

Universidade Federal do Rio Grande do Sul
Instituto de Biociências
Bacharelado em Biotecnologia - Ênfase em Biotecnologia Molecular

EDUARDA HEIDRICH PEZZI

**BIOLOGIA DE SISTEMAS APLICADA A POSSÍVEIS NOVAS ASSOCIAÇÕES DE
MÚLTIPLOS GENES NA OSTEOGÊNESE IMPERFEITA**

Porto Alegre

2021

EDUARDA HEIDRICH PEZZI

**BIOLOGIA DE SISTEMAS APLICADA A POSSÍVEIS NOVAS ASSOCIAÇÕES DE
MÚLTIPLOS GENES NA OSTEOGÊNESE IMPERFEITA**

Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do título de bacharel em
Biotecnologia - Ênfase em Biotecnologia
Molecular do Instituto de Biociências da
Universidade Federal do Rio Grande do Sul.

Orientadora: Professora Dra. Têmis Maria
Félix

Co-orientadora: Dra. Thayne Woycinck
Kowalski

Porto Alegre

2021

FOLHA DE APROVAÇÃO

EDUARDA HEIDRICH PEZZI

BIOLOGIA DE SISTEMAS APLICADA A POSSÍVEIS NOVAS ASSOCIAÇÕES DE MÚLTIPLOS GENES NA OSTEOGÊNESE IMPERFEITA

Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do título de bacharel em
Biotecnologia - Ênfase em Biotecnologia
Molecular do Instituto de Biociências da
Universidade Federal do Rio Grande do Sul.
Orientadora: Professora Dra. Têmis Maria
Félix

Co-orientadora: Dra. Thayne Woycinck
Kowalski

Aprovada em: Porto Alegre, 24 de maio de 2021.

BANCA EXAMINADORA:

Dra. Têmis Maria Félix

Hospital de Clínicas de Porto Alegre

Dra. Sandra Leistner-Segal

Hospital de Clínicas de Porto Alegre

Dra. Liliame Todeschini de Souza
Hospital de Clínicas de Porto Alegre

AGRADECIMENTOS

Aos meus pais por terem me dado todo o apoio e amor ao longo da minha trajetória.

À minha irmã por ser a minha grande parceira e amiga, me incentivando nos piores e melhores momentos.

Aos meus colegas de barra e minhas melhores amigas, Luiza e Catherine, pelas risadas e os momentos de descontração ao longo do curso.

Ao pessoal do Laboratório de Medicina Genômica que muito me ensinou e me deu oportunidades. Um agradecimento em especial ao Igor que além de me orientar se tornou um grande amigo.

À minha orientadora, Têmis, por ter aceitado me orientar e me guiou nessa jornada.

À minha coorientadora, Thayne, por ter toda a paciência em me explicar e ensinar sobre biologia de sistemas.

E à Deus e meu anjo da guarda que cuidaram de mim e me deram forças nesta caminhada.

Muito obrigada!

RESUMO

Osteogênese imperfeita (OI) é um grupo de doenças ósseas com heterogeneidade fenotípica e genética. Atualmente, apresenta uma classificação de 21 tipos estando associada à diferentes genes. Acredita-se que ainda existam outros genes e variantes que podem estar associadas a OI. Assim, o objetivo do presente estudo foi buscar associações de múltiplos genes com OI. Foram utilizadas análises de interação de rede, expressão e co-expressão gênica, métodos de bioinformática aplicados a biologia de sistemas para processar os dados. A primeira abordagem realizada com genes relacionados aos principais fenótipos de OI, conforme dados disponíveis no HPO, acabou gerando uma rede de interação muito grande e pouco conclusiva. Seguimos para uma segunda abordagem, em que analisamos a co-expressão gênica de osteoblastos comparada a células mesenquimais. Ao analisar as vias de GO e KEGG dos genes co-expressos já associados à OI se destacaram ontologias de ossificação e vias de adesão focal. Por fim, verificamos a expressão diferencial dos genes em ossos de camundongo. Ao cruzarmos a lista dos genes diferencialmente expressos com os genes relacionados a ossificação e adesão focal encontramos 13 genes presentes em ambos, *COL1A1*, *COL1A2*, *CREB3L1*, *MMP2*, *FBN2*, *IHH*, *COL11A2*, *COL2A1*, *COL11A1*, *SOX9*, *SATB2*, *PHEX* e *GLI2*. Os três primeiros genes já são associados à OI, no entanto, os demais seriam possíveis candidatos a serem associados, sendo necessário estudos *in vivo* e funcionais para averiguar a hipótese.

Palavras-chave: Osteogênese imperfeita, genes, biologia de sistemas.

ABSTRACT

Osteogenesis imperfecta (OI) is a group of bone disorders with phenotypic and genetic heterogeneity. Currently, 21 types of OI is classified and associated with different genes. It has been suggested that there are other genes and variants that may be associated with OI. Thus, the objective of the present study was to search for associations of multiple genes with OI. To generate the data, network interaction, gene expression and co-expression, bioinformatics techniques applied to systems biology were performed. The first approach was carried out using related genes of the main OI phenotypes and a very large and inconclusive interaction network was generated. We move on to a second approach where we analyzed the gene co-expression of osteoblasts and mesenchymal cells. When analyzing the GO and KEGG pathways of the co-expressed genes previously associated with OI, ossification and focal adhesion stood out. Finally, we verified the differential expression of genes in mouse bones. Crossing the list of differentially expressed genes with genes related to ossification and focal adhesion, we found 13 genes present in both, *COL1A1*, *COL1A2*, *CREB3L1*, *MMP2*, *FBN2*, *IHH*, *COL11A2*, *COL2A1*, *COL11A1*, *SOX9*, *SATB2*, *PHEX* and *GLI2*. The first three genes are already associated with OI, however, the others would be possible candidates to be associated, requiring *in vivo* and functional studies to ascertain the hypothesis.

Keywords: Osteogenesis imperfecta, gene, systems biology.

