na prevenção de fraturas em pacientes com OI (Rauch & Glorieux, 2004; Forlino et al., 2011). Em comparação a controles históricos, o uso de bisfosfonatos foi relacionado a melhora da dor óssea, sensação de bem estar, maior altura final atingida, aumento da força muscular, melhora da massa óssea em coluna e ossos longos, assim como redução da taxa de fraturas (Cundy, 2012; Palomo et al., 2015).

Com o avanço no entendimento das bases moleculares da OI, propõe-se que novas medidas terapêuticas, mais adequadas ao defeito molecular de base, possam ser desenvolvidas. Neste sentido, terapia com inibidores de esclerostina se mostrou eficaz em modelos animais de OI relacionada a defeitos em WNT1 (Sinder et al., 2016), propiciando investigação em humanos (Glorieux et al., 2017), e a desregulação da sinalização por TGF-β identificada em formas graves de OI por defeitos no colágeno (Grafe et al., 2014) embasa a investigação de terapia anti-TGF-β nestes casos (Lee, 2019).

Do ponto de vista prognóstico, em casos com fenótipos mais graves, com numerosas fraturas e deformidades, a expectativa de vida é reduzida e apenas 20% dos casos atingem a idade adulta (Cundy, 2012). A principal causa de mortalidade está relacionada a deformidades torácicas com redução da capacidade pulmonar e hipertensão de artéria pulmonar. Já nos casos leves a moderados observa-se redução da ocorrência de fraturas após a puberdade, podendo ocorrer novo aumento da taxa de fraturas após a menopausa nas mulheres e em idade avançada nos homens (Cundy, 2012). Dados do estudo retrospectivo dinamarquês envolvendo 687 pacientes com acometimento predominantemente leve revelaram redução da expectativa de vida e mortalidade relacionada a causas respiratórias, gastrointestinais e trauma (Folkestad et al., 2016). Os estudos de evolução e prognóstico de OI até aqui publicados não distinguem os pacientes de acordo com o defeito molecular de base.

1.6. Justificativa do estudo

A obtenção de diagnóstico molecular de OI na prática clínica traz precisão ao reconhecimento da doença, já que, conforme foi discutido, o diagnóstico puramente clínico de OI pode ser difícil em alguns casos, ou mesmo não reconhecido. Além disso, a identificação de variantes genéticas causando OI na nossa população contribui para o melhor entendimento da doença em nosso meio.

O acúmulo de conhecimento associando as manifestações clínicas esqueléticas e extra-esqueléticas de OI e a evolução da doença ao defeito molecular de base propicia embasamento para melhor previsão prognóstica e planejamento do seguimento a longo prazo, incluindo o rastreio de comorbidades e a indicação de exames subsidiários. Adicionalmente, possibilita adequado aconselhamento genético de portadores de OI e seus familiares.

Por fim, análises do efeito de diferentes terapias de acordo com o defeito molecular potencialmente permitirão escolha terapêutica individualizada e desenvolvimento de novos tratamentos com atuação direcionada à alteração encontrada. O diagnóstico molecular é um primeiro passo nessa direção.

2. Objetivo

Identificar o diagnóstico molecular em pacientes com osteogênese imperfeita através da pesquisa de variantes em genes candidatos utilizando sequenciamento paralelo em larga escala, buscando-se associar características clínicas ao achado molecular.

3. Materiais e métodos

3.1. Casuística

Pacientes com diagnóstico clínico de osteogênese imperfeita em seguimento no ambulatório de doenças osteometabólicas da Divisão de Endocrinologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HCFMUSP) foram convidados a participar do projeto de pesquisa "Diagnóstico molecular de osteogênese imperfeita através de sequenciamento de nova geração", aprovado pela CAPPesq HCFMUSP em 25/05/2015 (Registro CAAE: 43319415.2.0000.0068) e somente foram incluídos no estudo após assinatura do termo de consentimento livre e esclarecido. Também foram incluídos familiares de pacientes para estudo de segregação genética de variantes encontradas, após obtenção de consentimento informado.

3.1.1. Diagnóstico clínico de osteogênese imperfeita

Conforme discutido, não há definição de critérios clínicos mínimos para o diagnóstico de OI. Assim, foram selecionados para o presente estudo indivíduos nos quais o diagnóstico clínico de osteogênese imperfeita foi baseado em história de repetidas fraturas de fragilidade desde a infância e baixa densidade mineral óssea, associados, ou não, a deformidades esqueléticas, baixa estatura, antecedente familiar de fraturas de fragilidade na infância, esclera azulada atual ou pregressa, ou DGI atual ou pregressa.

Foram consideradas como fraturas de fragilidade aquelas decorrentes de injúria insuficiente para fraturar o osso normal, como, por exemplo, atividades cotidianas ou queda da própria altura (Brown & Josse, 2002; Marini & Blissett, 2013).

A densidade mineral óssea foi avaliada por absortiometria por dupla fonte de raio-X (DXA, densitometria óssea), seguindo-se as orientações vigentes (Baim et al., 2008). Baixa densidade mineral óssea para a idade foi definida por Z-score < -2,0 (em relação a

indivíduos de mesmo sexo e faixa etária) na coluna lombar (L1-L4) ou no corpo total (Marini et al., 2017).

Deformidades foram investigadas através de inspeção ao exame físico e radiografias dos ossos acometidos. A estatura foi avaliada utilizando-se estadiômetro de parede quando o indivíduo foi capaz de ficar em posição ortostática, ou fita métrica. Baixa estatura foi definida quando o escore de desvio-padrão da altura para idade e sexo (Z) foi inferior a -2,0, a partir dos dados da Organização Mundial de Saúde. Esclera azulada e DGI foram avaliadas por inspeção ao exame físico, e interrogatório.

A gravidade clínica da doença foi atribuída de acordo com o sistema de pontuação proposto por Mrosk e colaboradores em 2018 (Mrosk et al., 2018), conforme o Quadro 1.

Quadro 1 - Classificação da gravidade clínica da OI por Mrosk et al. em 2018

Critérios: Ritmo de fraturas (1–3 pontos) + Fratura vertebral (0–1 ponto) + Deformidades (0–4 pontos) + Escoliose (0–1 ponto) + Número de fraturas (1–3 pontos)
1º critério: Ritmo de fraturas (número de fraturas por ano) ≤1 fratura = 1 ponto; 2 a 3 fraturas = 2 pontos; >3 fraturas = 3 pontos 2º critério: Presença de fratura vertebral Ausente = 0; presente = 1 ponto 3º critério: Deformidades de ossos longos Ausente = 0; presente em tíbia, fêmur, úmero ou antebraço = 1 ponto cada (1-4) 4º critério: Presença de escoliose Ausente = 0; presente = 1 ponto 5º critério: Número de fraturas na vida ≤10 fraturas = 1 ponto; 10 a 30 fraturas = 2 pontos; ≥30 fraturas = 3 pontos
Classificação final (total de pontos): **OI leve: 1-4 pontos; OI moderada: 5-8 pontos; OI grave: 9-12 pontos**

3.2. Avaliação clínica da coorte

Foram coletados dados retrospectivos do acompanhamento clínico e de exames subsidiários para caracterização dos pacientes e busca de correlação com os achados moleculares: idade à primeira fratura; número total de fraturas; grau de mobilidade; histórico ou presença de esclera azulada, DGI e deformidades esqueléticas; estatura; avaliação de massa óssea em L1-L4 e/ou corpo total por DXA; avaliação audiométrica; avaliação cardíaca por ecocardiograma transtorácico; histórico de tratamento ortopédico e de tratamento medicamentoso com bisfosfonatos.

3.3. Análise molecular

O diagnóstico molecular foi buscado através de sequenciamento paralelo em larga escala (SPLE) de um painel de genes candidatos para OI. Variantes consideradas potencialmente associadas a OI foram confirmadas por sequenciamento Sanger.

3.3.1. Obtenção de amostras de DNA

Após a assinatura do termo de consentimento informado, foram coletados 16 mL de sangue de cada paciente ou familiar para extração de DNA genômico leucocitário através de adaptação da técnica de *salting out*. A extração foi realizada pelo técnico Rodolfo Batista no Laboratório de Investigação Médica 18 (LIM-18). As amostras foram anonimizadas e identificadas por números para armazenamento a -20°C até utilização.

3.3.2. Painel de genes candidatos

Até a elaboração deste projeto, 15 genes candidatos haviam sido implicados na gênese molecular de OI e foram incluídos no painel de sequenciamento: *COL1A1*, *COL1A2*, *CRTAP*, *P3H1*, *PPIB*, *SERPINH1*, *FKBP10*, *PLOD2*, *BMP1*, *IFITM5*,

SERPINF1, *WNT1*, *TMEM38B*, *SP7* e *CREB3L1*. Dados destes quinze genes e seus transcritos principais (canônicos) estão apresentados na Tabela 1.

3.3.3. Sequenciamento paralelo em larga escala

As regiões genômicas de interesse foram capturadas utilizando-se a plataforma SureSelect^{XT} (Agilent Technologies, Santa Clara, EUA). Sondas de captura foram desenhadas com o auxílio da ferramenta on-line SureDesign (Agilent) com base na versão 19 do genoma humano (GRCh37, Fev2009), produzidas a partir de cRNA biotinilado com 120 nucleotídeos cada, cobrindo as regiões codificadoras dos genes candidatos e, no mínimo, 25 pares de bases das junções íntron-éxon. Adicionalmente, foram desenhadas sondas para captura da região 5' não traduzida (5'-UTR) do gene *IFITM5*, uma vez que a variante patogênica associada a OI neste gene candidato está situada 14 pares de base à montante do sítio de início da tradução. Cada nucleotídeo da região genômica de interesse foi coberto por pelo menos 3 sondas (*tiling* de 3 vezes). A cobertura dos genes foi verificada manualmente, garantindo cobertura superior a 99%.

Amostras de 3 µg de DNA genômico foram submetidas a ultrassonicação centrada (E220 Focused Ultrasonicator Covaris, Woburn, EUA) a fim de gerar bibliotecas de DNA genômico fragmentado (fragmentos de 150 a 200 pb), de acordo com instruções do fabricante. Durante o procedimento de captura, identificadores foram ligados aos fragmentos (*barcoding*) permitindo múltiplas análises em uma mesma corrida.

As sequências capturadas foram analisadas na plataforma de SPLE NextSeq 500 (Illumina, San Diego, EUA) no Laboratório de Sequenciamento em Larga Escala (SELA) do Programa Rede de Equipamentos Multiusuários (PREMiUM) da Faculdade de Medicina da USP em colaboração com a Dra. Mônica França. Foi realizado o

sequenciamento em ambas as direções (*paired-end*) dos fragmentos resultando em arquivos FastQ.

3.3.4. Análise bioinformática dos resultados do SPLE

Os resultados do SPLE foram analisados em colaboração com o Dr. Antônio Lerário, da Universidade de Michigan, utilizando-se um pipeline próprio. A partir dos arquivos FastQ gerados pelo sequenciador, foi realizado o alinhamento dos fragmentos sequenciados com o software BWA (Burrows-Wheeler Aligner), utilizando como referência a versão 19 do genoma humano (GRCh37, Fev2009). Após controle de qualidade dos arquivos BAM gerados pelo alinhamento, variantes de ponto e pequenas indels foram identificadas utilizando-se o software Platypus. Gerou-se uma planilha com todas as variantes identificadas nos indivíduos analisados em uma mesma corrida contendo informações relativas às variantes, utilizando-se a ferramenta ANNOVAR. Os indivíduos desse projeto foram analisados em duas corridas, a primeira incluindo 27 pacientes e a segunda, 22 pacientes.

A partir dos arquivos VCF anotados, buscou-se identificar variantes com potencial patogênico que pudessem justificar o quadro clínico de OI levando-se em consideração a frequência populacional das variantes, o impacto sobre a função da proteína codificada e relatos prévios da variante em associação com OI.

Diversos bancos de dados criados a partir de exomas e/ou genomas sequenciados em diferentes populações são acessíveis e nos permitem estimar a frequência das variantes nessas populações. Um exemplo é o Arquivo Brasileiro Online de Mutações (ABraOM), que reúne dados do sequenciamento exômico de 609 idosos brasileiros (Naslavsky et al., 2017).

Variantes gerando códon de parada, ou de início, alterações em sítios de *splicing*, mudança no quadro de leitura dos códons (*frameshift*) ou grandes deleções ou duplicações (variantes de número de cópias, *copy number variant*, CNV) tendem a levar à perda de função. Já variantes intrônicas e as variantes de ponto sinônimas, caracterizadas por troca de nucleotídeo mantendo o aminoácido da proteína, costumam ser toleradas. Para avaliação do potencial potegênico das variantes de ponto que levam à troca de aminoácido na proteína (variantes não sinônimas), foram utilizadas ferramentas *in silico* que integram diferentes métodos para prever o impacto da troca do aminoácido na função da proteína.

No presente estudo, dentre as ferramentas *in silico* disponíveis, utilizou-se o S*orting Itolerant from Tolerant* (SIFT), que classifica a variante como deletéria ou tolerada baseando-se no grau de conservação da sequência proteica (Kumar et al., 2009); o *Polymorphism phenotyping version 2* (PolyPhen2), que divide as variantes em deletérias, possivelmente deletérias ou benignas de acordo com o efeito sobre a sequência e estrutura da proteína (Adzhubei et al., 2013) e o *Combined Annotation-Dependent Depletion* (CADD), que gera uma pontuação (*C-score*) às variantes de acordo com conservação, dados funcionais e análise de outras ferramentas *in silico*. Variantes com pontuação ≥10 representam variantes entre as 10% mais deletérias e ≥20 entre as 1% mais deletérias. O valor de corte sugerido para classificação da variante como deletéria é pontuação ≥15 (Kircher et al., 2014).

Para a busca de descrição prévia das variantes em associação com OI, além da pesquisa na literatura, bancos de dados *online* forma consultados, como *Osteogenesis Imperfecta Variant Database* (OIVD, http://oi.gene.le.ac.uk), *Online Mendelian Inheritance in Man* (OMIM), *Human Gene Mutation Database* (HGMD) e ClinVar (Amberger et al., 2015; Krawczak et al., 2000; Landrum et al., 2016).

Neste projeto, os seguintes critérios foram utilizados para a priorização de variantes:

- Frequência populacional menor ou igual a 0,5% nas bases de dados 1000Genome ExAC (Exome Aggregation Consortium) e ABraOM;
- Variantes localizadas em região exônica codificadora ou em sítio de *splicing* (com a exceção da região 5'UTR do gene *IFITM5*);
- Variantes com impacto na proteína codificada, gerando códon de parada, erro no quadro de leitura (*frameshift*) ou troca de aminoácido codificado (não sinônima);
- Variantes não sinônimas com predição *in silico* de efeito deletério utilizando-se os softwares SIFT, PolyPhen2e CADD;
- Variantes previamente descritas nas bases de dados PubMed, OIVD, OMIM, ClinVar e HGMD.

A partir desta priorização, as variantes de interesse foram confirmadas visualmente utilizando o software Integrative Genomics Viewer (IGV) a partir do arquivo BAM, e classificadas quanto ao seu potencial patogênico de acordo com as diretrizes propostas pelo Colégio Americano de Genética e Genômica Médica – Associação de Patologia Molecular (ACMG-AMP) (Richards et al., 2015) em variante patogênica (VP), variante provavelmente patogênica (VPP) ou variante de significado incerto (VSI).

3.3.5. Sequenciamento confirmatório pelo método de Sanger

As variantes encontradas foram confirmadas nos pacientes e pesquisadas em seus controles familiares por sequenciamento automático pelo método de Sanger. Resumidamente, foram desenhados oligonucleotideos para permitir a amplificação por reação em cadeia da polimerase (PCR) das regiões genômicas de interesse utilizando o software Primer3Plus. Para a PCR, utilizou-se o reagente PCR MasterMix (Promega, Madison, EUA) e 50 ng de DNA genômico, seguindo o protocolo do fabricante; as reações foram realizadas no termociclador Veriti (Thermo Fisher, Waltham, EUA). Após

a verificação da especificidade e rendimento da reação por eletroforese em géis de agarose, os produtos de PCR foram purificados enzimaticamente utilizando o reagente ExoProStar (GE Healthcare, Little Chalfont, Reino Unido) e submetidos a reação de sequenciamento utilizando o kit BigDye Terminator v3.1 (Thermo Fisher). O sequenciamento automático por eletroforese capilar foi realizado no equipamento ABI 3130xl (Thermo Fisher) e os eletroferogramas analisados no software Sequencher v4.6.