LISTA DE ABREVIATURAS E SIGLAS

OI	Osteogênese Imperfeita
DI	Dentinogênese imperfeita
AD	Autossômico dominante
AR	Autossômico recessivo
LX	Ligado ao X
GO	Ontologia gênica
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
HPO	<i>Human Phenotype Ontology</i>
GEO	<i>Gene Expression Omnibus</i>
PPI	Interações proteína-proteína
FDR	<i>False discovery rate</i>
NCBI	<i>National Center for Biotechnology Information</i>
TMM	<i>Trimmed means</i>
RNAseq	Sequenciamento de RNA
ECM	Matriz extracelular

SUMÁRIO

1 INTRODUÇÃO	8
1.1 OSTEOGÊNESE IMPERFEITA	8
1.2 GENÉTICA DA OSTEOGÊNESE IMPERFEITA	9
1.3 COLÁGENO TIPO 1 E OSTEOGÊNESE IMPERFEITA.....	10
1.4 OSTEOBLASTOS E OSTEOGÊNESE IMPERFEITA	11
1.5 MINERALIZAÇÃO E METABOLISMO ÓSSEO E OSTEOGÊNESE IMPERFEITA	12
1.6 MANIFESTAÇÕES CLÍNICAS DE OSTEOGÊNESE IMPERFEITA	12
1.7 BIOLOGIA DE SISTEMAS	13
2 JUSTIFICATIVA	14
3 OBJETIVO	15
3.1 OBJETIVO GERAL	15
3.2 OBJETIVOS ESPECÍFICOS	15
4 METODOLOGIA	16
4.1 SELEÇÃO DE FENÓTIPOS DA OSTEOGÊNESE IMPERFEITA.....	16
4.2 ANÁLISE DE INTERAÇÃO DE REDE	16
4.3 ANÁLISES GENÓTIPO-FENÓTIPO	16
4.4 ANÁLISE DE ONTOLOGIAS GÊNICAS.....	16
4.5 ANÁLISE DE CO-EXPRESSÃO DE RNASEQ.....	17
4.6 ANÁLISES DE EXPRESSÃO DIFERENCIAL	17
4.7 DIAGRAMA DE VENN	17
5 RESULTADOS	18
5.1 ABORDAGEM A PARTIR DO FENÓTIPO	18
5.2 ABORDAGEM A PARTIR DE OSTEOBLASTOS.....	21
5.3 VERIFICAÇÃO DE EXPRESSÃO DIFERENCIAL	24
6 DISCUSSÃO	27
REFERÊNCIAS	31

1 INTRODUÇÃO

1.1 OSTEOGÊNESE IMPERFEITA

Osteogênese imperfeita (OI) é um grupo de doenças ósseas com heterogeneidade fenotípica e genética. É caracterizada pela diminuição da densidade óssea devido à defeitos na síntese do colágeno tipo 1, tendo como genes codificantes principais o *COL1A1* e *COL1A2* (Marini et al., 2007). Nos indivíduos com OI a baixa massa óssea resulta em fragilidade óssea, fraturas de repetição, baixa estatura e progressiva deformidade óssea, principalmente nos ossos longos. Outras manifestações clínicas extra-esqueléticas observadas incluem escleróticas acinzentadas ou azuladas, dentinogênese imperfeita (DI), frouxidão ligamentar e surdez progressiva (Engelbert et al., 2004).

OI é uma doença rara com uma prevalência geral de 1:10.000 à 1:20.000 nascimentos (Roughley et al., 2003). No entanto, a prevalência pode ser muito maior, uma vez que os indivíduos com a forma mais branda da doença, por não procurarem ajuda médica, acabam não sendo diagnosticados (Martin et al., 2007; Pillion et al., 2011). A Associação Brasileira de Osteogênese estima que existam mais de 12.000 indivíduos com esta patologia no Brasil (Fiocruz, 2017).

Inicialmente, a OI foi classificada clínica e radiologicamente em quatro tipos (I-IV) pelo médico australiano David Sillence em 1979. OI tipo I é a forma mais leve, o tipo II é a forma letal perinatal, enquanto o tipo III é o mais grave e a OI tipo IV é caracterizada pela forma moderada (Sillence et al., 1979 ; Rauch et al., 2010 ; Lin et al., 2015 ; Mrosk et al., 2018). Com o aumento dos estudos e da compreensão da doença, mais genes foram sendo relacionados à OI, levando a um aumento na classificação genética para 21 tipos conforme descrito na Tabela 1 (Mortier et al., 2019; OMIM).

Tabela 1: Classificação de OI de acordo com a base molecular.

OI Tipo Classificação Genética	Herança	Gene	Gravidade
I	AD	<i>COL1A1, COL1A2</i>	Leve
II	AD	<i>COL1A1, COL1A2</i>	Letal
III	AD	<i>COL1A1, COL1A2</i>	Grave

IV	AD	<i>COL1A1, COL1A2</i>	Moderada
V	AD	<i>IFITM5</i>	Moderada a grave
VI	AR	<i>SERPINF1</i>	Grave
VII	AR	<i>CRTAP</i>	Grave a letal
VIII	AR	<i>P3H1</i>	Grave a letal
IX	AR	<i>PPIB</i>	Moderada a letal
X	AR	<i>SERPINH1</i>	Grave
XI	AR	<i>FKBP10</i>	Moderada a grave
XII	AR	<i>SP7</i>	Moderada
XIII	AR	<i>TMEM38B</i>	Grave
XIV	AR	<i>BMP1</i>	Grave
XV	AR	<i>WNT1</i>	Moderada a grave
	AD (osteoporose)		
XVI	AR	<i>CREB3L1</i>	Grave
XVII	AR	<i>SPARC</i>	Grave
XVIII	AR	<i>FAM46A</i>	Grave
XIX	LX	<i>PLS3</i>	Leve
XX	AR	<i>MESD</i>	Grave a letal
XXI	AR	<i>KDEL2</i>	Grave

AD: autossômico dominante; AR: autossômico recessivo; LX: ligado ao X. Adaptado de Mortier et al., 2019 e de dados coletados do site OMIM (<https://www.omim.org/>).