3.3.6. Análise de variantes de número de cópia

Naqueles pacientes sem identificação de variantes de interesse na sequência nucleotídica dos genes candidatos, foi realizada análise comparativa de número de cópias com o software *Copy Number Analysis for Targeted Resequencing* (CONTRA) utilizando como controle (comparador) pacientes nos quais variantes patogênicas haviam sido identificadas. Possíveis CNVs identificadas pelo CONTRA foram avaliadas visualmente no IGV e, se mantida a suspeita, essas variantes foram submetidas a análise confirmatória por *array* citogenômico que foi realizado em colaboração com a Dra. Leslie Kulokowski, do Laboratório de Citogenética, LIM-03, FMUSP, utilizando-se CytoSNP-850K da Illumina (San Diego, California, EUA). Tal método contempla 843888 marcadores, sendo a distância entre sondas de 1800pb, resultando em uma resolução média de 18kb. Os dados foram extraídos pelo iScan System (Illumina) e as amostras controles anteriormente analisadas foram usadas como referência.

Dados obtidos foram inseridos no software BlueFuse Multi v1.1 (Blue Gnome) para normalização e cálculo de razões log2, equivalente ao quociente de intensidade normalizada da amostra pela intensidade média da amostra controle. Para avaliação da patogenicidade, as CNVs confirmadas foram pesquisadas na literatura em associação com OI.

3.3.7. Análise de segregação familiar

Familiares de primeiro grau dos pacientes com OI foram convidados a participar do estudo, para análise de segregação familiar das variantes identificadas. Os controles familiares foram questionados sobre o antecedente de fraturas de fragilidade, esclera azulada, dentinogênese imperfeita e cardiopatia conhecida. A análise molecular dos familiares foi realizada pelo método de Sanger, conforme descrito acima.

A avaliação da segregação foi individualizada para cada caso. Nos casos isolados com suspeita de defeito *de novo*, quando a variante encontrada no paciente esteve ausente em pai e mãe não afetados, consideramos que a segregação favorecia a patogenicidade da variante. Nos casos com suspeita de herança autossômica dominante, quando a variante encontrada no paciente esteve presente em pai ou mãe com fenótipo sugestivo da doença, também consideramos que a segregação favorecia patogenicidade. Nesta mesma situação, quando um dos pais foi inacessível à avaliação e a variante encontrada no paciente não esteve presente no familiar no qual a análise foi possível, a segregação foi considerada inconclusiva. Nos casos com suspeita de defeito recessivo, histórico de consanguinidade e achado de variante em homozigose, quando um dos pais foi inacessível à avaliação e a variante encontrada no paciente esteve presente em heterozigose no familiar analisado considerou-se que a segregação favorecia a patogenicidade da variante. Nos casos familiares, quando controles não afetados não eram portadores da variante encontrada nos pacientes, consideramos que a segregação favorecia a patogenicidade da variante. Quando familiares foram inacessíveis, a análise de segregação foi determinada indisponível.

3.4 Análise de associação entre achado molecular e características clínicas

Buscou-se associação entre os achados moleculares e as seguintes características clínicas: gravidade do acometimento esquelético, presença de esclera azulada atual ou pregressa, presença de dentinogênese imperfeita atual ou pregressa, perda auditiva detectada à audiometria, e alteração ecocardiográfica relevante (ou seja, excluindo-se alterações mínimas). A gravidade esquelética foi dividida em duas categorias, acometimento leve ou acometimento moderado a grave, seguindo a tendência da literatura (Bardai et al., 2016).

Para a análise de associação entre estas variáveis categóricas, realizou-se o teste qui-quadrado de Pearson utilizando o software PAWS Statistics v.17 (previamente conhecido como SPSS, IBM). Quando mais de 20% das células apresentavam valores menores do que o valor mínimo esperado, violando as suposições de independência e distribuição do qui-quadrado, utilizou-se para análise de associação o teste exato de Fisher nas tabelas 2x2, ou a razão de verossimilhança nas tabelas 3x2. A associação foi considerada estatisticamente significativa quando $p<0,05$.

4. Resultados

As características gerais da coorte estudada estão apresentadas na Tabela 2. As Tabelas 3 e 4 apresentam individualmente as caraterísticas clínicas principais dos casos isolados e familiares, respectivamente.

4.1. Coorte

Quarenta e nove pacientes em seguimento com diagnóstico clínico de OI foram incluídos neste estudo, correspondendo a 38 casos de OI: 30 casos isolados e 8 casos familiares (19 pacientes). Os heredogramas dos oito casos familiares são mostrados na Figura 2. As idades dos pacientes variaram de 7 a 69 anos sendo apenas 9 deles (18%) menores de 18 anos; a mediana de idade foi 24 anos. Vinte e nove indivíduos (59%) são do sexo feminino.

Classificando esses pacientes conforme a gravidade, 18 deles (37%) apresentam acometimento leve, 15 pacientes (30%) acometimento moderado e 16 pacientes (33%) apresentam acometimento grave pela doença. Houve maior proporção de OI leve dentre os casos familiares (74%) do que dentre os casos isolados (13%).

Trinta e oito pacientes (78%) apresentam baixa estatura e 36 pacientes (76%) apresentam alguma deformidade óssea. Do ponto de vista de mobilidade, 23 pacientes (47%) não apresentam qualquer prejuízo da mobilidade e deambulam sem necessidade de apoio, 6 pacientes (12%) se utilizam de algum apoio para deambulação, 19 pacientes (39%) são cadeirantes e 1 paciente (2%) é acamado. Dez pacientes (20%) apresentaram mais de 100 fraturas ao longo da vida e 11 pacientes (22%) apresentaram até 5 fraturas. Apenas 1 paciente nunca teve fraturas (Tabela 4, F3a).

Dez pacientes (20%) apresentaram DGI na vida adulta e 4 pacientes (8%) apresentaram DGI apenas na primeira dentição. Trinta e três pacientes (66%) apresentam

esclera azulada, sendo que dentre estes cinco são menores de 18 anos, e 6 pacientes (12%) apresentaram esclera azulada na infância.

Sete casos isolados e 3 famílias relataram antecedente de união consanguínea na família.

Tabela 2 - Características gerais da coorte

	Número de indivíduos (%)
Coorte	49 (100%)
Mulheres	29 (59%)
Homens	20 (41%)
Crianças (<12 anos)	2 (4%)
Adolescentes (12 – 18 anos)	7 (14%)
Adultos (>18 anos)	40 (82%)
Leve	18 (37%)
Moderado	15 (30%)
Grave	16 (33%)
Baixa estatura	38 (78%)
Presença de deformidade	36 (76%)
Sem limitação a deambulação	23 (47%)
Uso de apoio para deambulação	6 (12%)
Cadeirante	19 (39%)
Acamado	1 (2%)
Presença de dentinogênese imperfeita atual	10 (20%)
Presença de dentinogênese imperfeita na infância	4 (8%)
Presença de esclera azulada atual	33 (67%)
Presença de esclera azulada na infância	6 (12%)
Antecedente de consanguinidade	13 (26%)

Tabela 3 - Características clínicas dos casos isolados

Pt	Sexo	Idade	Total de fraturas	Altura (Z)	Gravidade	Consang.	Mobilidade	Esclera azulada	DGI
1	M	15	>150	-6,8	Grave	Não	Acamado	Infância	Não
2	F	23	>150	-6,6	Grave	Não	Cadeirante	Sim	Não
3	F	31	>150	ND	Grave	Sim	Cadeirante	Não	Não
4	M	20	>150	-6,4	Grave	Não	Cadeirante	Sim	1ª dentição
5	F	23	>150	-8,6	Grave	Sim	Cadeirante	Não	Não
6	M	23	110	-12,4	Grave	Não	Cadeirante	Infância	1ª dentição
7	M	19	106	-6,3	Grave	Não	Cadeirante	Sim	Não
8	F	32	100	-11,0	Grave	Sim	Cadeirante	Sim	1ª dentição
9	M	20	100	-11,3	Grave	Não	Cadeirante	Sim	1ª dentição
10	M	20	100	-9,8	Grave	Não	Cadeirante	Sim	Sim
11	F	21	50	-6,1	Mod.	Não	Cadeirante	Sim	Não
12	F	24	50	-0,9	Grave	Sim	Cadeirante	Infância	Não
13	F	31	50	-5,7	Mod.	Não	Com apoio	Sim	Sim
14	M	25	40	-6,1	Grave	Sim	Cadeirante	Sim	Não
15	F	27	40	-7,2	Grave	Sim	Sem apoio	Infância	Não
16	F	18	40	-1,9	Mod.	Sim	Sem apoio	Sim	Não
17	M	24	35	-6,4	Grave	Não	Cadeirante	Infância	Não
18	F	19	35	-9,6	Grave	Não	Cadeirante	Não	Não
19	M	22	30	-9,8	Mod.	Não	Cadeirante	Sim	Sim
20	M	18	24	-3,1	Mod.	Não	Com apoio	Não	Não
21	F	17	24	-5,4	Mod.	Não	Com apoio	Sim	Não
22	F	20	20	-5,0	Mod.	Não	Cadeirante	Sim	Não
23	F	46	14	-4,4	Grave	Não	Com apoio	Sim	Não
24	F	59	10	-3,1	Mod.	Não	Com apoio	Sim	Não
25	M	14	9	-2,9	Mod.	Não	Sem apoio	Sim	Não
26	M	24	8	-0,7	Leve	Não	Sem apoio	Sim	Não
27	F	7	8	-0,3	Leve	Não	Sem apoio	Sim	Não
28	M	22	7	-2,1	Leve	Não	Sem apoio	Sim	Sim
29	F	16	5	-9,0	Mod.	Não	Cadeirante	Infância	Sim
30	F	53	50	-2,8	Leve	Não	Sem apoio	Sim	Não

Pt, identificador do paciente; M, masculino; F, feminino; Z, desvio-padrão da altura para mesma idade/sexo; Consang., antecedente de consanguinidade; DGI, dentinogênese imperfeita; Mod., Moderado; Com apoio/Sem apoio, deambula com/sem apoio; ND, não disponível

Tabela 4 - Características clínicas dos casos familiares

Pt	Sexo	Idade	Total de fraturas	Altura (Z)	Gravidade	Consang.	Mobilidade	Esclera azulada	DGI
F1a	F	8	4	0,6	Leve	Não	Sem apoio	Sim	Não
F1b	M	37	10	-1,2	Leve	Não	Sem apoio	Sim	Não
F2a	F	48	4	-3,0	Leve	Não	Sem apoio	Não	Sim
F2b	F	42	9	-7,9	Mod.	Não	Sem apoio	Não	Sim
F3a	F	19	0	-1,9	Leve	Sim	Sem apoio	Não	Não
F3b	M	59	2	-2,9	Leve	Sim	Sem apoio	Não	Não
F3c	F	69	2	-3,7	Leve	Sim	Sem apoio	Não	Não
F3d	F	48	1	-1,6	Leve	Sim	Sem apoio	Não	Não
F4a	F	27	30	-4,8	Mod.	Não	Cadeirante	Sim	Sim
F4b	F	54	5	-3,1	Mod.	Não	Cadeirante	Sim	Não
F5a	F	16	2	-2,3	Leve	Não	Sem apoio	Sim	Sim
F5b	M	50	12	-4,6	Leve	Não	Com apoio	Sim	Sim
F6a	M	20	20	-2,3	Mod.	Sim	Sem apoio	Sim	Não
F6b	F	48	50	-4,1	Mod.	Sim	Sem apoio	Sim	Não
F7a	M	20	14	-1,0	Leve	Sim	Sem apoio	Sim	Não
F7b	F	25	15	-1,3	Leve	Sim	Sem apoio	Sim	Não
F8a	M	32	5	-3,7	Leve	Não	Sem apoio	Sim	Não
F8b	M	29	5	-2,9	Leve	Não	Sem apoio	Sim	Não
F8c	F	54	3	-2,9	Leve	Não	Sem apoio	Sim	Não

Pt, identificador do paciente; M, masculino; F, feminino; Z, desvio-padrão da altura para mesma idade/sexo; Consang., antecedente de consanguinidade; DGI, dentinogênese imperfeita; Mod., Moderado; Com apoio/Sem apoio, deambula com/sem apoio; ND, não disponível

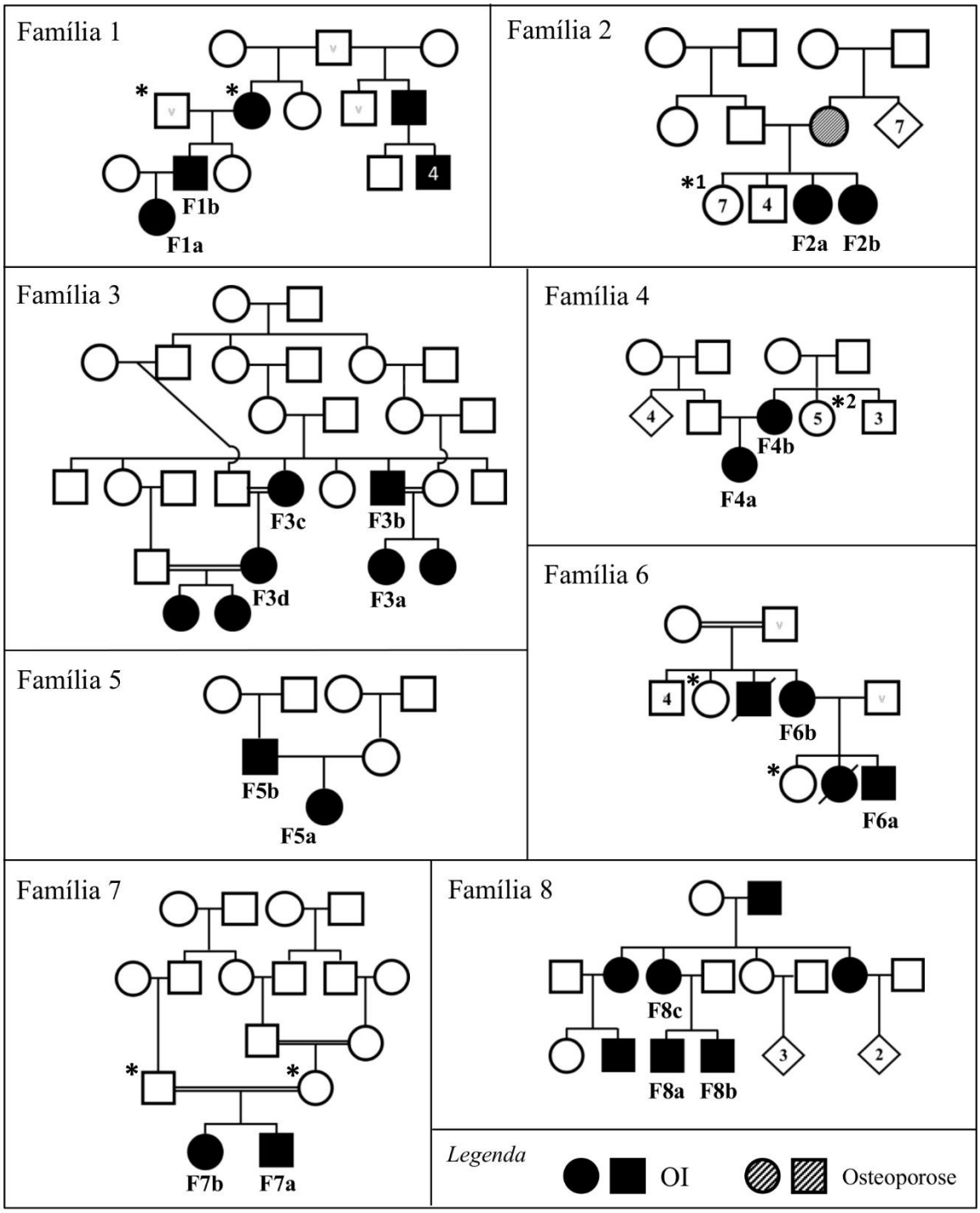

Figura 2 - Heredogramas dos casos familiares
Os pacientes analisados através do painel de SPLE dentro de cada família estão identificados; outros indivíduos dos quais foram obtidas amostras de DNA para análise de segregação familiar estão destacados pelo *.

Dados adicionais do seguimento clínico da coorte são mostrados na Tabela 5. Noventa e quatro por cento dos pacientes tiveram a primeira fratura durante a infância, e as deformidades ósseas mais comuns foram em membros (67%) e coluna (43%). A maioria dos pacientes apresentava baixa massa óssea (61%, Z-escore < -2,0) entretanto em muitos indivíduos a análise da densidade mineral óssea era falseada por material de síntese metálico ou deformidades. A avaliação audiométrica foi obtida em 35 pacientes (71%), e resultou normal na maioria (83%). A avaliação cardíaca por ecocardiograma foi realizada em 40 pacientes (82%), identificando alterações relevantes em 9 pacientes (23%). Dados retrospectivos referentes ao tratamento ortopédico mostram que em 35% dos pacientes as fraturas foram tratadas apenas de modo conservador, e que 55% tinham sido submetidos a fixação metálica. Dados retrospectivos do tratamento medicamentoso prévio com bisfosfonatos estão apresentados no Apêndice 1.

Tabela 5 - Dados clínicos adicionais dos pacientes estudados

Pt	Idade 1ª fratura	Deformidades	Z-escore corpo total	Z-escore L1-L4	Audiometria	Ecocardiograma	Tratamento ortopédico prévio
1	5 m	Membros, coluna e tórax	-1,4	-2,0*	ND	ND	Conservador
2	6 m	MMII	-0,2*	-0,7	ND	ND	Fixação metálica MMII
3	15 d	Membros, coluna e tórax	ND	ND	Normal	Normal	Fixação metálica MIE, redução cirúrgica MMSS
4	IU	Membros, coluna e tórax	0,9*	-3,1	Normal	Dilatação discreta de VD	Fixação metálica MMII e MSD
5	1 a	Membros, coluna	-1,0	-2,9*	ND	ND	Conservador
6	RN	Membros, coluna e tórax	-2,8	ND	ND	Normal	Redução cirúrgica MMII
7	7 m	Membros e coluna	4,4*	-1,5*	Normal	Normal	Fixação metálica MMII
8	RN	Coluna	-3,8	-5,6*	Normal	ND	Conservador
9	RN	Membros e coluna	-2,3	-5,4*	Normal	Normal	Conservador
10	RN	Membros e coluna	-0,3	-2,0*	Normal	Normal	Conservador
11	RN	MMII e coluna	0,6*	-3,8*	Normal	Normal	Fixação metálica MMII
12	6 m	Membros e coluna	-0,5	-1,3*	Normal	Refluxo mitral e tricúspide mínimos	Fixação metálica MMII, artrodese de coluna
13	6 a	MMII	ND	-3,8	ND	Refluxo mitral e tricúspide mínimos	Fixação metálica MMII
14	RN	MMII, MSD e coluna	0,1*	-3,5	ND	Normal	Fixação metálica MMII
15	7 a	Membros	-2,7*	-3,0	Normal	Refluxo mitral mínimo	Fixação metálica MMII
16	RN	MID	3,1*	2,5	Normal	Insuficiência mitral de moderada a grave, aumento do átrio E	Fixação metálica MID
17	RN	MIE e coluna	-3,8*	-4,2*	Perda auditiva à E	Normal	Fixação metálica MID e MSD
18	RN	Coluna	1,4*	-5,3*	Normal	Normal	Fixação metálica MMII e MSD
19	RN	Membros, coluna e tórax	-3,1*	-6,3*	Normal	Normal	Fixação metálica MMII, artrodese de coluna

continua

Tabela 5 - Dados clínicos adicionais dos pacientes estudados *(continuação)*

Pt	Idade 1ª fratura	Deformidades	Z-escore corpo total	Z-escore L1-L4	Audiometria	Ecocardiograma	Tratamento ortopédico prévio
20	3 a	MMII	0,5*	0,5	Normal	Normal	Fixação metálica MID
21	RN	MSE e coluna	6,5*	-2,8	Normal	Normal	Fixação metálica MMII
22	2 a	MIE e coluna	-1,6	-3,6*	Normal	Refluxo mitral, tricúspide e aórtico mínimos	Fixação metálica MIE
23	2 a	MMSS, MID e coluna	-0,9*	-3,8*	ND	Discreto refluxo mitral e tricúspide; indícios de forame oval patente	Fixação metálica MIE
24	5 a	MID e coluna	ND	-1,5	Perda auditiva à E	Hipertrofia concêntrica de VE, discreto aumento de AE, discreto refluxo mitral, aórtico e tricúspide	Conservador
25	6 m	*Ausentes*	-1,9*	-3,3	Normal	Normal	Fixação metálica MMII
26	1 a	MSD	ND	-2,2	ND	Normal	Fixação metálica MSD
27	7 m	*Ausentes*	ND	-1,5	ND	ND	Redução cirúrgica MSD
28	2 a	*Ausentes*	0,1	-1,3	Normal	Normal	Fixação metálica MID
29	RN	MMII e coluna	-3,7*	-5*	Normal	Normal	Fixação metálica MMII
30	4 a	MSD	ND	-2,6	Perda auditiva à D moderada a grave	Hipertrofia concêntrica de VE, disf. diastólica moderada, discreto aumento do AE e refluxo aórtico	Conservador
F1a	11 m	*Ausentes*	0,1	-0,5	ND	ND	Conservador
F1b	4 a	Clavícula E e MSD	ND	-3,1	Normal	Aumento da espessura da parede miocárdica	Fixação metálica MSE
F2a	3 a	MSE	ND	-1,9	Normal	Normal	Fixação metálica MIE
F2b	4 a	MMII	-1,5*	-2,9	ND	Normal	Conservador
F3a	-	*Ausentes*	-4,1	-4,1	Normal	Normal	Não

continua

Tabela 5 - Dados clínicos adicionais dos pacientes estudados *(conclusão)*

Pt	Idade 1ª fratura	Deformidades	Z-escore corpo total	Z-escore L1-L4	Audiometria	Ecocardiograma	Tratamento ortopédico prévio
F3b	25 a	*Ausentes*	ND	-1,5	Perda auditiva bilateral	Normal	Conservador
F3c	2 a	Coluna	ND	-2,5	Perda auditiva leve para agudos	Dilatação discreta de AE e disfunção diastólica discreta	Conservador
F3d	48 a	*Ausentes*	ND	-3,7	Normal	Normal	Conservador
F4a	RN	MMII	-0,4*	-1,8	Perda auditiva à E	ND	Fixação metálica MMII
F4b	5 a	MMII	ND	-0,9	Normal	Disfunção diastólica discreta	Redução cirúrgica MMII
F5a	IU	*Ausentes*	1,4	2,3	Normal	Normal	Fixação metálica MID
F5b	RN	MID	-2,5	-2,6	Normal	Normal	Conservador
F6a	2 a	MSE	1,6	1,8	Normal	Normal	Fixação metálica MSD
F6b	4 m	MMII	ND	-3,2	Normal	Normal	Redução cirúrgica MID
F7a	RN	*Ausentes*	-1,5	-2,0	Normal	Normal	Conservador
F7b	5 a	*Ausentes*	ND	-1,0	Normal	Refluxo mínimo de tricúspide	Conservador
F8a	3 m	*Ausentes*	ND	-3,4	ND	ND	Conservador
F8b	8 a	*Ausentes*	ND	-3,8	ND	ND	Fixação metálica MSD
F8c	10 a	*Ausentes*	ND	-1,3	ND	Aumento atrial moderado, hipertrofia excêntrica VE, disfunção diastólica moderada; refluxo valvar discreto	Conservador

Pt, identificador do paciente; IU, intra-útero; RN, recém-nascido; d, dia; m, mês; a, ano; MMII, membros inferiores; MIE, membro inferior esquerdo, MID, membro inferior direito; MMSS, membros superiores; MSE, membro superior esquerdo; MSD, membro superior direito; ND, não disponível; E, esquerda; D, direita; AE, átrio esquerdo; VE, ventrículo esquerdo; VD, ventrículo direito;
*exame com artefato (presença de material de síntese metálico ou deformidade)

4.2. Resultados do SPLE

O SPLE do painel de genes candidatos foi realizado em todos os 49 pacientes da coorte. A cobertura média do SPLE variou entre 354 a 1382 leituras. Em todos os pacientes, acima de 99% das regiões-alvo tiveram cobertura superior a 50 vezes.

Ao todo, foram encontradas 42 variantes de interesse (Tabela 6; descrição completa da variante ao nível do DNA e da proteína no Apêndice 2). Destas, 23 já foram descritas anteriormente em associação a OI e 19 são variantes novas. Dentre as variantes de ponto e pequenas deleções/inserções, 40 variantes foram confirmadas por sequenciamento Sanger; apesar de inúmeras tentativas, não se obteve amplificação por PCR da região envolvendo a variante PLOD2 p.(Trp561*), que foi detectada em 386 de 853 leituras no caso Pt 6. A análise de segregação familiar foi realizada através de sequenciamento Sanger e seus resultados são mostrados na Tabela 7.

Aplicando-se as diretrizes de interpretação propostas pelo ACMG-AMP (Richards et al., 2015), 24 variantes foram classificadas como patogênicas (VP), 7 como provavelmente patogênicas (VPP) e 11 como variante de significado incerto (VSI). Nove VSIs foram identificadas em combinação com outras VSIs ou VPPs.

Em trinta e sete casos (97%) foi identificado o diagnóstico molecular. Em 18 casos foram encontradas variantes no gene *COL1A1* e em 9 casos, variantes em *COL1A2*, ou seja, 71% dos casos apresenta defeito molecular nos genes do colágeno tipo 1. Em três destes casos, foi encontrada uma combinação de variantes: 2 casos com 2 variantes em *COL1A1* (Pt 9 e F6, Tabela 6), e 1 caso com 2 variantes em *COL1A2* (F4, Tabela 6).

Em 9 casos foram encontradas variantes nos genes candidatos *IFITM5* (n=1), *P3H1* (n=1), *SERPINF1* (n=2), *FKBP10* (n=2), *PLOD2* (n=2) e *WNT1* + *P3H1* (n=1). Em um caso foi identificada deleção de 2 éxons do gene *TMEM38B*.

Não foram identificadas variantes de interesse no caso isolado Pt 15.

Tabela 6 - Variantes genéticas identificadas na coorte

Pt	Variante genética	Frequência populacional			Predição *in silico*			Segregação familiar	Descrição prévia	ACMG -AMP
		1000g	ExAC	ABraOM	SIFT	PP2	CADD			
1	SERPINF1 p.(Phe384Leufs*9)(;)(Phe384Leufs*9)	0	0	0	ND	ND	ND	Inconclusiva	Sim	VP
2	COL1A1 p.(Gln1280Pro)	0	0	0	D	D	23,7	Inconclusiva	Não	VPP
3	FKBP10 p.(Gly278Argfs*95)(;)(Gly278Argfs*95)	0	0,0001	0	ND	ND	ND	Segrega	Sim	VP
4	COL1A1 p.(Gly857Cys)	0	0	0	D	D	27	Inconclusiva	Sim	VP
5	*SERPINF1* c.[283+2T>C];[(283+2T>C)]	0	0	0	ND	ND	25,1	Segrega	Não	VP
6	PLOD2 p.(Trp561*)	0	0	0	T	ND	40	Indisponível	Não	VP
	PLOD2 p.(Glu499Aspfs*29)	0	0	0	ND	ND	ND	Inconclusiva	Não	VP
7	*COL1A1* c.643-2A>G	0	0	0	ND	ND	24,8	Segrega	Sim	VP
8	*P3H1* c.[1080+1G>T];[1080+1G>T]	0,0006	0,0002	0	ND	ND	25,9	Segrega	Sim	VP
9	COL1A1 p.(Gly293Asp)	0	0	0	D	D	26,5	Segrega	Sim	VP
	COL1A1 p.(Ser291Arg)	0	0	0	T	B	22,7	Segrega	Não	VSI
10	COL1A2 p.(Gly367Glu)	0	0	0	D	D	18,6	Segrega	Sim	VP
11	COL1A2 p.(Gly328Ser)	0	0	0	D	D	29,7	Inconclusiva	Sim	VP
12	FKBP10 p.(Gln60Pro)(;)(Gln60Pro)	0	0	0	T	D	33	Segrega	Não	VSI
13	COL1A2 p.(Gly772Ser)	0	0	0	D	PD	34	Indisponível	Sim	VPP
14	COL1A1 p.(Gly719Cys)	0	0	0	D	D	35	Inconclusiva	Não	VSI
16	*TMEM38B* deleção de éxons 1 e 2	ND	ND	ND	ND	ND	ND	Indisponível	Sim	VP

continua

Tabela 6 - Variantes genéticas identificadas na coorte (*continuação*)

Pt	Variante genética	Frequência populacional			Predição *in silico*			Segregação familiar	Descrição prévia	ACMG -AMP
		1000g	ExAC	ABraOM	SIFT	PP2	CADD			
17	P3H1 p.(Ala117Cysfs*64)	0	0	0	ND	ND	ND	Segrega	Não	VPP
	P3H1 p.(Cys483Tyr)	0	0	0	D	D	17	Segrega	Não	VSI
	WNT1 p.(Asn103Thr)	0	0	0	D	D	23,9	Segrega	Não	VSI
	WNT1 p.(Gly259_Gly262dup)	0	0	0,000821	ND	ND	ND	Segrega	Não	VSI
18	COL1A1 p.(Gly1040Ser)	0	0	0	D	D	21,5	Inconclusiva	Sim	VP
19	COL1A1 p.(Gly767Ser)	0	0	0	D	D	23,8	Segrega	Sim	VP
20	PLOD2 p.(Tyr547His)	0,0016	0,0006	0,003284	T	PD	17,1	Indisponível	Não	VSI
	PLOD2 p.(Cys282Arg)	0,0006	0,0002	0	D	D	21,6	Indisponível	Não	VSI
21	COL1A2 p.(Gly193Asp)	0	0	0	D	D	25,6	Segrega	Não	VPP
22	COL1A1 p.(Gln250*)	0	0	0	ND	ND	37	Inconclusiva	Não	VP
23	*IFITM5* c.-14C>T	0	0	0	ND	ND	ND	Indisponível	Sim	VP
24	COL1A1 p.(Pro871Leufs*237)	0	0	0	ND	ND	ND	Indisponível	Sim	VP
25	COL1A2 p.(Gly772Ser)	0	0	0	D	PD	34	Inconclusiva	Sim	VPP
26	COL1A1 p.(Glu106*)	0	0	0	T	ND	32	Inconclusiva	Não	VP
27	*COL1A1* c.2452-2A>G	0	0	0	ND	ND	23,8	Segrega	Sim	VP
28	COL1A1 p.(Gly788Ser)	0	0	0	D	D	28,2	Segrega	Sim	VP
29	COL1A2 p.(Gly460Ser)	0	0	0	D	D	28,7	Segrega	Sim	VP

continua

Tabela 6 - Variantes genéticas identificadas na coorte (*conclusão*)

Pt	Variante genética	Frequência populacional			Predição *in silico*			Segregação familiar	Descrição prévia	ACMG -AMP
		1000g	ExAC	ABraOM	SIFT	PP2	CADD			
30	*COL1A1* c.334-9A>G	0	0	0	ND	ND	ND	Segrega	Sim	VPP
F1	*COL1A1* c.334-9A>G	0	0	0	ND	ND	ND	Segrega	Sim	VPP
F2	COL1A2 p.(Gly319Arg)	0	0	0	D	D	22,8	Segrega	Sim	VPP
F3	COL1A2 p.(Gly772Ser)	0	0	0	D	PD	34	Indisponível	Sim	VPP
F4	COL1A2 p.(Gly229Asp)	0	0	0	D	D	19,7	Segrega	Não	VSI
	COL1A2 p.(Pro230Gln)	0	0		D	D	17,7	Segrega	Não	VSI
F5	COL1A1 p.(Gly527Ala)	0	0	0	D	D	20,7	Indisponível	Sim	VPP
F6	COL1A1 p.(Arg1141*)	0	0	0	ND	ND	42	Segrega	Sim	VP
	COL1A1 p.(Arg528His)	0,0001	0,0001	0,000821	D	D	17,7	Segrega	Não	VSI
F7	COL1A1 p.(Ser271Glnfs*16)	0	0	0	ND	ND	ND	Segrega	Sim	VP
F8	COL1A1 p.(Gly257Arg)	0	0	0	D	D	28,6	Indisponível	Sim	VP

Pt, identificador; ND, não disponível; D, deletéria; T, tolerada; PD, possivelmente deletéria; B, benigna; VP, variante patogênica; VPP, variante provavelmente patogênica; VSI, variante de significado incerto.
Bases de dados de frequência populacional:1000g, 1000 Genomes; ExAC, Exome Aggregation Consortium; ABraOM, Arquivo Brasileiro Online de Mutações.
Ferramentas de predição *in silico*: SIFT, Sorting Intolerant from Tolerant; PP2, PolyPhen2; CADD, Combined Annotation Dependent Depletion (pontuação escalonada, "phred-like"; pontuação ≥15 sugere patogenicidade).
ACMG-AMP: Interpretação do potencial patogênico de variantes de acordo com o Colégio Americano de Genética e Genômica Médica e a Associação de Patologia Molecular (Richards et al., 2015).