1.2 GENÉTICA DA OSTEOGÊNESE IMPERFEITA

Os 4 primeiros tipos da classificação, que representam a grande maioria dos casos de OI, são associados a variações nos genes *COL1A1* e *COL1A2*. Entretanto, alterações nos genes do colágeno não foram identificadas em diversos casos com características clínicas de OI, principalmente formas graves ou letais. Nas últimas décadas, com o aumento dos estudos na doença, foram associadas diferentes variantes patogênicas de genes não colágenos, proteínas codificadoras envolvidas na biossíntese de colágeno ou fatores de transcrição e moléculas de sinalização relacionadas à diferenciação e mineralização de células ósseas (Bregou Bourgeois et al., 2016 ; Lindert et al., 2016 ; Marom et al., 2016 ; Marini et al., 2017).

Dentre as descobertas realizadas, foi visto que as variantes patogênicas nos genes *COL1A1*, *COL1A2* e *IFITM5* apresentam herança autossômica dominante. Enquanto nos genes *SERPINF1*, *CRTAP*, *P3H1* (*LEPRE1*), *PPIB*, *SERPINH1*, *FKBP10*, *SP7*, *BMP1*, *TMEM38B*, *WNT1*, *CREB3L1*, *SPARC*, *FAM46A* (*TENT5A*), *MESD* e *KDEL2* a herança é autossômica recessiva. Já o gene *PLS3* é associado a herança ligada ao cromossomo X, Tabela 1 (Mortier et al., 2019; OMIM).

Esses genes já classificados geneticamente com tipos de OI estão correlacionados com a diminuição da massa óssea por alterações na formação e estruturação do colágeno tipo 1, função e diferenciação de osteoblastos e na mineralização e metabolismo ósseo (Figura 1) (Marini et al., 2017). Acredita-se que novos genes e novas variantes patogênicas podem vir a ser identificadas e relacionadas à OI.

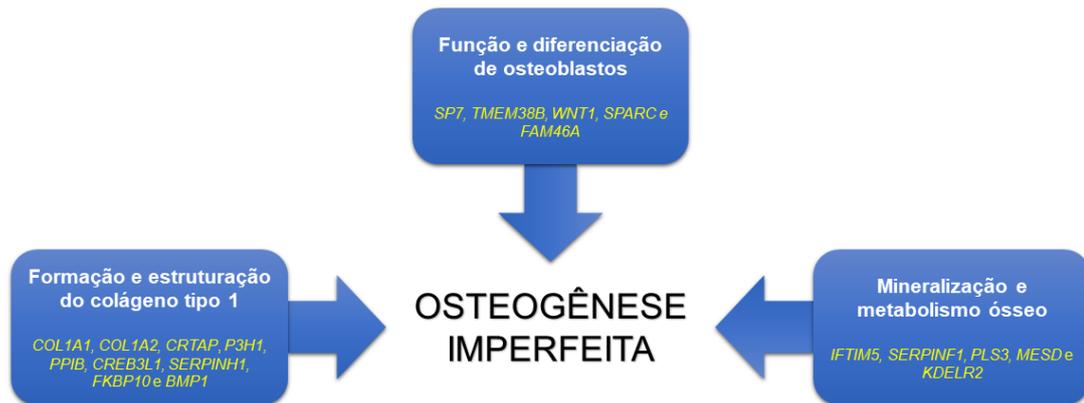


Figura 1: Mecanismo alterados dos genes associado a Osteogênese Imperfeita

1.3 COLÁGENO TIPO 1 E OSTEOGÊNESE IMPERFEITA

O colágeno tipo I é o principal constituinte da matriz extracelular óssea humana, ditando a resistência óssea a partir de sua quantidade e integridade (Eyre and Weis, 2013). Esse é composto por duas cadeias de procolágeno $\alpha 1$ e uma cadeia $\alpha 2$, formando uma molécula de tripla hélice. Diferente da maioria das proteínas, os procolágenos são montados a partir da região C-terminal e não da região N-terminal. Após o processo de tradução, essas ainda sofrem diversas modificações até serem incorporadas na matriz extracelular óssea. Além disso, eles são caracterizados por sequência de aminoácidos repetidas, cada cadeia contém Gly-XY, sendo geralmente prolina e hidroxiprolina nas posições X e Y (Marini et al, 2010; Makareeva et al., 2011).

Aproximadamente 85% dos casos de OI são causadas por variações patogênicas nos genes *COL1A1* e *COL1A2* que codificam as cadeias de procolágenos $\alpha 1$ e $\alpha 2$, respectivamente. Variantes que alteram a estrutura do colágeno tipo I, grande maioria substituições de glicina por outro aminoácido, levam a forma moderada, grave ou letal de OI e são consideradas alterações qualitativas. Enquanto, as alterações quantitativas, causadas por haploinsuficiência, em geral, por códon de parada

prematurados nos genes levando a uma menor expressão das cadeias α , são associadas à quadros clínicos leves da doença (Marini et al., 2010; Willing et al., 1996).

A formação da estrutura em tripla hélice pós-traducional do colágeno tipo I tem como complexo proteico responsável a colágeno 3-prolil-hidroxilase, codificada pelos genes *CRTAP*, *P3H1* e *PP1B*. Variantes patogênicas nestes genes são associadas às formas graves em OI (Morello et al., 2006; Cabral et al., 2007; van Dijk et al., 2009). Também responsável pela formação da tripla hélice, a isomerização da prolina é realizada pela proteínas codificadas dos genes *SERPINH1* e *FKBP10*, e suas alterações levam a quadros clínicos moderados a graves (Alanay et al., 2010; Christiansen et al., 2010). Outra associação de fenótipo grave a OI são variações patogênicas no gene *BMP1*, codificante de uma protease que atua na clivagem do propeptídeo C-terminal do procolágeno tipo 1 (Martinez-Glez et al., 2012). Ainda, o gene *CREB3L1*, que atua na ativação da expressão do gene *COL1A1*, também apresenta alterações que foram relacionadas com OI (Symoens et al., 2013).

1.4 OSTEOLASTOS E OSTEOGÊNESE IMPERFEITA

O osso está em constante estado de renovação, uma das principais células nesse processo são os osteoblastos, que atuam na formação óssea (Matsuo and Irie, 2008). Existem alguns genes que estão relacionados a diferenciação na linhagem dos osteoblastos como *SP7*, *TMEM38B* e *WNT1*. Estes apresentam variantes patogênicas que são associadas a OI, e o último também está relacionado com osteoporose (Lapunzina et al., 2010; Shaheen et al., 2012; Keupp et al., 2013; Laine et al., 2013; Palomo et al., 2014).