Tabela 7 - Análise de segregação familiar

Pt	Variante genética	Análise de segregação familiar	Conclusão
1	SERPINF1 p.(Phe384Leufs*9)(;)(Phe384Leufs*9)	Pai não afetado heterozigoto, mãe inacessível	Inconclusiva
2	COL1A1 p.(Gln1280Pro)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
3	FKBP10 p.(Gly278Argfs*95)(;)(Gly278Argfs*95)	Pai inacessível, mãe não afetada heterozigota; pais primos de 1º grau	Segrega
4	COL1A1 p.(Gly857Cys)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
5	*SERPINF1* c.[283+2T>C];[(283+2T>C)]	Pai inacessível, mãe não afetada heterozigota, irmã não afetada heterozigota; pais primos de 2º grau	Segrega
6	PLOD2 p.(Trp561*)	*Não foi possível realizar o sequenciamento Sanger*	Indisponível
	PLOD2 p.(Glu499Aspfs*29)	Pai não afetado heterozigoto, mãe não afetada selvagem; pais não consanguíneos	Inconclusiva
7	*COL1A1* c.643-2A>G	Pais não afetados selvagens	Segrega
8	*P3H1* c.[1080+1G>T];[1080+1G>T]	Pais não afetados heterozigotos; primos de 1º grau	Segrega
9	COL1A1 p.(Gly293Asp)	Pais não afetados selvagens, irmão não afetado selvagem	Segrega
	COL1A1 p.(Ser291Arg)	Pais não afetados selvagens, irmão não afetado selvagem	Segrega
10	COL1A2 p.(Gly367Glu)	Pais não afetados selvagens	Segrega
11	COL1A2 p.(Gly328Ser)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
12	FKBP10 p.(Gln60Pro)(;)(Gln60Pro)	Pai inacessível, mãe não afetada heterozigota; pais primos de 1º grau	Segrega
13	COL1A2 p.(Gly772Ser)	Familiares inacessíveis	Indisponível

continua

Tabela 7 – Análise de segregação familiar (*continuação*)

Pt	Variante genética	Análise de segregação familiar	Conclusão
14	COL1A1 p.(Gly719Cys)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
16	*TMEM38B* deleção de éxons 1 e 2	*Não foi possível realizar SPLE ou SNP array nas amostras familiares*	Indisponível
17	P3H1 p.(Ala117Cysfs*64)	Pai não afetado heterozigoto, mãe não afetada selvagem	Segrega
	P3H1 p.(Cys483Tyr)	Pai não afetado selvagem, mãe não afetada heterozigota	
	WNT1 p.(Asn103Thr)	Pai não afetado heterozigoto, mãe não afetada selvagem	Segrega
	WNT1 p.(Gly259_Gly262dup)	Pai não afetado selvagem, mãe não afetada heterozigota	
18	COL1A1 p.(Gly1040Ser)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
19	COL1A1 p.(Gly767Ser)	Pais não afetados selvagens	Segrega
20	PLOD2 p.(Tyr547His)	Familiares inacessíveis	Indisponível
	PLOD2 p.(Cys282Arg)		
21	COL1A2 p.(Gly193Asp)	Pais não afetados selvagens	Segrega
22	COL1A1 p.(Gln250*)	Pai não afetado selvagem, mãe inacessível	Inconclusiva
23	*IFITM5* c. -14 C->T	Familiares inacessíveis	Indisponível
24	COL1A1 p.(Pro871Leufs*237)	Familiares inacessíveis	Indisponível
25	COL1A2 p.(Gly772Ser)	Pai inacessível, mãe heterozigota com antecedente de 1 fratura traumática	Inconclusiva
26	COL1A1 p.(Glu106*)	Pai inacessível, mãe não afetada selvagem	Inconclusiva
27	*COL1A1* c.2452-2A>G	Pai inacessível, mãe heterozigota tem esclera azulada e 2 fraturas vertebrais	Segrega

continua

Tabela 7 – Análise de segregação familiar (*conclusão*)

Pt	Variante genética	Análise de segregação familiar	Conclusão
28	COL1A1 p.(Gly788Ser)	Pais não afetados selvagens	Segrega
29	COL1A2 p.(Gly460Ser)	Pais não afetados selvagens	Segrega
30	*COL1A1* c.334-9A>G	Pais inacessíveis, 3 irmãos não afetados selvagens	Segrega
F1	*COL1A1* c.334-9A>G	3 pacientes, avó, pai e filha, heterozigotos; avô não afetado selvagem	Segrega
F2	COL1A2 p.(Gly319Arg)	Pais inacessíveis, 2 pacientes irmãs heterozigotas, 1 irmã não afetada selvagem	Segrega
F3	COL1A2 p.(Gly772Ser)	Familiares inacessíveis	Indisponível
F4	COL1A2 p.(Gly229Asp)	2 pacientes, mãe e filha, heterozigotas, 2 irmãs da mãe não afetadas selvagens	Segrega
	COL1A2 p.(Pro230Gln)	2 pacientes, mãe e filha, heterozigotas, 2 irmãs da mãe não afetadas selvagens	Segrega
F5	COL1A1 p.(Gly527Ala)	Familiares inacessíveis	Indisponível
F6	COL1A1 p.(Arg1141*)	2 pacientes, mãe e filho, heterozigotos, 1 irmã da mãe e 1 irmã do filho não afetadas e selvagens	Segrega
	COL1A1 p.(Arg528His)	2 pacientes, mãe e filho, heterozigotos, 1 irmã da mãe e 1 irmã do filho não afetadas e selvagens	Segrega
F7	COL1A1 p.(Ser271Glnfs*16)	2 pacientes irmãos heterozigotos, pai heterozigoto com baixa estatura e valvopatia, mãe não afetada selvagem	Segrega
F8	COL1A1 p.(Gly257Arg)	Familiares inacessíveis	Indisponível

4.2.1. Variantes encontradas no gene *COL1A1*

Dezenove variantes heterozigóticas foram encontradas no gene *COL1A1* em 13 casos isolados e 5 famílias. Em um caso isolado (Pt 9) e em uma família (F6) foi encontrada combinação de 2 variantes em *COL1A1*. A variante de *splicing COL1A1* c.334-9A>G, já descrita na literatura, foi encontrada em um caso isolado (Pt 30) e uma família (F1), aparentemente não relacionados. Destas 19 variantes em *COL1A1*, 13 já foram descritas anteriormente na literatura e 6 são variantes novas.

Oito variantes (42%) encontradas envolvem o resíduo glicina da cadeia α1 do colágeno tipo 1, cuja presença a cada três aminoácidos é importante para a formação da estrutura em tripla-hélice do colágeno. Sete destas variantes já foram descritas na literatura associadas a OI.

As variantes encontradas estão dispersas ao longo da proteína (ausência de *hot spot*) e a variabilidade clínica dentre os pacientes é grande, com 13 pacientes apresentando acometimento leve, 5 apresentando acometimento moderado e 6 com acometimento grave (Figura 3). Todos os indivíduos portadores da variante c.334-9A>G tiveram acometimento leve, e dentro das famílias com defeito em *COL1A1*, a gravidade da doença foi semelhante entre os portadores.

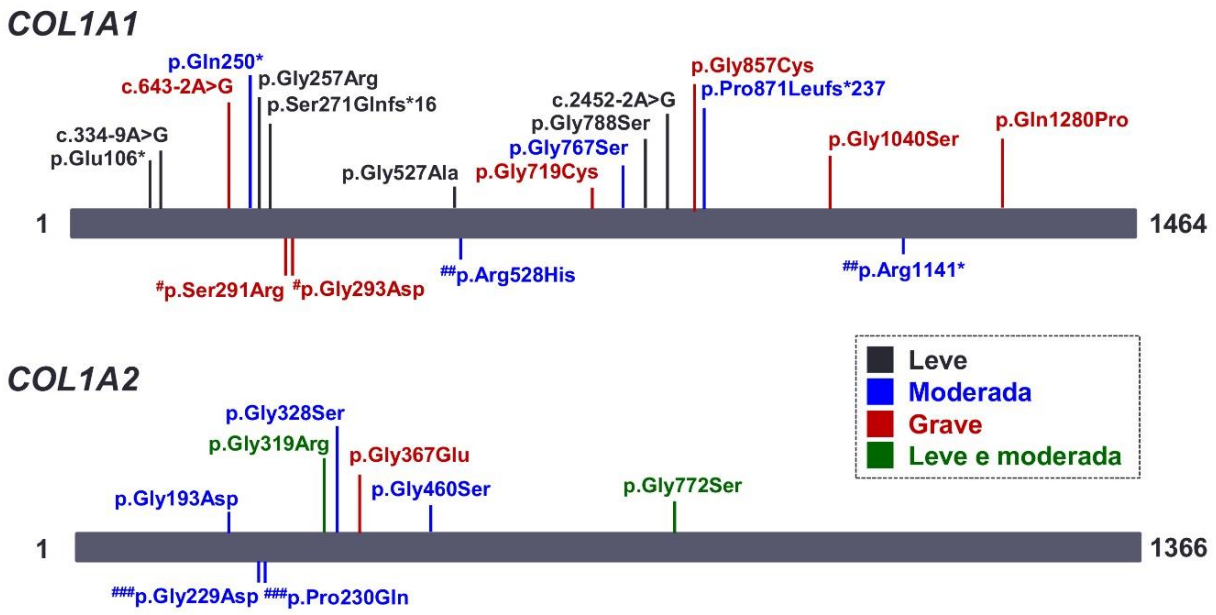

Figura 3 - Representação das variantes identificadas em COL1A1 e COL1A2
As proteínas COL1A1 e COL1A2 são representadas pelas barras cinzas, e os números nas extremidades denotam os resíduos iniciais e finais, respectivamente. As variantes são mostradas de acordo com sua posição relativa na proteína; a cor denota a gravidade de OI com a qual a variante foi associada na coorte, de acordo com a legenda. Acima das barras constam as variantes identificadas isoladamente, e abaixo das barras aquelas identificadas em combinação em um mesmo caso; o marcador # antecedendo o nome da variante denota os pares destas combinações.

4.2.2. Variantes encontradas no gene *COL1A2*

Oito variantes heterozigóticas foram encontradas no gene *COL1A2* em 6 casos isolados e em 3 famílias, sendo que 6 destas variantes já foram descritas anteriormente e 2 são novas. Em uma família foi encontrada a combinação de 2 variantes em *COL1A2* (F4). Sete variantes (87,5%) envolvem o resíduo glicina da cadeia α2 do colágeno tipo 1; destas, 6 já foram descritas na literatura associadas a OI.

As variantes em COL1A2 encontradas estão concentradas no primeiro terço da proteína, e a variabilidade clínica dentre os pacientes também é grande, com 5 pacientes apresentando acometimento leve, 8 pacientes com acometimento moderado e 1 paciente com acometimento grave (Figura 3).

A variante COL1A2 p.(Gly772Ser) foi encontrada na família 3 e nos casos isolados Pt 13 e Pt 25 (Tabela 6). Aparentemente estes 3 casos não são relacionados: a família é procedente de Senhor do Bonfim/BA, um dos casos é procedente de Francisco Morato/SP, e o outro de Ipiaú/BA (a distância entre as duas cidades baianas é de cerca de 540 km). De acordo com informação da base de dados *Osteogenesis Imperfecta Variant Database* (https://oi.gene.le.ac.uk), esta variante já foi previamente reportada em 17 pacientes com OI, em diversos países (Balasubramanian et al., 2015; Lu et al., 2014; Nuytinck et al., 1997). O quadro clínico dos indivíduos portadores da variante nesta coorte é bastante diverso (Tabela 8). Não foram encontradas outras variantes que pudessem modificar a expressão fenotípica nesses pacientes.

Tabela 8 - Características clínicas dos portadores de COL1A2 p.(Gly772Ser)

	Pt 13	**Pt 25**	**F3a**	**F3b**	**F3c**	**F3d**
Sexo	F	M	F	M	F	F
Idade (anos)	31	14	69	59	48	19
Total de fraturas	50	9	2	2	1	0
Altura (Z)	-5,7	-2,9	-3,7	-2,9	-1,6	-1,9
Deformidades	MMII	ausentes	escoliose	ausentes	ausentes	ausentes
Mobilidade	Com apoio	Sem apoio	Sem apoio	Sem apoio	Sem apoio	Sem apoio
Esclera azulada	Sim	Sim	Não	Não	Não	Não
DGI	Sim	Não	Não	Não	Não	Não

F, feminino; M, masculino; MMII, membros inferiores

4.2.3. Variantes encontradas nos demais genes candidatos

Em dez casos isolados foram encontradas variantes de interesse nos demais genes candidatos. Nos casos familiares não foram encontradas variantes nestes genes. Dentre os pacientes com variantes nos genes não colágenos prevaleceu o fenótipo grave da doença, manifesto em 8 pacientes deste grupo (Tabela 4). Dois pacientes apresentam a forma moderada e nenhum paciente apresenta forma leve da doença.

Ao todo, foram encontradas 15 variantes nos demais genes candidatos, sendo 5 previamente descritas em associação a OI e 10 novas. Nenhuma variante nestes genes foi encontrada em mais de um caso.

Em cinco casos foram encontradas variantes de ponto em homozigose, em genes associados a OI recessiva: *SERPINF1*, *FKBP10* e *P3H1*. Destes 5 casos, apenas o Pt 1 com a variante SERPINF1 p.(Phe384Leufs*9)(;)(Phe384Leufs*9) não tinha histórico de consanguinidade, porém os pais eram naturais de uma mesma cidade no interior de Minas Gerais (Bueno Brandão, população estimada em 11.000 habitantes em 2017). Ao se

realizar a busca desta variante na literatura, encontrou-se descrição prévia deste mesmo caso, resultante de sequenciamento exômico realizado em outra instituição, e de outros casos de Bueno Brandão (Minillo et al., 2014).

Em 2 casos sem histórico de consanguinidade (Pt 6 e Pt 20) foram encontradas combinações de duas variantes heterozigóticas diferentes em *PLOD2*, associado a OI recessiva, configurando, assim, herança por heterozigose composta.

Em um paciente com OI grave, fruto de união não consanguínea, foi encontrada a combinação de quatro variantes heterozigóticas diferentes em dois genes candidatos: duas em *P3H1* e duas em *WNT1*, ambos genes associados a OI recessiva (Pt17, Tabela 6). Tais variantes são extremamente raras ou ausentes na população, com impacto esperado ou predito na proteína (Tabela 6, Figura 4). A análise de segregação familiar revelou que, em cada um destes genes, cada variante foi herdada de um dos pais, que têm massa óssea normal, nunca tiveram fraturas nem têm deformidades ósseas, configurando dupla heterozigose composta, e um potencial defeito digênico (Figura 4).

		Frequência populacional	Predição *in silico*			Pt 17	Mãe	Pai
			SIFT	PP2	CADD			
P3H1	p.(Cys483Tyr)	0,0008%	D	D	17,0	HTZ	HTZ	WT
	p.(Ala117Cysfs*64)	0	ND	ND	ND	HTZ	WT	HTZ
WNT1	p.(Asn103Thr)	0	D	D	23,9	HTZ	WT	HTZ
	p. (Gly259_Gly262dup)	0	ND	ND	ND	HTZ	HTZ	WT

Figura 4 - Potencial defeito digênico envolvendo *P3H1* e *WNT1*
A análise por imagem do paciente 17 através do esquadrinhamento de corpo inteiro na densitometria óssea (painel superior esquerdo), radiografia anteroposterior de coluna vertebral (painel superior central) e escanometria de membros inferiores (painel superior direito) revela acometimento ósseo grave com múltiplas deformidades e escoliose. O quadro mostra as variantes heterozigóticas identificadas em *P3H1* e *WNT1*, com padrão de herança em dupla heterozigose composta. Estas variantes são extremamente raras ou ausentes na população, com impacto na proteína esperado (*frameshift* resultando em ganho de códon de parada em *P3H1*; inserção de 4 aminoácidos em WNT1) ou predito pelas ferramentas SIFT, PolyPhen2 (PP2) e CADD. D, deletéria; ND, não disponível; HTZ, heretozigoto; WT, selvagem.