O gene *SPARC*, que é expresso nas células de osteoblastos, codifica a proteína SPARC que se liga ao colágeno tipo 1 e outras proteínas da matriz. Duas alterações nesse gene já foram associadas a forma grave de OI (Mendoza-Londono et al., 2015). Outro gene que se manifesta nessas células é o *FAM46A*, que ainda tem a sua função indefinida, mas a família FAM46 tem sido relacionada a diferentes tipos de doença, incluindo este a OI (Doyard et al., 2018).

1.5 MINERALIZAÇÃO E METABOLISMO ÓSSEO E OSTEOGÊNESE IMPERFEITA

Diversos genes controlam e atuam na mineralização óssea. Um deles, *IFTIM5*, apresenta uma alteração na região 5' UTR (c.-14C>T) que cria um novo sítio de iniciação levando a adição de 5 aminoácidos na porção N-terminal do gene foi identificada como responsável pela OI tipo V (Semler et al., 2012). Já o tipo VI foi associado a variantes patogênicas no gene *SERPINF1*, que leva a uma deficiência de mineralização óssea (Becker et al., 2011; Homan et al., 2011).

O gene *PLS3*, localizado no cromossomo X, é um regulador do metabolismo ósseo, e não só está associado com OI como com osteoporose. No entanto, suas vias de regulação ainda não são bem esclarecidas (van Dijk et al., 2013; Cao et al., 2014).

A associação do gene *KDELR2* apresenta como hipótese que a proteína codificada por este interfere na formação de fibras (van Dijk et al., 2020). *MESD* atua na via WNT e alterações nesta podem afetar negativamente o acúmulo de massa óssea e a resistência óssea (Moosa et al., 2019)

1.6 MANIFESTAÇÕES CLÍNICAS DE OSTEOGÊNESE IMPERFEITA

A OI é caracterizada por fragilidade e repetidas fraturas ósseas, deformidade esquelética e baixa estatura (Engelbert et al., 2004). Com a malformação na coluna vertebral, ossos longos e caixa torácica reduzida indivíduos com OI tendem a ter problemas nas funções pulmonares (LoMauro et al., 2018). Manifestações extra-esqueléticas também podem ser desenvolvidas, uma vez que OI é uma desordem sistêmica do tecido conjuntivo, incluindo escleróticas acinzentadas ou azuladas, DI, fraqueza muscular, complicações cardiovasculares, frouxidão ligamentar e surdez progressiva (Engelbert et al., 2004).

No tipo mais leve da doença, tipo I, as fraturas ósseas tendem a diminuir com o avanço da idade (Tauer et al., 2019). A baixa estatura é comum nos pacientes diagnosticados com a classificação clínica do tipo I, III e IV (Barber et al., 2019; Jain et al., 2019). Além disso, no tipo III as complicações respiratórias, fragilidade e a má oclusão dos dentes são mais graves (LoMauro et al., 2018; Najirad et al.; 2018, Tam et al., 2018). O tipo II acaba sendo letal tanto no período pré-natal quanto neonatal (Sillence et al., 1984).

1.7 BIOLOGIA DE SISTEMAS

O crescente aumento de bancos de dados genômicos públicos possibilita a realização de uma série de análises de bioinformática, sendo possível estudos com cenários mais abrangentes para avaliar a complexidade biológica de um organismo. A biologia de sistemas, por sua vez, combina experimentos de biologia molecular e análise computacional, buscando entender as interações mais simples de complexos biológicos e como essas funcionam no sistema (Le Novère, 2015; Kowalski et al., 2019). Quando integrado informações de genômica, proteômica e fenotípica, é possível avaliar como cada um desses atua e como suas perturbações podem modificar o fenótipo da interação (Sahni et al., 2013).

Na biologia de sistemas se tem como principal análise as redes biológicas, integrando os processos biológicos de acordo com a abordagem do estudo (Le Novère, 2015). Para auxiliar na montagem de um cenário mais completo existe uma série de análises que podem ser utilizadas. As ontologias gênicas (GO) e as vias de sinalização enriquecidas (KEGG) podem indicar como os genes se organizam em vias metabólicas (Ashburner et al., 2000; Kanehisa et al., 2017). A co-expressão diferencial, por sua vez, avaliará se a alteração em um gene repercute em outros (Yang et al., 2013). Enquanto, a expressão diferencial compara a expressão do gene em diferentes ambientes (Robinson et al., 2010).

Essa metodologia consiste em uma abordagem muito eficaz e econômica para a exploração dos mecanismos moleculares de doenças, principalmente doenças raras, pois pode ser aplicada com poucas informações obtidas em estudos.

2 JUSTIFICATIVA

Osteogênese Imperfeita é um grupo heterogêneo de doenças ósseas causadas por diferentes genes. Até o momento 21 genes foram relacionados a OI, sendo estes relacionados a estrutura e síntese do colágeno tipo 1, modificação, processamento e cross-link do colágeno, mineralização óssea ou diferenciação e função dos osteoblastos.

Considerando as recentes descobertas de múltiplos genes associados e o aumento dos estudos em OI, juntamente com o avanço da tecnologia, diversas associações ainda podem ser identificadas e estudadas, ampliando o conhecimento sobre a interação dos genes causadores de doenças e suas repercussões fenotípicas. A biologia de sistemas, por sua vez, pode auxiliar na identificação de possíveis novos genes a serem associados a OI.

3 OBJETIVO

3.1 OBJETIVO GERAL

Analisar as associações de múltiplos genes com Osteogênese Imperfeita utilizando Biologia de Sistemas.

3.2 OBJETIVOS ESPECÍFICOS

- a) Construir uma rede de interação dos genes associados com os principais fenótipos de OI.
- b) Analisar as principais ontologias genicas (GO) e vias de sinalização (KEGG) enriquecidas de OI.
- c) Analisar co-expressão gênica dos genes associados a OI a partir de amostras de osteoblastos e células mesenquimais de estudos anteriores.
- d) Analisar expressão diferencial de RNA-seq a partir de ossos de camundongo de estudos anteriores.