Em apenas uma paciente da coorte (Pt 23, Tabela 6) foi identificada a variante heterozigótica na região 5' não traduzida de *IFITM5* associada a OI dominante com peculiaridades radiológicas. Após o diagnóstico molecular, a reanálise de exames de imagem da paciente revelou presença de calo hipertrófico em membro inferior direito e calcificação das membranas interósseas nos antebraços, com luxações das cabeças dos rádios (Figura 5).

Figura 5 - Estudo por imagem da portadora da variante *IFITM5* c.-14C>T
À esquerda, esquadrinhamento de corpo inteiro na densitometria óssea da paciente 23 revelou escoliose grave, deformidades em membros superiores e inferiores, e imagem compatível com calo ósseo hiperplásico no fêmur esquerdo. Ao meio, a radiografia de fêmur confirmou o achado de calo ósseo hiperplásico reorganizado, e à direita, radiografias de antebraços mostraram calcificação de membranas interósseas e luxação das cabeças dos rádios. Estes achados radiológicos são típicos da forma de OI associada à variante c.-14C>T que resulta em novo códon de iniciação de tradução, e, portanto, em uma proteína com 5 aminoácidos a mais.

4.2.3.1. Variante de número de cópias gênicas

Em uma paciente (Pt 16) fruto de união consanguínea, com acometimento clínico moderado, foi encontrada a deleção homozigótica dos éxons 1 e 2 do gene *TMEM38B* (Figura 6).

Na análise do SPLE do painel gênico não foi encontrada nenhuma variante de interesse na sequência nucleotídica dos genes candidatos, instigando a análise comparativa do número de cópias gênicas utilizando CONTRA. Uma possível diminuição do número de cópias foi encontrada em região do cromossomo 9 envolvendo o gene *TMEM38B*. A visualização do sequenciamento paralelo no IGV revelou ausência de leituras nos éxons 1 e 2 na amostra da paciente. Para confirmação, foi realizada análise por *array* citogenômico, que revelou perda de heterozigose em região do braço longo do cromossomo 9 abrangendo seis sondas subsequentes para as quais não houve hibridação, delimitando, portanto, área de deleção em homozigose de chr9:108,449,986 a chr9:108,474,679, incluindo os éxons 1 e 2 de *TMEM38B*. A deleção dos éxons 1 e 2 de *TMEM38B* já foi descrita anteriormente em associação a OI (Rubinato et al., 2014).

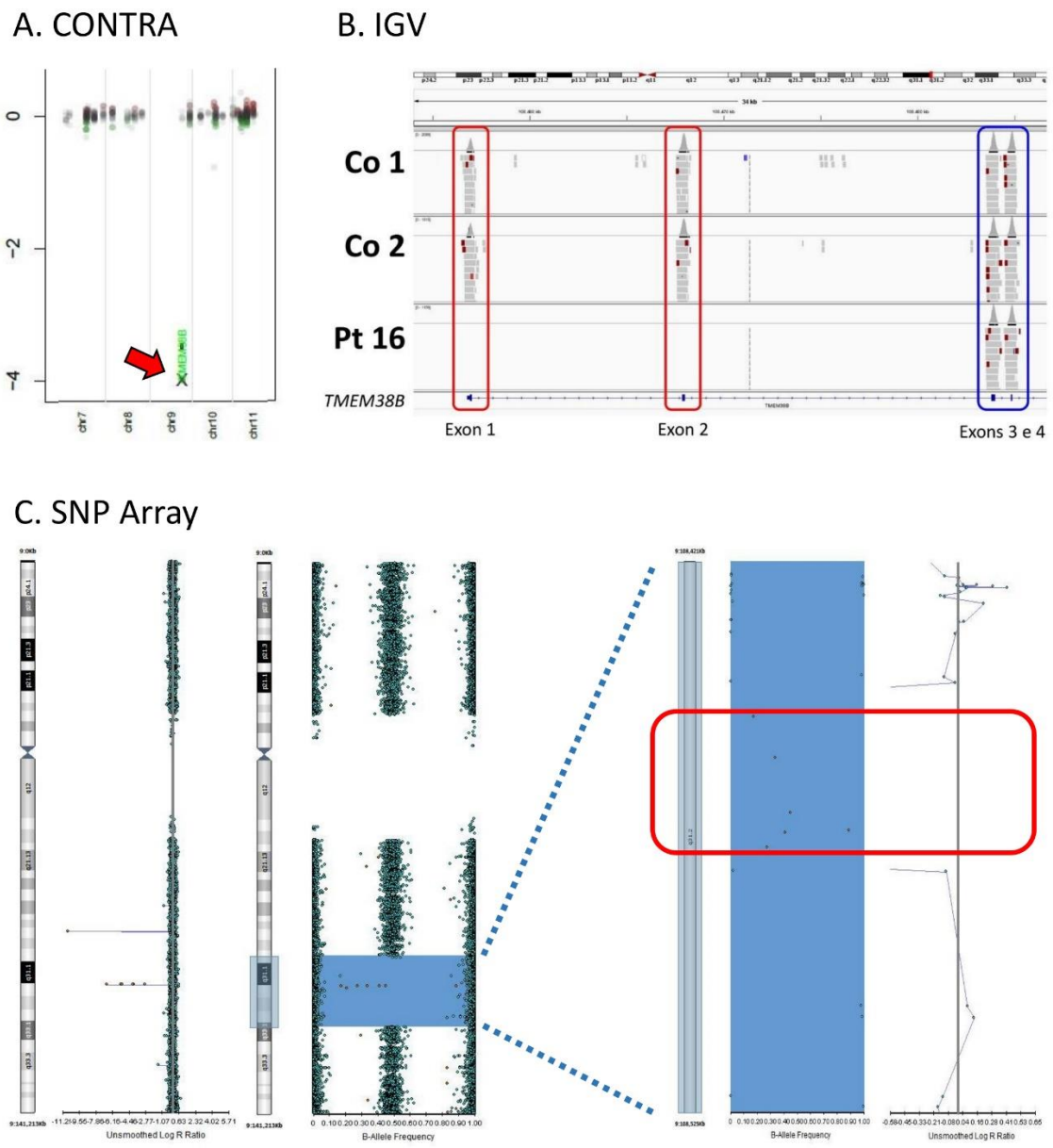

Figura 6 - Deleção homozigótica dos éxons 1 e 2 do gene *TMEM38B*
A) A análise dos dados do sequenciamento paralelo da paciente 16 utilizando o software CONTRA revelou perda de cópias gênicas no locus *TMEM38B* (seta vermelha). B) A visualização do sequenciamento no IGV, comparando a paciente 16 (Pt 16) a dois outros indivíduos sequenciados concomitantemente (Co 1 e Co 2) mostrou equivalência do número de leituras dos éxons 3 e 4 nas três amostras (caixa azul), porém ausência de leitura nos éxons 1 e 2 da paciente (caixas vermelhas). C) Confirmação por *array* citogenômico: em azul denota-se região do braço longo do cromossomo 9 com perda de heterozigose; no destaque desta região, delimitam-se 6 sondas subsequentes (área demarcada pela caixa vermelha) para as quais não houve hibridação, revelando deleção em homozigose de chr9:108,449,986 a chr9:108,474,679, englobando os éxons 1 e 2 de *TMEM38B*.

4.3. Associação entre achado molecular e características clínicas

Buscou-se associar gravidade do acometimento esquelético, esclera azulada, DGI, perda auditiva e alteração ecocardiográfica ao achado molecular. As informações clínicas a respeito da gravidade da OI, esclera azulada e dentinogênese imperfeita estavam disponíveis para toda a coorte. A avaliação audiométrica estava disponível em 34 indivíduos, sendo 16 com variantes em *COL1A1*, 12 com variantes em *COL1A2* e 6 com variantes em outros genes. A avaliação por ecocardiograma estava disponível em 39 indivíduos, sendo 19 com variantes em *COL1A1*, 13 com variantes em *COL1A2* e 7 com variantes em outros genes.

Na análise inicial, os indivíduos foram agrupados em: defeitos em *COL1A1* (n=24), defeitos em *COL1A2* (n=14), e defeitos em outros genes (n=10). Houve associação significativa com gravidade (Tabela 9, p=0,012), observando-se maior proporção de apresentação leve nos defeitos de *COL1A1* e de apresentação moderada a grave nos defeitos nos genes não colágenos. Também houve associação significativa com esclera azulada (Tabela 9, p=0,009), presente em maior proporção nos defeitos de *COL1A1*.

Subsequentemente, comparou-se defeitos nos genes que codificam as cadeias do colágeno tipo 1 (*COL1A1* e *COL1A2*) com defeitos nos demais genes candidatos. Houve associação significativa apenas com gravidade (Tabela 10, p=0,008), observando-se maior proporção de apresentação moderada a grave nos defeitos não colágenos.

Por fim, dentre os defeitos dos genes do colágeno 1, comparou-se os de *COL1A1* aos de *COL1A2*, e trocas de glicina às demais variantes. Houve associação significativa de variantes em *COL1A1* e presença de esclera azulada (Tabela 11, p=0,006) e de variantes com troca de glicina e presença de DGI (Tabela 12, p=0,003).

Tabela 9 - Associação entre defeitos em *COL1A1*, *COL1A2* ou nos demais genes candidatos e achados clínicos

			COL1A1	*COL1A2*	Outros genes	
Gravidade	Leve	n	13 (54%)	5 (36%)	0 (0%)	p=0,012[a]
		esperado	9,0 (37%)	5,2 (37%)	3,7 (37%)	
	Moderada a grave	n	11 (46%)	9 (64%)	10 (100%)	
		esperado	15,0 (63%)	8,8 (63%)	6,3 (63%)	
Esclera azulada	Sim	n	23 (96%)	8 (57%)	7 (70%)	p=0,009[b]
		esperado	19,0 (79%)	11,1 (79%)	7,9 (79%)	
	Não	n	1 (4%)	6 (43%)	3 (30%)	
		esperado	5,0 (21%)	2,9 (21%)	2,1 (21%)	
DGI	Sim	n	6 (26%)	6 (43%)	2 (20%)	p=0,423[b]
		esperado	6,9 (30%)	4,2 (30%)	3,0 (30%)	
	Não	n	17 (74%)	8 (47%)	8 (80%)	
		esperado	16,1 (70%)	9,8 (70%)	7,0 (70%)	
Perda auditiva	Sim	n	2 (12%)	3 (25%)	1 (17%)	p=0,695[b]
		esperado	2,8 (17%)	2,1 (17%)	1,1 (17%)	
	Não	n	14 (88%)	9 (75%)	5 (83%)	
		esperado	13,2 (83%)	9,9 (83%)	4,9 (83%)	
Alteração ecocardio-gráfica	Sim	n	5 (26%)	2 (15%)	2 (29%)	p=0,706[b]
		esperado	4,4 (23%)	3,0 (23%)	1,6 (23%)	
	Não	n	14 (74%)	11 (85%)	5 (71%)	
		esperado	14,6 (77%)	10,0 (77%)	5,4 (77%)	

[a]Teste qui-quadrado de Pearson; [b]Razão de verossimilhança

Tabela 10 - Associação entre defeitos nos genes do colágeno 1 (*COL1A1* e *COL1A2*) ou nos demais genes candidatos e achados clínicos

			Colágeno 1	Outros genes	
Gravidade	Leve	n	18 (47%)	0 (0%)	*p*=0,008[a]
		esperado	14,3 (37%)	3,7 (37%)	
	Moderada a grave	n	20 (53%)	10 (100%)	
		esperado	23,8 (63%)	6,3 (63%)	
Esclera azulada	Sim	n	31 (82%)	7 (70%)	*p*=0,414[a]
		esperado	30,1 (79%)	7,9 (79%)	
	Não	n	7 (18%)	3 (30%)	
		esperado	7,9 (21%)	2,1 (21%)	
DGI	Sim	n	12 (32%)	2 (20%)	*p*=0,700[a]
		esperado	11,0 (30%)	3,0 (30%)	
	Não	n	25 (68%)	8 (80%)	
		esperado	26,0 (70%)	7,0 (70%)	
Perda auditiva	Sim	n	5 (18%)	1 (17%)	*p*=1,000[a]
		esperado	4,9 (18%)	1,1 (18%)	
	Não	n	23 (82%)	5 (83%)	
		esperado	23,1 (82%)	4,9 (82%)	
Alteração ecocardio-gráfica	Sim	n	7 (22%)	2 (29%)	*p*=0,653[a]
		esperado	7,4 (23%)	1,6 (23%)	
	Não	n	25 (78%)	5 (71%)	
		esperado	24,6 (77%)	5,4 (77%)	

[a]Teste exato de Fisher

Tabela 11 - Associação entre defeitos em *COL1A1* ou *COL1A2* e achados clínicos

			COL1A1	*COL1A2*	
Gravidade	Leve	n	13 (54%)	5 (36%)	p=0,272[a]
		esperado	11,4 (47%)	6,6 (47%)	
	Moderada a grave	n	11 (46%)	9 (64%)	
		esperado	12,6 (53%)	7,4 (53%)	
Esclera azulada	Sim	n	23 (96%)	8 (57%)	p=0,006[b]
		esperado	19,6 (82%)	11,4 (82%)	
	Não	n	1 (4%)	6 (43%)	
		esperado	4,4 (18%)	2,6 (18%)	
DGI	Sim	n	6 (26%)	6 (43%)	p=0,470[b]
		esperado	7,5 (32%)	4,5 (32%)	
	Não	n	17 (74%)	8 (57%)	
		esperado	15,5 (68%)	9,5 (68%)	
Perda auditiva	Sim	n	2 (12%)	3 (25%)	p=0,624[b]
		esperado	2,9 (18%)	2,1 (18%)	
	Não	n	14 (88%)	9 (75%)	
		esperado	13,1 (82%)	9,9 (82%)	
Alteração ecocardio-gráfica	Sim	n	5 (26%)	2 (15%)	p=0,671[b]
		esperado	4,2 (22%)	2,8 (22%)	
	Não	n	14 (74%)	11 (85%)	
		esperado	14,8 (78%)	10,2 (78%)	

[a]Teste qui-quadrado de Pearson; [b]Teste exato de Fisher

Tabela 12 - Associação entre substituição de glicina ou outras variantes em COL1A1 e COL1A2 e achados clínicos

			Substituição de Glicina	Outras variantes	
Gravidade	Leve	n	11 (44%)	7 (54%)	*p*=0,564[a]
		esperado	11,8 (47%)	6,2 (47%)	
	Moderada a grave	n	14 (56%)	6 (46%)	
		esperado	13,2 (53%)	6,8 (53%)	
Esclera azulada	Sim	n	18 (72%)	13 (100%)	*p*=0,072[b]
		esperado	20,4 (82%)	10,6 (82%)	
	Não	n	7 (28%)	0 (0%)	
		esperado	4,6 (18%)	2,4 (18%)	
DGI	Sim	n	12 (48%)	0 (0%)	*p*=0,003[b]
		esperado	8,1 (32%)	3,9 (32%)	
	Não	n	13 (52%)	12 (100%)	
		esperado	16,9 (68%)	8,1 (68%)	
Perda auditiva	Sim	n	3 (16%)	2 (22%)	*p*=1,000 [b]
		esperado	3,4 (18%)	1,6 (18%)	
	Não	n	16 (84%)	7 (78%)	
		esperado	15,6 (82%)	7,4 (82%)	
Alteração ecocardio-gráfica	Sim	n	4 (18%)	3 (30%)	*p*=0,648[b]
		esperado	4,8 (22%)	2,2 (22%)	
	Não	n	18 (82%)	7 (70%)	
		esperado	17,2 (78%)	7,8 (78%)	

[a]Teste qui-quadrado de Pearson; [b]Teste exato de Fisher

5. Discussão

Este projeto foi desenvolvido com o objetivo de obter o diagnóstico molecular de pacientes com diagnóstico clínico de OI através de SPLE. A utilização de um painel com 15 genes candidatos permitiu a identificação do diagnóstico molecular em 37 de 38 casos (97%).

A coorte estudada é predominantemente composta por adultos, com mediana de idade de 24 anos, sendo apenas 9 indivíduos menores de 18 anos. A maioria dos estudos publicados acerca de OI é realizada em crianças e adolescentes, já que a doença tende a se manifestar mais pronunciadamente nestas fases da vida (Patel et al., 2015; Rauch et al., 2010). Tal peculiaridade do estudo atual reflete as prerrogativas do atendimento no ambulatório de doenças osteometabólicas da Divisão de Endocrinologia do Instituto Central do HCFMUSP, que, dentro da estrutura do HCFMUSP, é responsável pelo seguimento de adultos com OI.

O cenário no qual este estudo foi desenvolvido também impacta na composição da gravidade clínica da OI nesta coorte. Os pacientes estudados apresentam características clínicas heterogêneas, desde quadros leves com poucas fraturas e sem deformidades até quadros graves com centenas de fraturas e múltiplas deformidades, mas predomina o acometimento moderado a grave (63%). De fato, 53% tem comprometimento da mobilidade. A complexidade da assistência à saúde em um centro terciário de referência como o HCFMUSP justifica este perfil de pacientes, que não necessariamente reflete a gravidade da manifestação de OI na população como um todo.

Do ponto de vista terapêutico, verificou-se que o tratamento ortopédico e medicamentoso destes indivíduos também foi heterogêneo. Apesar de ser preconizado que, na atenção ortopédica a OI, todas as fraturas sejam reduzidas a fim de se evitar deformidades, em 35% dos pacientes o tratamento das fraturas foi apenas conservador, e

muitas vezes sem alinhamento adequado, potencialmente implicando na alta prevalência (67%) de deformidades em ossos longos. O tratamento medicamentoso com bisfosfonatos foi heterogêneo já que muitos pacientes iniciaram seguimento em outros serviços, sendo apenas encaminhados após a idade adulta, e também pela indefinição de um protocolo uniforme de tratamento até 2013. Assim, não se considerou adequado tentar se associar os achados moleculares da coorte a desfechos de tratamento.

O advento do sequenciamento paralelo em larga escala vem facilitando o diagnóstico molecular em diversas doenças genéticas, mas também resultando na identificação de grande número de variantes às quais significado biológico ou clínico é de difícil atribuição (Richards et al., 2015; Tucker et al., 2009). Neste sentido, as diretrizes propostas pelo ACMG-AMP buscam estabelecer critérios a serem seguidos na definição de patogenicidade de variantes alélicas (Richards et al., 2015). Dos 37 casos aos quais foi aqui atribuído um diagnóstico molecular, 4 apresentavam apenas variantes classificadas como VSIs. Se estes diagnósticos fossem desconsiderados, a taxa de sucesso do diagnóstico molecular reduziria de 97% para 87%. Entretanto, acreditou-se que estas VSIs têm papel na patogênese destes casos, como será discutido a seguir, e por isso foram incluídas nos resultados do estudo.

Na paciente 12, fruto da união de primos de primeiro grau, foi encontrada em homozigose a VSI não sinônima p.(Gln60Pro)(;)(Gln60Pro) em *FKBP10*, gene associado a OI recessiva. Esta variante não foi descrita anteriormente em associação a OI, é ausente em bancos de dados populacionais e tem predição de ser deletéria pelas ferramentas *in silico* PolyPhen2 e CADD. Na análise de segregação familiar, confirmou-se que a mãe não afetada é portadora da variante em heterozigose, mas infelizmente não foi possível obter amostra de DNA nem avaliação fenotípica do pai (falecido) para estudo. No paciente 14 foi encontrada em heterozigose a VSI não sinônima COL1A1 p.(Gly719Cys)

com substituição de glicina; este tipo de substituição tem, habitualmente, grande impacto na formação da tripla hélice do colágeno (Forlino & Marini, 2016). Tal variante não é encontrada nos bancos de dados populacionais e é predita como deletéria pelas três ferramentas *in silico* utilizadas. Trata-se de variante nova, ainda não descrita em associação a OI, e a análise de segregação familiar foi inconclusiva já que, apesar da mãe não afetada não ser portadora da variante, não foi possível obter amostra de DNA do pai.. No paciente 20 foram encontradas as duas VSIs heterozigóticas p.(Tyr547His) e p.(Cys282Arg) em *PLOD2*, gene associado a OI recessiva. Esse paciente nega antecedente de uniões consanguíneas na família. Ambas as variantes não foram previamente reportadas na literatura em associação a OI, apresentam frequência populacional extremamente baixas, e predição de impacto deletério *in silico*. Não foi possível obter amostra de DNA de nenhum familiar deste paciente. Por fim, nas pacientes mãe e filha com OI da família 4 foi encontrada a combinação de duas VSIs heterozigóticas não sinônimas em *COL1A2*, p.(Gly229Asp) e p.(Pro230Gln), a primeira envolvendo a troca de um resíduo glicina. Ambas não descritas anteriormente em associação com OI, não são encontradas em bancos de dados populacionais e apresentam predição de deletéria *in silico*. Na análise de segregação, duas irmãs não afetadas da mãe foram estudadas e não carreavam estas variantes. Frente aos elementos apresentados, considerou-se que estas VSIs tinham potencial papel patogênico nestes casos, e, por isso, foram atribuídas como diagnóstico molecular. Destaque-se que dos 9 maiores estudos de diagnóstico molecular de OI publicados entre 2015 e 2019, apenas um reportou as variantes de acordo com a classificação ACMG-AMP, e também incluiu VSIs na atribuição de diagnóstico (Mrosk et al., 2018).

Em apenas um caso estudado (Pt 15) não se identificou o diagnóstico molecular. O diagnóstico clínico de OI neste indivíduo foi revisado e confirmado. Trata-se de um

quadro grave, com total de 40 fraturas ao longo da vida, presença de deformidades em membros superiores e inferiores e baixa estatura (Z -7,2). É improvável que o diagnóstico molecular tenha sido perdido por motivos técnicos já que a cobertura das regiões alvo nesta amostra foi superior a 50 vezes. É possível, entretanto, que o defeito esteja em áreas não codificadoras não analisadas (íntron, promotor), que decorra de novos mecanismos genéticos em OI, em genes ainda não identificados, ou que esteja nos genes *SPARC* ou *MBTPS2,* associados a OI após o início deste trabalho e, portanto, não incluídos no painel (Mendoza-Londono et al., 2015; Lindert et al., 2016).

5.1. Variantes identificadas em *COL1A1* e *COL1A2*

Em acordo com a literatura, defeitos em *COL1A1* e *COL1A2* foram a principal causa de OI nesta coorte (71%), incluindo todos os casos com OI leve. Nestes, aproximadamente metade das variantes levam a alterações quantitativas do colágeno 1 (variantes resultando em códon de parada prematuro), enquanto a outra metade leva a alterações qualitativas (variantes não sinônimas). Este achado vai contra o paradigma histórico de que OI leve estaria sempre associada à haploinsuficiência de uma das cadeias do colágeno tipo 1 (Forlino et al., 2011; Rauch & Glorieux, 2004), e reflete o novo entendimento molecular que vem tomando forma nos últimos anos (Forlino & Marini, 2016). Por exemplo, Lindahl et al. analisaram características clínicas e genéticas de 99 pacientes com OI leve por defeitos em *COL1A1* ou *COL1A2*, e reportaram 62% de alterações quantitativas e 32% de alterações qualitativas, sem diferença significativa na manifestação clínica (Lindahl et al., 2015). Assim, este estudo vem corroborar que alterações não sinônimas no colágeno, inclusive substituições de glicina, podem resultar em OI leve.

Não se encontrou região *hotspot* de variantes em *COL1A1*, já que as variantes encontradas estão distribuídas ao longo de toda a proteína (Figura 3). Já em *COL1A2*, a maioria das variantes se concentrou no primeiro terço da proteína, achado que não é corroborado pela literatura (Forlino & Marini, 2016; Li et al., 2019; Marini et al., 2017). Esta diferença pode ser explicada pelo menor número de variantes encontradas em *COL1A2*. Tanto para *COL1A1* quanto para *COL1A2* não se observou correlação de posição da variante com a apresentação clínica, já que variantes associadas a quadros leves, moderados e graves estiveram dispersas ao longo das proteínas (Figura 3).

Além da pobre correlação genótipo fenótipo no que tange a posição das variantes no colágeno 1, observou-se variabilidade fenotípica em indivíduos portadores de uma mesma variante. Na família 2, com OI causada por COL1A2 p.(Gly319Arg), o indivíduo F2a apresentou acometimento leve e o indivíduo F2b acometimento moderado. Nos 6 portadores da variante recorrente COL1A2 p.(Gly772Ser), a manifestação da OI também foi variável conforme mostrado na Tabela 8. Não se encontrou nesses indivíduos variantes nos genes candidatos que expliquem esta variabilidade fenotípica que, também na literatura permanece inexplicada. Esta variabilidade tem implicação na previsão prognóstica dos pacientes com OI mesmo após diagnóstico molecular.

Mesmo assim, coletivamente, OI causada por defeitos nos genes do colágeno 1 pode estar associada a particularidades de apresentação. Neste estudo, variantes nestes genes estiveram associadas a OI de menor gravidade do que aquelas encontradas nos genes não colágenos, variantes em *COL1A1* estiveram mais associadas a esclera azulada do que variantes em *COL1A2*, e variantes com substituição de glicina estiveram mais associadas à dentinogênese imperfeita do que os outros defeitos de *COL1A1* e *COL1A2*. Estes achados encontram respaldo na literatura (Bardai et al., 2016; Li et al., 2019).

Nas famílias 3, 6 e 7 e no paciente 14 foi referido antecedente de consanguinidade. Apesar da consanguinidade, foram identificadas variantes heterozigóticas em *COL1A1* ou *COL1A2* nestes indivíduos, destacando a vantagem de uma abordagem não direcionada através do painel de SPLE para elucidação diagnóstica. Em especial na família 7, os dois indivíduos acometidos F7a e F7b eram os únicos dois filhos de pais consanguíneos, e nenhum outro caso de OI era referido na família, direcionando a suspeita para OI recessiva. Mesmo assim, foi identificada a variante heterozigótica COL1A1 p.(Ser271Glnfs*16) como a causa da OI. Descobriu-se posteriormente que o pai (falecido) era portador da variante, e que apesar de nunca ter fraturado, tinha baixa estatura e doença valvar cardíaca, compatíveis com manifestações do espectro de OI. Em dois outros casos (Pt 25 e Pt 27) o diagnóstico molecular dos indivíduos índices também permitiu reconhecer familiares afetados, mas com manifestações leves e sem diagnóstico clínico prévio de OI. Coletivamente estes achados demonstram a relevância do diagnóstico molecular para a medicina de precisão.

5.2. Contextualização dos achados moleculares

Nove estudos prévios principais publicados a partir de 2015 buscaram identificar o diagnóstico molecular de OI em coortes de diferentes países (Patel et al., 2015; Lindahl et al., 2015; Bardai et al., 2016; Liu et al., 2017; Caparros-Martin et al., 2017; Mrosk et al., 2018; Mohd Nawawi et al., 2018; Li et al., 2019; Maioli et al., 2019). Conforme mostrado na Tabela 13, estes estudos foram bastante heterogêneos em termos de desenho experimental, metodologia utilizada, genes candidatos avaliados e características clínicas dos casos estudados. Mesmo assim, a análise conjunta dos seus resultados permite contextualizar os achados deste projeto (Figura 7).

Tabela 13 - Estudos de diagnóstico molecular em OI

Referência	Contexto do estudo	Casos índices	OI moderada a grave	Metodologia	Genes candidatos a OI analisados
Patel et al., 2015	EUA e Canadá (86% brancos)	370	58%	Sanger	5 *COL1A1, COL1A2, IFITM5; CRTAP, SERPINF1*
Lindahl et al., 2015	Suécia	164	32%	Sanger	2 *COL1A1, COL1A2*
Bardai et al., 2016	Canadá (76% brancos)	487	57%	87% Sanger + 17% Painel SPLE	16 *COL1A1, COL1A2, CRTAP, P3H1, SERPINF1, IFITM5, BMP1, CREB3L1, FKBP10, PLOD2, PPIB, SERPINH1, SP7, SPARC, TMEM38B, WNT1*
Liu et al., 2017	China	101	72%	Painel gênico SPLE	14 *COL1A1, COL1A2, IFITM5, CRTAP, SERPINF1, LEPRE1, PPIB, FKBP10, SERPINH1, SP7, PLOD2, TMEM38B, BMP1, WNT1*
Caparros-Martin et al., 2017	Egito, Espanha e Turquia	42	90%	Painel SPLE + Sanger*	15 *BMP1, COL1A1, COL1A2, CREB3L1, CRTAP, FKBP10, IFITM5, P3H1, PLOD2, PPIB, SERPINF1, SERPINH1, SP7, TMEM38B, WNT1*
Mrosk et al., 2018	Índia	50	67%	Painel gênico SPLE	15 *BMP1, COL1A1, COL1A2, CREB3L1, CRTAP, FKBP10, IFITM5, P3H1, PLOD2, PPIB, SERPINF1, SERPINH1, SP7, TMEM38B, WNT1*
Mohd Nawawi et al., 2018	Malásia	28	83%	Painel gênico SPLE	14 *BMP1, COL1A1, COL1A2, CREB3L1, CRTAP, FKBP10, IFITM5, P3H1, PLOD2, PPIB, SERPINF1, SERPINH1, TMEM38B, WNT1*
Li et al., 2019	China	378	61%	Sanger + diversas**	17 *COL1A1, COL1A2, CRTAP, P3H1, SERPINF1, IFITM5, BMP1, CREB3L1, FKBP10, PLOD2, MBTPS2, PPIB, SERPINH1, SP7, SPARC, TMEM38B, WNT1*
Maioli et al., 2019	Itália	295	28%	Desnaturação de alta resolução	2 *COL1A1, COL1A2*
Este estudo	*Brasil*	*38*	*63%*	*Painel gênico SPLE*	*15* *BMP1, COL1A1, COL1A2, CREB3L1, CRTAP, FKBP10, IFITM5, P3H1, PLOD2, PPIB, SERPINF1, SERPINH1, SP7, TMEM38B, WNT1*

*20 casos sem consanguinidade foram analisados por painel, e 22 casos com consanguinidade foram analisados por Sanger; **Foram realizados exoma e painel em alguns casos, mas o estudo não deixa claro em quantos casos

		EUA + Canadá	Suécia	Canadá	China	Egito, Espanha e Turquia	Índia	Malásia	China	Itália	Brasil
Síntese e estrutura do colágeno 1	***COL1A1***	62%	69%	61%	41%	19%	33%	54%	48%	63%	47%
	COL1A2	27%	18%	27%	32%	14%	16%	15%	29%	22%	24%
Modificação pós-traducional do colágeno 1	***CRTAP***	1%		2%		7%			0,3%		
	P3H1			1%		2%	10%	12%			4%
	PPIB			0,2%							
Crosslinking e processamento do colágeno 1	***SERPINH1***								0,5%		
	BMP1						4%	4%			
	FKBP10			0,5%	3%	5%	8%		1%		5%
	PLOD2			0,2%	1%	5%			0,3%		5%
Mineralização óssea	***IFITM5***	2%		5%	10%	7%	4%	12%	3%		3%
	SERPINF1	1%		2%	4%	12%	8%	4%	2%		5%
Diferenciação e função de osteoblastos	***SP7***										
	TMEM38B				3%	5%					3%
	WNT1			0,7%	4%	7%	16%		5%		1%
	CREB3L1										

Figura 7 - Panorama da distribuição diagnóstica de OI em diferentes contextos
Representação da distribuição dos defeitos moleculares identificados nos estudos detalhados na Tabela 13, agrupados de acordo com mecanismo fisiopatológico. Dentro de cada estudo, a intensidade da cor de fundo da célula é proporcional à prevalência relativa dos genes identificados.

No presente estudo, em 71% dos casos foram identificadas variantes nos genes que codificam o colágeno tipo 1 (*COL1A1* 47%; *COL1A2* 24%). Historicamente, a literatura sugere que 85 a 90% dos casos de OI sejam relacionados a defeitos nesses genes (Forlino & Marini, 2016). De fato, nos estudos de Lindahl et al., Bardai et al., e Maioli et al., realizados, respectivamente, nas populações sueca, canadense e italiana, e que empregaram predominantemente como método de análise molecular o sequenciamento Sanger de *COL1A1* e *COL1A2*, as proporções diagnósticas de defeitos no colágeno 1 ficaram entre 85 e 88% (Lindahl et al., 2015; Bardai et al., 2016; Maioli et al., 2019). Ressalta-se que predominam fenótipos leves de OI nos estudos de Lindahl et al. e Maioli et al., e que aproximadamente metade da coorte de Bardai et al. também tinha acometimento leve, potencialmente influenciando estes resultados (Tabela 13).