4 METODOLOGIA

4.1 SELEÇÃO DE FENÓTIPOS DA OSTEOGÊNESE IMPERFEITA

Foram selecionados os principais fenótipos de Osteogênese Imperfeita para obtenção de seus genes associados, dos quais foram coletados do banco de dados *Human Phenotype Ontology* (HPO).

4.2 ANÁLISE DE INTERAÇÃO DE REDE

As redes foram geradas a partir da lista de genes obtida no HPO ou dos genes diferencialmente expressos obtidos através dos estudos do repositório *Gene Expression Omnibus* (GEO), usando o banco de dados STRING v.11, compreendendo interações proteína-proteína (PPI) para *Homo sapiens*. Apenas interações experimentais foram selecionadas, com uma pontuação de interação mínima necessária definida em 0,400 (padrão). As redes montadas foram transferidas para o software Cytoscape v.3.8.0, no qual foram realizadas as estatísticas da rede. A comparação das redes geradas foi analisada com o aplicativo DyNet no Cytoscape v.3.8.0.

4.3 ANÁLISES GENÓTIPO-FENÓTIPO

Foram utilizados os dados do HPO gerados pela seleção de fenótipos descrito na seção 4.1. Todos os dados foram normalizados no software RStudio Desktop v. 1.3.1093 em linguagem R, conforme pacotes disponibilizados pelas plataformas.

4.4 ANÁLISE DE ONTOLOGIAS GÊNICAS

O enriquecimento das ontologias foi calculado com o pacote clusterprofileR (Yu et al., 2012), em linguagem R, utilizando-se dos bancos de dados Gene Ontology (GO) e KEGGdb, para vias de sinalização. Todos os Valores-P obtidos foram ajustados pelo método de *false discovery rate* (FDR). Após o ajuste, todos os Valores-P < 0,05 foram considerados estatisticamente significativos.

4.5 ANÁLISE DE CO-EXPRESSÃO DIFERENCIAL

Utilizando os dados de sequenciamento de RNA (RNAseq) do estudo “*Osteogenesis depends on commissioning of a network of stem cell transcription factors that act as repressors of adipogenesis*” (Rauch et al., 2019), obtido no banco de dados GEO, sob código GSE113253. Foi calculada a co-expressão diferencial dos genes conforme descrito na secção 4.2., com o pacote DCGL (Yang et al., 2013), em linguagem R, a partir da correlação de Pearson, comparando os genes par-a-par. O pacote proporciona qual o gene que está regulando a correlação diferencial dentro do par; esse gene foi chamado de “mandante” nesse trabalho.

4.6 ANÁLISES DE EXPRESSÃO DIFERENCIAL

Após busca no GEO, escolhemos o estudo “*The Osteocyte Transcriptome Is Extensively Dysregulated in Mouse Models of Osteogenesis Imperfecta*” (Zimmerman et. al, 2019), sob código GSE154748 para análise de expressão diferencial. Os arquivos fastq coletados do banco de dados *Trace* (National Center for Biotechnology Information - NCBI) e foram tratados no servidor *useGalaxy* (Afgan et al., 2018), realizando alinhamento com *Bowtie2* e a contagem da expressão gênica com o *featurecounts*. A expressão diferencial foi avaliada em pacote *edgeR* (Robinson et al., 2010), em linguagem R, após normalização por *trimmed means* (TMM). Todos os genes com valores de $\log_{2}FC > 1$ foram considerados superregulados, já todos os genes com $\log_{2}FC < -1$ foram considerados subregulados. Foram considerados resultados estatisticamente significativos todos os genes com valor-P $< 0,05$, após a realização do ajuste pelo método FDR.

4.7 DIAGRAMA DE VENN

Para identificar a coexistência de genes nos dados gerados nas secções 4.4, 4.5 e 4.6 utilizamos o diagrama de Venn.

5 RESULTADOS

5.1 ABORDAGEM A PARTIR DO FENÓTIPO

Selecionamos os 18 fenótipos mais frequentes em pacientes com Osteogênese Imperfeita para a nossas análises, descritos na Tabela 2.

Tabela 2: Fenótipos selecionados e seus códigos HP.

FENOTIPO	HP
Cifoescoliose	HP:0002751
Deficiência auditiva	HP:0000365
Deficiência auditiva condutiva progressiva	HP:0008607
Deficiência auditiva neurossensorial de início na idade adulta	HP:0008615
Densidade mineral óssea reduzida	HP 0004349
Dentinogênese imperfeita	HP:0000703
Esclera azul	HP:0000592
Escoliose	HP:0002650
Escoliose lombar	HP:0004626
Escoliose torácica	HP:0002943
Escoliose toracolombar	HP:0002944
Fratura óssea	HP:0020110
Hipermobilidade articular	HP:0001382
Morfologia anormal da válvula aórtica	HP:0001646
Morfologia anormal da válvula cardíaca	HP:0001654
Ossos de <i>wormianos</i>	HP:0002645
Prolapso da válvula mitral	HP:0001634

A partir dos dados gerados no STRING foi possível construir a rede de interação (Figura 2). O resultado inicial gerou uma rede muito grande e de difícil análise. Portanto, optamos por selecionar na rede somente os genes que apareciam em três ou mais fenótipos. Dos genes já associados a OI apenas 7 estão presentes na rede de interação: *COL1A1*, *COL1A2*, *CRTAP*, *PPIB*, *SP7*, *BMP1* e *WNT1*.

Os fenótipos deficiência auditiva condutiva progressiva, deficiência auditiva neurossensorial de início na idade adulta, escoliose lombar, escoliose toracolombar e fratura óssea não apresentaram dados suficientes para serem incluídos na rede gerada.

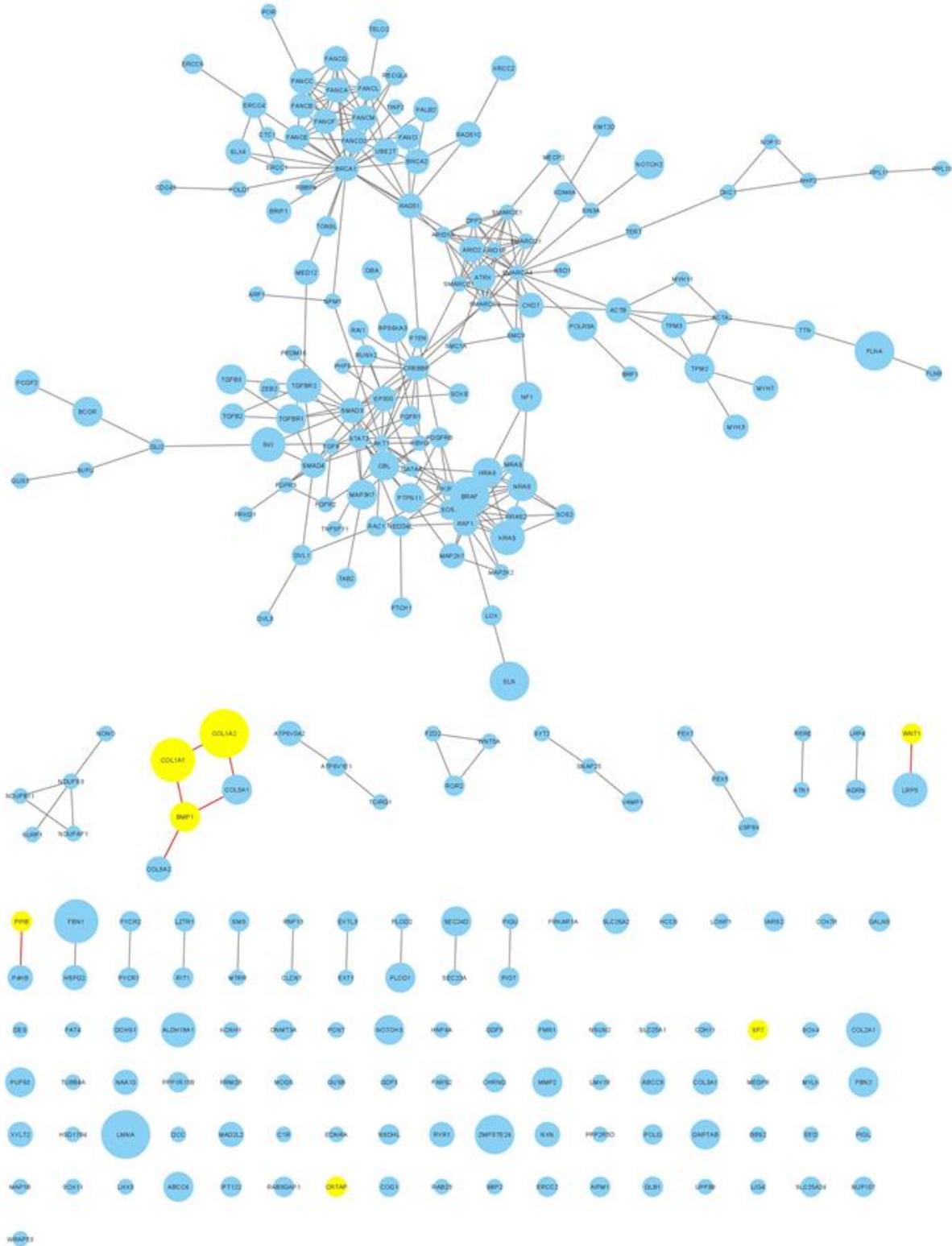


Figura 2: Rede de interação de genes gerada a partir do aplicativo DyNet no Cytoscape v.3.8.0 com a cor amarela aplicada à genes associados à OI.

Com a lista de genes gerados a partir dos fenótipos selecionados também foi pesquisado as ontologias gênicas (GO) e vias de KEGG. A Figura 3 mostra as 20 mais enriquecidas desses genes.