Por outro lado, os estudos de Liu et al., Mrosk et al., e Mohd Nawawi et al. em coortes chinesa, indiana e malaia, respectivamente, com predominância de OI moderada a grave (Tabela 13) e empregando painel gênico por SPLE, encontraram prevalência bastante variável de defeitos no colágeno 1, de 49 a 73% (Liu et al., 2017; Mohd Nawawi et al., 2018; Mrosk et al., 2018), distribuição que se assemelha à deste estudo. Neste sentido, pode-se aventar que características étnicas e socioculturais possam influenciar as prevalências relativas das causas moleculares de OI. De fato, 26% da coorte referia histórico de consanguinidade na família, tornando possível aventar que a maior frequência de consanguinidade em regiões do Brasil esteja influenciando o cenário da OI em nosso meio (Santos et al., 2010). Além disso, é possível que a análise concomitante dos diversos genes candidatos por SPLE permita atribuir o diagnóstico com maior acurácia.

De todo modo, é provável que a composição da coorte aqui estudada tenha influenciado o achado de 26% de defeitos não colágenos já que, pelas características do

atendimento terciário à saúde no HCFMUSP, 63% dos pacientes estudados tinham acometimento moderado a grave. Nos estudos prévios e também no estudo atual evidencia-se que, dentro das coortes, acometimento moderado a grave esteve mais associado a defeitos não colágenos, e acometimento leve a defeitos em *COL1A1* e *COL1A2*. Por exemplo, no estudo de Bardai et al., nos pacientes com acometimento leve apenas foram identificados defeitos em *COL1A1* e *COL1A2*, enquanto que nos pacientes com acometimento mais grave identificou-se 22% de defeitos em genes não colágenos (Bardai et al., 2016). No presente estudo, todos os pacientes com OI leve também tiveram o diagnóstico molecular restrito a *COL1A1* e *COL1A2*, enquanto que 32% dos pacientes com formas moderadas a graves apresentaram defeitos nos demais genes candidatos.

No que diz respeito à distribuição das causas não colágenas nas diversas coortes, chama a atenção a grande heterogeneidade entre os estudos. Descrições clínicas de coortes de OI a partir dos anos 2000 atribuíam prevalência de aproximadamente 5% ao que se chamava de OI tipo V, posteriormente associada ao defeito na região 5' não traduzida de *IFITM5* (Cho et al., 2012; Glorieux et al., 2000; Semler et al., 2012). Perdura desde então o conceito de que esta é a causa molecular não colágena mais comum. Observando os achados dos estudos moleculares resumidos na Figura 7, nota-se que em algumas coortes a prevalência relativa do defeito em *IFITM5* é consideravelmente baixa. No estudo atual, em apenas uma paciente (Pt 24) foi encontrada a variante heterozigótica c.-14 C->T em *IFITM5*, tornando essa uma causa não colágena menos frequente do que defeitos em *SERPINF1*, *FKBP10*, *PLOD2* e *P3H1* nesta coorte. Nas coortes de Caparros-Martin et al., Mrosk et al. e Li et al., *IFITM5* também não se mostrou como a principal causa de OI não colágena, sugerindo que há diversidade no diagnóstico molecular de OI e que conceitos estabelecidos em determinadas populações não se aplicam a outras (Caparros-Martin et al., 2017; Li et al., 2019; Mrosk et al., 2018).

Ainda dentre as causas de OI não colágena, chama a atenção neste estudo 10% de defeitos em *FKBP10* e *PLOD2*, genes associados a OI e Síndrome de Bruck, uma forma de fragilidade óssea acompanhada de contraturas musculares e pterígio (Alanay et al., 2010; Leal et al., 2018; Puig-Hervas et al., 2012). O reconhecimento das contraturas musculares pode ser difícil nos adultos frente às múltiplas deformidades ósseas resultantes da fragilidade óssea exuberante. A prevalência relativa de defeitos nesses genes como causa de OI é bastante variável: enquanto Bardai et al. e Li et al. encontraram prevalência ao redor de 1%, Mrosk et al. e Caparros-Martin et al. encontraram prevalência de 8 a 10%, semelhante ao achado deste estudo. Novamente, é possível que diferenças na frequência de uniões consanguíneas nestas populações explique essa discrepância.

A taxa de sucesso do diagnóstico molecular nas coortes prévias é relativamente homogênea frente às grandes heterogeneidades metodológicas entre os estudos. Estudos empregando painel gênico por SPLE como os de Liu et al., Mrosk et al. e Mohd Nawawi et al. obtiveram respectivamente 87%, 96% e 90% de sucesso diagnóstico (Liu et al., 2017; Mrosk et al., 2018; Mohd Nawawi et al., 2018). Estudos com desenho misto, incluindo abordagem direcionada a gene candidato seguida de abordagem não direcionada genômica, como os de Bardai et al. e Li et al., reportam sucesso de 98% e 90%. No presente estudo, a taxa de sucesso de 97% se destaca, indicando que o emprego de painel gênico por SPLE deva se tornar metodologia de escolha para diagnóstico molecular de OI.

6. Conclusões

- O diagnóstico molecular da osteogênese imperfeita através do SPLE de painel de 15 genes candidatos foi obtido em 37 de 38 casos (97%).
- A maioria dos casos (71%) apresentou defeitos em *COL1A1* ou *COL1A2*, que codificam o colágeno tipo 1, enquanto 26% apresentaram defeitos em outros genes candidatos, destacando-se *P3H1*, *FKBP10*, *PLOD2* e *SERPINF1*. Frente à literatura, houve maior proporção de defeitos em genes não colágenos, que pode ser atribuída à gravidade dos pacientes incluídos no estudo, já que 63% tinham acometimento moderado a grave.
- Houve variabilidade clínica entre indivíduos acometidos de mesma família e entre os seis indivíduos portadores da variante COL1A2 p.(Gly772Ser). Não foram encontradas variantes genéticas adicionais que pudessem explicar esta variabilidade de expressão fenotípica.
- Neste estudo, variantes nos genes não colágenos estiveram associados a OI de maior gravidade, variantes em *COL1A1* à manifestação de esclera azulada, e, dentre as variantes de *COL1A1* e *COL1A2*, aquelas que resultaram em substituição de glicina estiveram associadas à dentinogênese imperfeita.

7. Referências

Adzhubei, I., Jordan, D.M., and Sunyaev, S.R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet *Chapter 7*, Unit7.20.

Alanay, Y., Avaygan, H., Camacho, N., Utine, G.E., Boduroglu, K., Aktas, D., Alikasifoglu, M., Tuncbilek, E., Orhan, D., Bakar, F.T., *et al.* (2010). Mutations in the gene encoding the RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta. Am J Hum Genet *86*, 551-559.

Amberger, J.S., Bocchini, C.A., Schiettecatte, F., Scott, A.F., and Hamosh, A. (2015). OMIM.org: Online Mendelian Inheritance in Man (OMIM®), an online catalog of human genes and genetic disorders. Nucleic Acids Res *43*, D789-798.

Baim, S., Binkley, N., Bilezikian, J.P., Kendler, D.L., Hans, D.B., Lewiecki, E.M., and Silverman, S. (2008). Official Positions of the International Society for Clinical Densitometry and executive summary of the 2007 ISCD Position Development Conference. J Clin Densitom *11*, 75-91.

Balasubramanian, M., Wagner, B.E., Peres, L.C., Sobey, G.J., Parker, M.J., Dalton, A., Arundel, P., and Bishop, N.J. (2015). Ultrastructural and histological findings on examination of skin in osteogenesis imperfecta: a novel study. Clin Dysmorphol *24*, 45-54.

Bardai, G., Moffatt, P., Glorieux, F.H., and Rauch, F. (2016). DNA sequence analysis in 598 individuals with a clinical diagnosis of osteogenesis imperfecta: diagnostic yield and mutation spectrum. Osteoporos Int.

Becker, J., Semler, O., Gilissen, C., Li, Y., Bolz, H.J., Giunta, C., Bergmann, C., Rohrbach, M., Koerber, F., Zimmermann, K., *et al.* (2011). Exome sequencing identifies truncating mutations in human SERPINF1 in autosomal-recessive osteogenesis imperfecta. Am J Hum Genet *88*, 362-371.

Ben Amor, I.M., Glorieux, F.H., and Rauch, F. (2011). Genotype-phenotype correlations in autosomal dominant osteogenesis imperfecta. J Osteoporos *2011*, 540178.

Boright, A.P., Lancaster, G.A., and Scriver, C.R. (1984). Osteogenesis imperfecta: a heterogeneous morphologic phenotype in cultured dermal fibroblasts. Hum Genet *67*, 29-33.

Brown, J.P., and Josse, R.G. (2002). 2002 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada. CMAJ *167*, S1-34.

Cabral, W.A., Chang, W., Barnes, A.M., Weis, M., Scott, M.A., Leikin, S., Makareeva, E., Kuznetsova, N.V., Rosenbaum, K.N., Tifft, C.J., *et al.* (2007). Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. Nat Genet *39*, 359-365.

Caparros-Martin, J.A., Aglan, M.S., Temtamy, S., Otaify, G.A., Valencia, M., Nevado, J., Vallespin, E., Del Pozo, A., Prior de Castro, C., Calatrava-Ferreras, L., *et al.* (2017). Molecular spectrum and differential diagnosis in patients referred with sporadic or autosomal recessive osteogenesis imperfecta. Mol Genet Genomic Med *5*, 28-39.

Cho, T.J., Lee, K.E., Lee, S.K., Song, S.J., Kim, K.J., Jeon, D., Lee, G., Kim, H.N., Lee, H.R., Eom, H.H., *et al.* (2012). A single recurrent mutation in the 5'-UTR of IFITM5 causes osteogenesis imperfecta type V. Am J Hum Genet *91*, 343-348.

Christiansen, H.E., Schwarze, U., Pyott, S.M., AlSwaid, A., Al Balwi, M., Alrasheed, S., Pepin, M.G., Weis, M.A., Eyre, D.R., and Byers, P.H. (2010). Homozygosity for a missense mutation in SERPINH1, which encodes the collagen chaperone protein HSP47, results in severe recessive osteogenesis imperfecta. Am J Hum Genet *86*, 389-398.

Cundy, T. (2012). Recent advances in osteogenesis imperfecta. Calcif Tissue Int *90*, 439-449.

Folkestad, L., Hald, J.D., Canudas-Romo, V., Gram, J., Hermann, A.P., Langdahl, B., Abrahamsen, B., and Brixen, K. (2016). Mortality and Causes of Death in Patients With Osteogenesis Imperfecta: A Register-Based Nationwide Cohort Study. J Bone Miner Res.

Forlino, A., Cabral, W.A., Barnes, A.M., and Marini, J.C. (2011). New perspectives on osteogenesis imperfecta. Nat Rev Endocrinol *7*, 540-557.

Forlino, A., and Marini, J.C. (2016). Osteogenesis imperfecta. Lancet *387*, 1657-1671.

Glorieux, F.H., Devogelaer, J.P., Durigova, M., Goemaere, S., Hemsley, S., Jakob, F., Junker, U., Ruckle, J., Seefried, L., and Winkle, P.J. (2017). BPS804 Anti-Sclerostin Antibody in Adults With Moderate Osteogenesis Imperfecta: Results of a Randomized Phase 2a Trial. J Bone Miner Res *32*, 1496-1504.

Glorieux, F.H., Rauch, F., Plotkin, H., Ward, L., Travers, R., Roughley, P., Lalic, L., Glorieux, D.F., Fassier, F., and Bishop, N.J. (2000). Type V osteogenesis imperfecta: a new form of brittle bone disease. J Bone Miner Res *15*, 1650-1658.

Grafe, I., Yang, T., Alexander, S., Homan, E.P., Lietman, C., Jiang, M.M., Bertin, T., Munivez, E., Chen, Y., Dawson, B., *et al.* (2014). Excessive transforming growth factor-β signaling is a common mechanism in osteogenesis imperfecta. Nat Med *20*, 670-675.

Homan, E.P., Rauch, F., Grafe, I., Lietman, C., Doll, J.A., Dawson, B., Bertin, T., Napierala, D., Morello, R., Gibbs, R., *et al.* (2011). Mutations in SERPINF1 cause osteogenesis imperfecta type VI. J Bone Miner Res *26*, 2798-2803.

Keupp, K., Beleggia, F., Kayserili, H., Barnes, A.M., Steiner, M., Semler, O., Fischer, B., Yigit, G., Janda, C.Y., Becker, J., *et al.* (2013). Mutations in WNT1 cause different forms of bone fragility. Am J Hum Genet *92*, 565-574.

Kircher, M., Witten, D.M., Jain, P., O'Roak, B.J., Cooper, G.M., and Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet *46*, 310-315.

Krawczak, M., Ball, E.V., Fenton, I., Stenson, P.D., Abeysinghe, S., Thomas, N., and Cooper, D.N. (2000). Human gene mutation database-a biomedical information and research resource. Hum Mutat *15*, 45-51.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc *4*, 1073-1081.

Laine, C.M., Joeng, K.S., Campeau, P.M., Kiviranta, R., Tarkkonen, K., Grover, M., Lu, J.T., Pekkinen, M., Wessman, M., Heino, T.J., *et al.* (2013). WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. N Engl J Med *368*, 1809-1816.

Landrum, M.J., Lee, J.M., Benson, M., Brown, G., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Hoover, J., *et al.* (2016). ClinVar: public archive of interpretations of clinically relevant variants. Nucleic Acids Res *44*, D862-868.

Lapunzina, P., Aglan, M., Temtamy, S., Caparros-Martin, J.A., Valencia, M., Leton, R., Martinez-Glez, V., Elhossini, R., Amr, K., Vilaboa, N., *et al.* (2010). Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. Am J Hum Genet *87*, 110-114.

Leal, G.F., Nishimura, G., Voss, U., Bertola, D.R., Åström, E., Svensson, J., Yamamoto, G.L., Hammarsjö, A., Horemuzova, E., Papadiogannakis, N., *et al.* (2018). Expanding the Clinical Spectrum of Phenotypes Caused by Pathogenic Variants in PLOD2. J Bone Miner Res *33*, 753-760.

Lee, B. (2019). Safety of Fresolimumab in the Treatment of Osteogenesis Imperfecta (ClinicalTrials.gov: U.S. National Library of Medicine), pp. https://clinicaltrials.gov/ct2/show/NCT03064074.

Li, L., Mao, B., Li, S., Xiao, J., Wang, H., Zhang, J., Ren, X., Wang, Y., Wu, Y., Cao, Y., *et al.* (2019). Genotypic and phenotypic characterization of Chinese patients with osteogenesis imperfecta. Hum Mutat *40*, 588-600.

Lindahl, K., Astrom, E., Rubin, C.J., Grigelioniene, G., Malmgren, B., Ljunggren, O., and Kindmark, A. (2015). Genetic epidemiology, prevalence, and genotype-phenotype correlations in the Swedish population with osteogenesis imperfecta. Eur J Hum Genet *23*, 1042-1050.

Lindert, U., Cabral, W.A., Ausavarat, S., Tongkobpetch, S., Ludin, K., Barnes, A.M., Yeetong, P., Weis, M., Krabichler, B., Srichomthong, C., *et al.* (2016). MBTPS2 mutations cause defective regulated intramembrane proteolysis in X-linked osteogenesis imperfecta. Nat Commun *7*, 11920.

Liu, Y., Asan, Ma, D., Lv, F., Xu, X., Wang, J., Xia, W., Jiang, Y., Wang, O., Xing, X., *et al.* (2017). Gene mutation spectrum and genotype-phenotype correlation in a cohort of Chinese osteogenesis imperfecta patients revealed by targeted next generation sequencing. Osteoporos Int *28*, 2985-2995.

Lu, Y., Zhao, F., Ren, X., Li, Z., Yang, X., and Han, J. (2014). A rare case of osteogenesis imperfecta combined with complete tooth loss. J Pediatr Endocrinol Metab *27*, 143-147.

Maioli, M., Gnoli, M., Boarini, M., Tremosini, M., Zambrano, A., Pedrini, E., Mordenti, M., Corsini, S., D'Eufemia, P., Versacci, P., *et al.* (2019). Genotype-phenotype correlation study in 364 osteogenesis imperfecta Italian patients. Eur J Hum Genet.

Marini, J.C., Forlino, A., Cabral, W.A., Barnes, A.M., San Antonio, J.D., Milgrom, S., Hyland, J.C., Korkko, J., Prockop, D.J., De Paepe, A., *et al.* (2007). Consortium for osteogenesis imperfecta mutations in the helical domain of type I collagen: regions rich in lethal mutations align with collagen binding sites for integrins and proteoglycans. Hum Mutat *28*, 209-221.

Marini, J.C., and Blissett, A.R. (2013). New genes in bone development: what's new in osteogenesis imperfecta. J Clin Endocrinol Metab *98*, 3095-3103.