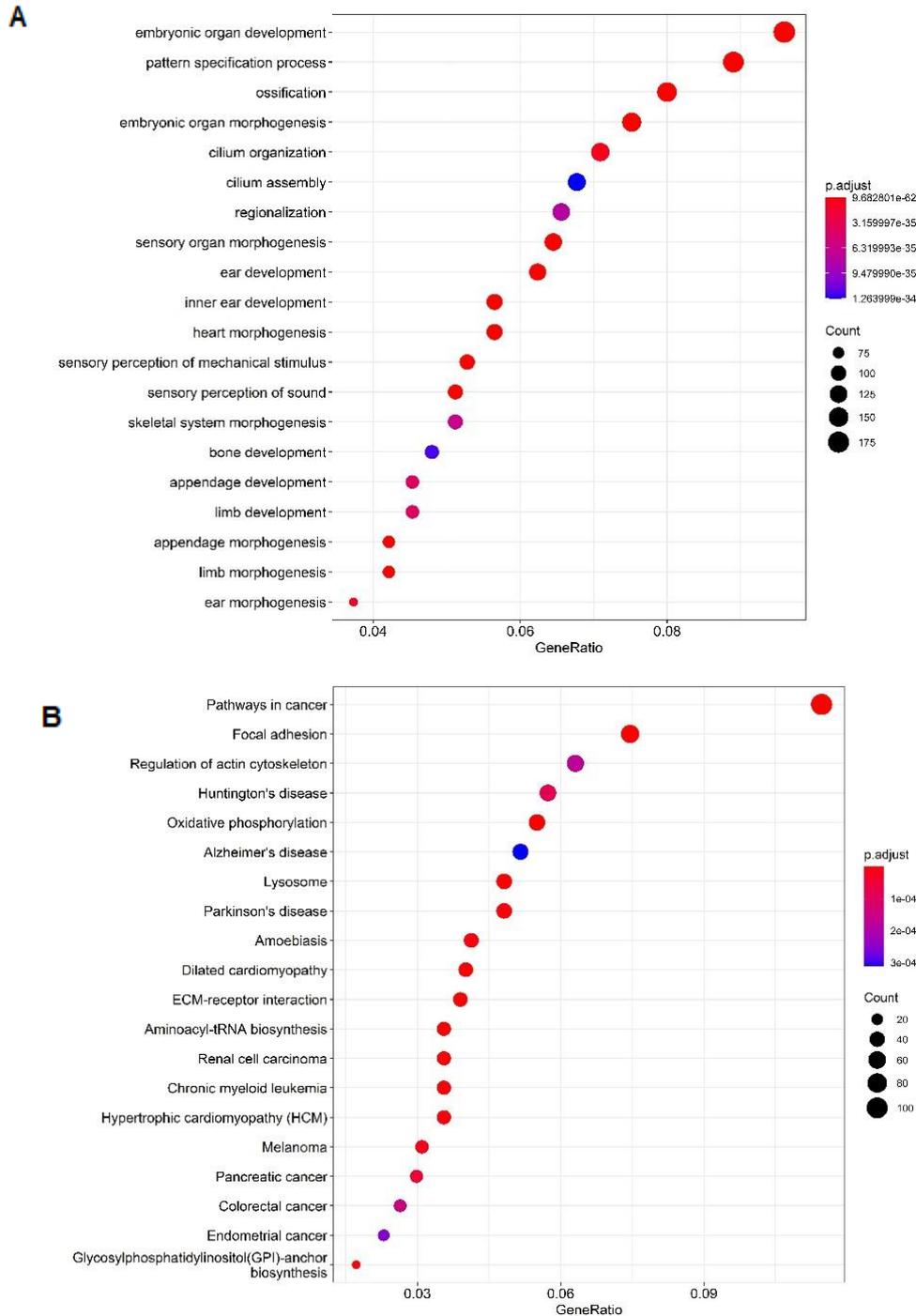


Figura 3: (A) Ontologias Gênicas (GO) e (B) vias de KEGG. Embryonic organ development: desenvolvimento de órgão embrionário; pattern specification process: processo de especificação de padrão; Ossification: Ossificação; embryonic organ morphogenesis: morfogênese de órgão embrionário; cilium organization: organização do cílio; cilium assembly: montagem do cílio; regionalization: regionalização;

sensory organ morphogenesis: morfogênese do órgão sensorial; ear development: desenvolvimento da orelha; inner ear development: desenvolvimento do ouvido interno; heart morphogenesis: morfogênese do coração; sensory perception of mechanical stimulus: percepção sensorial de estímulo mecânico; sensory perception of sound percepção: sensorial do som; skeletal system morphogenesis: morfogênese do sistema esquelético; bone development: desenvolvimento ósseo; appendage development: desenvolvimento de apêndice; limb development: desenvolvimento de membro; appendage morphogenesis: morfogênese do apêndice; limb morphogenesis: morfogênese do membro; ear morphogenesis: morfogênese da orelha; Pathways in cancer: Vias do câncer; Focal adhesion: Adesão focal; Regulation of actin cytoskeleton: Regulação do citoesqueleto de actina; Huntington's disease: Doença de Huntington; Oxidative phosphorylation: Fosforilação oxidativa; Alzheimer's disease: doença de Alzheimer; Lysosome: Lisossoma; Parkinson's disease: doença de Parkinson; Amoebiasis: Amebíase; Dilated cardiomyopathy: Cardiomiopatia dilatada; ECM-receptor interaction: Interação ECM-receptor; Aminoacyl-tRNA biosynthesis: Biossíntese de aminoacil-tRNA; Renal cell carcinoma: Carcinoma de células renais; Chronic myeloid leucemia: Leucemia mielóide crônica; Hypertrophic cardiomyopathy: Cardiomiopatia hipertrófica; Melanoma: Melanoma; Pancreatic câncer: Câncer de pâncreas; Colorectal câncer: Câncer colorretal; Endometrial cancer Glycosylphosphatidylinositol(GPI)-anchor: câncer endometrial Glicosilfosfatidilinositol (GPI) – âncora; Biosynthesis: biossíntese.

5.2 ABORDAGEM A PARTIR DE OSTEOLASTOS

Partindo para uma segunda abordagem, realizamos análises utilizando dados de osteoblastos e mesenquimais (Rauch et al., 2019), disponível sob código GSE113253 no GEO. Primeiramente, foi estudado a co-expressão diferencial dos genes selecionados na seção 4.2. Dos dados gerados, foi observado que somente *COL1A1*, *COL1A2*, *FAM46A*, *TMEM38B*, *CREB3L1* e *PSL3*, genes já associados com OI, apresentavam correlação significativa com outros genes. Além disso, *COL1A1* e *COL1A2* não foram os “mandantes” em nenhuma co-expressão diferencial, enquanto os outros genes todos foram os “mandantes”.

Quando cruzadas as listas de genes que apresentam co-expressão diferencial dos genes já associados à OI não é observado nenhum gene presente em todas as redes. A Tabela 3 apresenta as 15 maiores correlações encontradas no diagrama de Venn. No entanto, *TMEM38B* é co-expresso com *CREB3L1* e *FAM46A*.

Tabela 3: Diagrama de Venn das listas de genes associados à OI.

GENES OI	TOTAL	GENES COEXISTENTES
COL1A1 COL1A2 CREB3L1 PLS3	2	CA2 RNF113A
COL1A1 CREB3L1 FAM46A PLS3	6	MUSK FGF10 KLLN CRB1 CACNA1G LSS
COL1A2 CREB3L1 PLS3 TMEM38B	1	GRHL3
COL1A1 COL1A2 TMEM38B	2	CCDC65 LFNG
COL1A1 COL1A2 CREB3L1	1	CHRNA1
COL1A1 COL1A2 PLS3	1	CRLF1
COL1A1 CREB3L1 FAM46A	3	COLEC11 PLK4 WWOX
COL1A1 CREB3L1 PLS3	7	REEP6 A2ML1 CDKN1B CLCN4 MAFB GPC3 TRPV3
COL1A2 FAM46A TMEM38B	1	RYR1
COL1A2 CREB3L1 PLS3	3	IPW PUS3 PHYH
COL1A2 FAM46A PLS3	3	LRP6 NFASC COL7A1
CREB3L1 FAM46A TMEM38B	4	ABCC6 KCNQ4 AEBP1 TMEM38B
CREB3L1 PLS3 TMEM38B	2	SOX4 NR2E3
FAM46A PLS3 TMEM38B	1	MYO18B
CREB3L1 FAM46A PLS3	11	KCNJ2 SOS1 ACY1 NUBPL TBL1XR1 BBS10 PEX1 ZEB2 EYA4 TBC1D20 TDO2

Com os genes que apresentaram correlação diferencial com *COL1A1*, *COL1A2*, *FAM46A*, *TMEM38B*, *CREB3L1* e *PSL3*, foi possível analisar o enriquecimento de GO e KEGG (Figura 4). Nas ontologias e vias se destacaram principalmente ossificação e adesão focal, respectivamente. Curiosamente, dos genes relacionados com ossificação como *COL1A2*, *COL11A1*, *COL1A1*, *COL2A1*, *CTNNA1*, *COL6A1* também são vistos em adesão focal (Figura 5).