Marini, J.C., Forlino, A., Bächinger, H.P., Bishop, N.J., Byers, P.H., Paepe, A., Fassier, F., Fratzl-Zelman, N., Kozloff, K.M., Krakow, D., *et al.* (2017). Osteogenesis imperfecta. Nat Rev Dis Primers *3*, 17052.

Martinez-Glez, V., Valencia, M., Caparros-Martin, J.A., Aglan, M., Temtamy, S., Tenorio, J., Pulido, V., Lindert, U., Rohrbach, M., Eyre, D., *et al.* (2012). Identification of a mutation causing deficient BMP1/mTLD proteolytic activity in autosomal recessive osteogenesis imperfecta. Hum Mutat *33*, 343-350.

Mendoza-Londono, R., Fahiminiya, S., Majewski, J., Tétreault, M., Nadaf, J., Kannu, P., Sochett, E., Howard, A., Stimec, J., Dupuis, L., *et al.* (2015). Recessive osteogenesis imperfecta caused by missense mutations in SPARC. Am J Hum Genet *96*, 979-985.

Minillo, R.M., Sobreira, N., de Faria Soares Mde, F., Jurgens, J., Ling, H., Hetrick, K.N., Doheny, K.F., Valle, D., Brunoni, D., and Perez, A.B. (2014). Novel Deletion of SERPINF1 Causes Autosomal Recessive Osteogenesis Imperfecta Type VI in Two Brazilian Families. Mol Syndromol *5*, 268-275.

Mohd Nawawi, N., Selveindran, N.M., Rasat, R., Chow, Y.P., Abdul Latiff, Z., Syed Zakaria, S.Z., Jamal, R., Abdul Murad, N.A., and Abd Aziz, B.B. (2018). Genotype-phenotype correlation among Malaysian patients with osteogenesis imperfecta. Clin Chim Acta *484*, 141-147.

Morello, R., Bertin, T.K., Chen, Y., Hicks, J., Tonachini, L., Monticone, M., Castagnola, P., Rauch, F., Glorieux, F.H., Vranka, J., *et al.* (2006). CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. Cell *127*, 291-304.

Mrosk, J., Bhavani, G.S., Shah, H., Hecht, J., Krüger, U., Shukla, A., Kornak, U., and Girisha, K.M. (2018). Diagnostic strategies and genotype-phenotype correlation in a large Indian cohort of osteogenesis imperfecta. Bone *110*, 368-377.

Myllyharju, J., and Kivirikko, K.I. (2004). Collagens, modifying enzymes and their mutations in humans, flies and worms. Trends Genet *20*, 33-43.

Naslavsky, M.S., Yamamoto, G.L., de Almeida, T.F., Ezquina, S.A.M., Sunaga, D.Y., Pho, N., Bozoklian, D., Sandberg, T.O.M., Brito, L.A., Lazar, M., *et al.* (2017). Exomic variants of an elderly cohort of Brazilians in the ABraOM database. Hum Mutat *38*, 751-763.

Nuytinck, L., Wettinck, K., Freund, M., Van Maldergem, L., Fabry, G., and De Paepe, A. (1997). Osteogenesis imperfecta phenotypes resulting from serine for glycine substitutions in the alpha2(I) collagen chain. Eur J Hum Genet *5*, 161-167.

Palomo, T., Al-Jallad, H., Moffatt, P., Glorieux, F.H., Lentle, B., Roschger, P., Klaushofer, K., and Rauch, F. (2014). Skeletal characteristics associated with homozygous and heterozygous WNT1 mutations. Bone *67*, 63-70.

Palomo, T., Fassier, F., Ouellet, J., Sato, A., Montpetit, K., Glorieux, F.H., and Rauch, F. (2015). Intravenous Bisphosphonate Therapy of Young Children With Osteogenesis Imperfecta: Skeletal Findings During Follow Up Throughout the Growing Years. J Bone Miner Res *30*, 2150-2157.

Palomo, T., Vilaça, T., and Lazaretti-Castro, M. (2017). Osteogenesis imperfecta: diagnosis and treatment. Curr Opin Endocrinol Diabetes Obes *24*, 381-388.

Patel, R.M., Nagamani, S.C., Cuthbertson, D., Campeau, P.M., Krischer, J.P., Shapiro, J.R., Steiner, R.D., Smith, P.A., Bober, M.B., Byers, P.H., *et al.* (2015). A cross-sectional multicenter study of osteogenesis imperfecta in North America - results from the linked clinical research centers. Clin Genet *87*, 133-140.

Puig-Hervas, M.T., Temtamy, S., Aglan, M., Valencia, M., Martinez-Glez, V., Ballesta-Martinez, M.J., Lopez-Gonzalez, V., Ashour, A.M., Amr, K., Pulido, V., *et al.* (2012). Mutations in PLOD2 cause autosomal-recessive connective tissue disorders within the Bruck syndrome--osteogenesis imperfecta phenotypic spectrum. Hum Mutat *33*, 1444-1449.

Rauch, F., and Glorieux, F.H. (2004). Osteogenesis imperfecta. Lancet *363*, 1377-1385.

Rauch, F., Lalic, L., Roughley, P., and Glorieux, F.H. (2010). Relationship between genotype and skeletal phenotype in children and adolescents with osteogenesis imperfecta. J Bone Miner Res *25*, 1367-1374.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., *et al.* (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med *17*, 405-424.

Rubinato, E., Morgan, A., D'Eustacchio, A., Pecile, V., Gortani, G., Gasparini, P., and Faletra, F. (2014). A novel deletion mutation involving TMEM38B in a patient with autosomal recessive osteogenesis imperfecta. Gene *545*, 290-292.

Santos, S., Kok, F., Weller, M., de Paiva, F.R., and Otto, P.A. (2010). Inbreeding levels in Northeast Brazil: Strategies for the prospecting of new genetic disorders. Genet Mol Biol *33*, 220-223.

Semler, O., Garbes, L., Keupp, K., Swan, D., Zimmermann, K., Becker, J., Iden, S., Wirth, B., Eysel, P., Koerber, F., *et al.* (2012). A mutation in the 5'-UTR of IFITM5 creates an in-frame start codon and causes autosomal-dominant osteogenesis imperfecta type V with hyperplastic callus. Am J Hum Genet *91*, 349-357.

Shaheen, R., Alazami, A.M., Alshammari, M.J., Faqeih, E., Alhashmi, N., Mousa, N., Alsinani, A., Ansari, S., Alzahrani, F., Al-Owain, M., *et al.* (2012). Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by TMEM38B mutation. J Med Genet *49*, 630-635.

Sillence, D.O., Senn, A., and Danks, D.M. (1979). Genetic heterogeneity in osteogenesis imperfecta. J Med Genet *16*, 101-116.

Sinder, B.P., Lloyd, W.R., Salemi, J.D., Marini, J.C., Caird, M.S., Morris, M.D., and Kozloff, K.M. (2016). Effect of anti-sclerostin therapy and osteogenesis imperfecta on tissue-level properties in growing and adult mice while controlling for tissue age. Bone *84*, 222-229.

Symoens, S., Malfait, F., D'Hondt, S., Callewaert, B., Dheedene, A., Steyaert, W., Bachinger, H.P., De Paepe, A., Kayserili, H., and Coucke, P.J. (2013). Deficiency for the ER-stress transducer OASIS causes severe recessive osteogenesis imperfecta in humans. Orphanet J Rare Dis *8*, 154.

Tucker, T., Marra, M., and Friedman, J.M. (2009). Massively parallel sequencing: the next big thing in genetic medicine. Am J Hum Genet *85*, 142-154.

van Dijk, F.S., Nesbitt, I.M., Zwikstra, E.H., Nikkels, P.G., Piersma, S.R., Fratantoni, S.A., Jimenez, C.R., Huizer, M., Morsman, A.C., Cobben, J.M., *et al.* (2009). PPIB mutations cause severe osteogenesis imperfecta. Am J Hum Genet *85*, 521-527.

Van Dijk, F.S., Pals, G., Van Rijn, R.R., Nikkels, P.G., and Cobben, J.M. (2010). Classification of Osteogenesis Imperfecta revisited. Eur J Med Genet *53*, 1-5.

Van Dijk, F.S., and Sillence, D.O. (2014). Osteogenesis imperfecta: clinical diagnosis, nomenclature and severity assessment. Am J Med Genet A *164A*, 1470-1481.

Warman, M.L., Cormier-Daire, V., Hall, C., Krakow, D., Lachman, R., LeMerrer, M., Mortier, G., Mundlos, S., Nishimura, G., Rimoin, D.L., *et al.* (2011). Nosology and classification of genetic skeletal disorders: 2010 revision. Am J Med Genet A *155A*, 943-968.

Wenstrup, R.J., Willing, M.C., Starman, B.J., and Byers, P.H. (1990). Distinct biochemical phenotypes predict clinical severity in nonlethal variants of osteogenesis imperfecta. Am J Hum Genet *46*, 975-982.

Willing, M.C., Deschenes, S.P., Slayton, R.L., and Roberts, E.J. (1996). Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type I cell strains. Am J Hum Genet *59*, 799-809.

Apêndices

Apêndice 1. Dados retrospectivos do tratamento medicamentoso com bisfosfonatos

Apêndice 2. Identificação das variantes encontradas ao nível do DNA e da proteína

Apêndice 1. Dados retrospectivos do tratamento medicamentoso com bisfosfonatos

Pt	Tratamento com Pamidronato		Tratamento com Ácido Zoledrônico		Tratamento com Alendronato	
	Idade ao início	*Idade ao fim*	*Idade ao início*	*Idade ao fim*	*Idade ao início*	*Idade ao fim*
1	2 anos	8 anos	11 anos	15 anos	NR	NR
2	8 anos	18 anos	NR	NR	NR	NR
3	3 anos	17 anos	1º) 22 anos 2º) 28 anos	1º) 23 anos 2º) 29 anos	NR	NR
4	3 anos	16 anos	NR	NR	NR	NR
5	8 anos	20 anos	NR	NR	NR	NR
6	8 anos	18 anos	19 anos	23 anos	NR	NR
7	8 anos	17 anos	NR	NR	NR	NR
8	1º) 16 anos 2º) 30 anos	1º) 21 anos 2º) 30 anos	NR	NR	NR	NR
9	5 anos	17 anos	NR	NR	NR	NR
10	8 anos	17 anos	NR	NR	17 anos	19 anos
11	8 anos	15 anos	NR	NR	15 anos	18 anos
12	9 anos	16 anos	NR	NR	NR	NR
13	15 anos	20 anos	NR	NR	NR	NR
14	Incerto	Incerto	NR	NR	NR	NR
15	16 anos	16 anos	NR	NR	NR	NR
16	4 anos	14 anos	NR	NR	NR	NR
17	9 anos	16 anos	NR	NR	NR	NR
18	1º) 3 anos 2º) 14 anos	1º) 10 anos 2º) 16 anos	NR	NR	NR	NR
19	7 anos	19 anos	21 anos	22 anos	NR	NR
20	11 anos	15 anos	NR	NR	NR	NR

Pt	Tratamento com Pamidronato		Tratamento com Ácido Zoledrônico		Tratamento com Alendronato	
	Idade ao início	*Idade ao fim*	*Idade ao início*	*Idade ao fim*	*Idade ao início*	*Idade ao fim*
21	1º) 8 meses 2º) 10 anos	1º) 8 anos 2º) 15 anos	NR	NR	NR	NR
22	15 anos	16 anos	NR	NR	8 anos	11 anos
23	34 anos	34 anos	NR	NR	NR	NR
24	NR	NR	55 anos	58 anos	NR	NR
25	NR	NR	NR	NR	NR	NR
26	NR	NR	NR	NR	NR	NR
27	NR	NR	NR	NR	NR	NR
28	10 anos	18 anos	NR	NR	NR	NR
29	2 anos	14 anos	NR	NR	NR	NR
30	NR	NR	NR	NR	53 anos	53 anos
F1a	NR	NR	NR	NR	NR	NR
F1b	NR	NR	NR	NR	NR	NR
F2a	NR	NR	NR	NR	41 anos	46 anos
F2b	NR	NR	NR	NR	24 anos	28 anos
F3a	NR	NR	NR	NR	12 anos	18 anos
F3b	NR	NR	NR	NR	NR	NR
F3c	NR	NR	NR	NR	66 anos	69 anos
F3d	NR	NR	NR	NR	45 anos	46 anos
F4a	12 anos	16 anos	NR	NR	NR	NR
F4b	NR	NR	53 anos	54 anos	42 anos	51 anos
F5a	1 ano	11 anos	NR	NR	NR	NR
F5b	NR	NR	NR	NR	47 anos	50 anos
F6a	6 a	15 a	NR	NR	NR	NR

Pt	Tratamento com Pamidronato		Tratamento com Ácido Zoledrônico		Tratamento com Alendronato	
	Idade ao início	*Idade ao fim*	*Idade ao início*	*Idade ao fim*	*Idade ao início*	*Idade ao fim*
F6b	NR	NR	NR	NR	46 anos	48 anos
F7a	1 ano	14 anos	NR	NR	NR	NR
F7b	NR	NR	NR	NR	NR	NR
F8a	NR	NR	NR	NR	NR	NR
F8b	NR	NR	NR	NR	25 anos	28 anos
F8c	NR	NR	NR	NR	NR	NR

NR, não realizado

Apêndice 2. Identificação das variantes encontradas ao nível do DNA e da proteína

Pt	Gene	Variante (DNA)	Variante (Proteína)
1	*SERPINF1*	c.1145_1163del	p.(Phe384Leufs*9)(;)(Phe384Leufs*9)
2	*COL1A1*	c.A3839C	p.(Gln1280Pro)
3	*FKBP10*	c.825dupC	p.(Gly278Argfs*95)(;)(Gly278Argfs*95)
4	*COL1A1*	c.G2569T	p.(Gly857Cys)
5	*SERPINF1*	c.[283+2T>C];[(283+2T>C)]	
6	*PLOD2* *PLOD2*	c.1682G>A c.1496_1500del	p.(Trp561*) p.(Glu499Aspfs*29)
7	*COL1A1*	c.643-2A>G	
8	*P3H1*	c.[1080+1G>T];[1080+1G>T]	
9	*COL1A1* *COL1A1*	c.G878A c.C873A	p.(Gly293Asp) p.(Ser291Arg)
10	*COL1A2*	c.G1100A	p.(Gly367Glu)
11	*COL1A2*	c.G982A	p.(Gly328Ser)
12	*FKBP10*	c.A179C	p.(Gln60Pro)(;)(Gln60Pro)
13	*COL1A2*	c.G2314A	p.(Gly772Ser)
14	*COL1A1*	c.G2155T	p.(Gly719Cys)
16	*TMEM38B*	deleção de éxons 1 e 2	
17	*P3H1* *P3H1* *WNT1* *WNT1*	c.G1448A c.349_358del c.A308C c.775_786dup	p.(Ala117Cysfs*64) p.(Cys483Tyr) p.(Asn103Thr) p.(Gly259_Gly262dup)
18	*COL1A1*	c.G3118A	p.(Gly1040Ser)
19	*COL1A1*	c.G2299A	p.(Gly767Ser)
20	*PLOD2* *PLOD2*	c.1639T>C c.844T>C	p.(Tyr547His) p.(Cys282Arg)
21	*COL1A2*	c.G578A	p.(Gly193Asp)
22	*COL1A1*	c.C748T	p.(Gln250*)
23	*IFITM5*	c. -14 C->T	
24	*COL1A1*	c.2612delC	p.(Pro871Leufs*237)
25	*COL1A2*	c.G2314A	p.(Gly772Ser)
26	*COL1A1*	c.G316T	p.(Glu106*)
27	*COL1A1*	c.2452-2A>G	
28	*COL1A1*	c.G2362A	p.(Gly788Ser)
29	*COL1A2*	c.G1378A	p.(Gly460Ser)
30	*COL1A1*	c.334-9A>G	
F1	*COL1A1*	c.334-9A>G	
F2	*COL1A2*	c.G955A	p.(Gly319Arg)
F3	*COL1A2*	c.G2314A	p.(Gly772Ser)
F4	*COL1A2* *COL1A2*	c.G686A c.C689A	p.(Gly229Asp) p.(Pro230Gln)
F5	*COL1A1*	c.G1580C	p.(Gly527Ala)
F6	*COL1A1* *COL1A1*	c.C3421T c.G1583A	p.(Arg1141*) p.(Arg528His)
F7	*COL1A1*	c.809dupT	p.(Ser271Glnfs*16)
F8	*COL1A1*	c.G769A	p.(Gly257Arg)