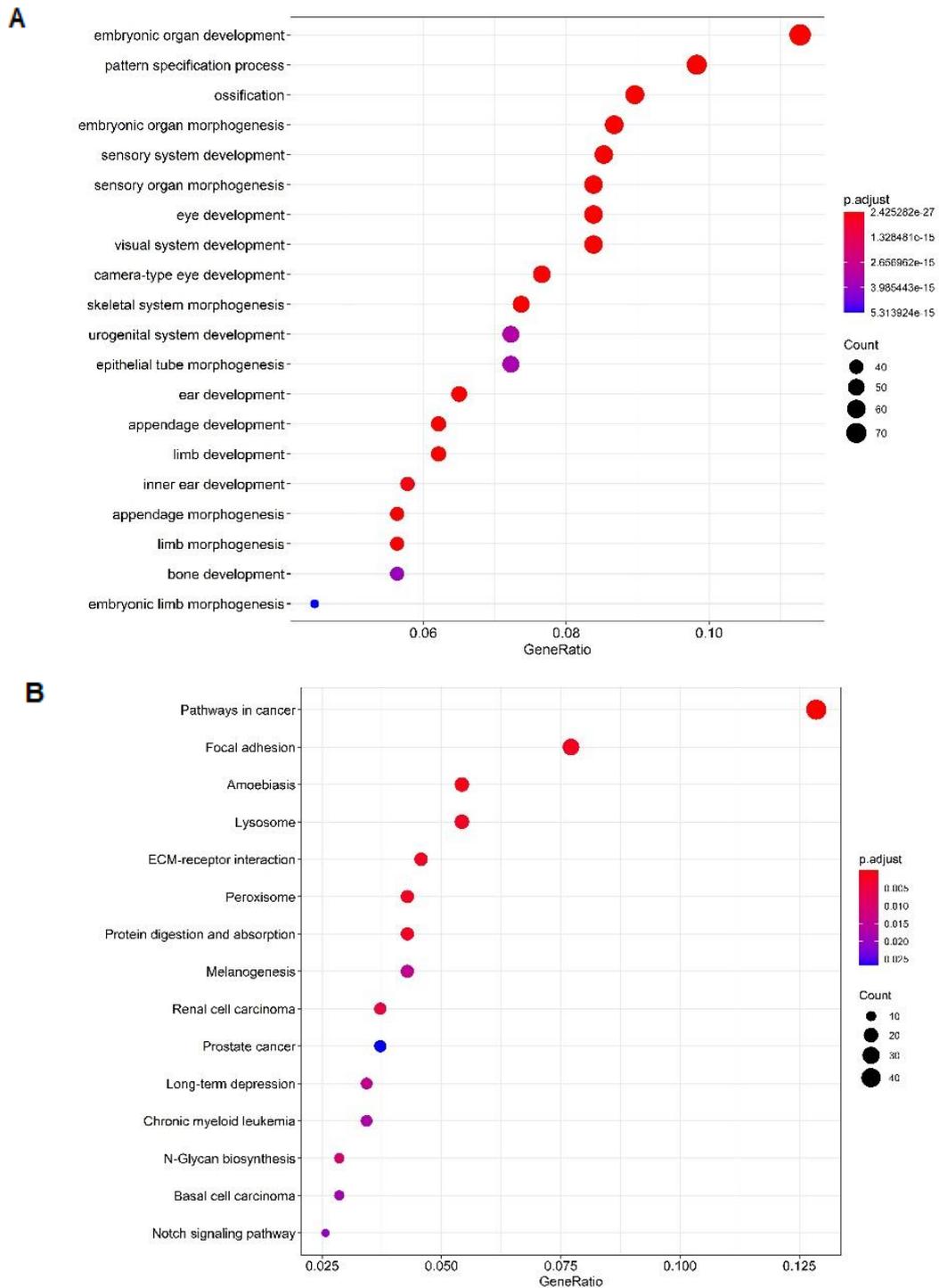


Figura 4: (A) Ontologias gênicas e (B) vias de KEGG. Embryonic organ development: desenvolvimento de órgão embrionário; Pattern specification process: processo de especificação de padrão; Ossification: ossificação; Embryonic organ morphogenesis: morfogênese de órgão embrionário; Sensory system development: desenvolvimento do sistema sensorial; Sensory organ morphogenesis: morfogênese do órgão sensorial; Eye development: desenvolvimento do olho; Visual system development: desenvolvimento do sistema visual; Camera-type eye development: desenvolvimento

de olho tipo câmara; Skeletal system morphogenesis: morfogênese do sistema esquelético; Urogenital system development: desenvolvimento do sistema urogenital; Epithelial tube morphogenesis: morfogênese do tubo epitelial; Ear development: desenvolvimento da orelha; Appendage development: desenvolvimento de apêndice; Limb development: desenvolvimento de membro; Inner ear development: desenvolvimento da orelha interna; Appendage morphogenesis: morfogênese do apêndice; Limb morphogenesis: morfogênese do membro; Bone development: desenvolvimento ósseo; Embryonic limb morphogenesis: morfogênese do membro embrionário; Pathways in cancer: vias de câncer; Focal adhesion: adesão focal; Amoebiasis: amebíase; Lysosome: lisossoma; ECM-receptor interaction: interação ECM-receptor; Peroxisome: peroxissomo; Protein digestion and absorption: digestão e absorção de proteínas; Melanogenesis: Melanogênese; Renal cell carcinoma: carcinoma de células renais; Prostate cancer: câncer de próstata; Long-term depression: depressão de longo prazo; Chronic myeloid leukemia: leucemia mielóide crônica; N-Glycan biosynthesis: biossíntese de N-Glicano; Basal cell carcinoma: carcinoma basocelular; Notch signaling pathway: via de sinalização Notch.

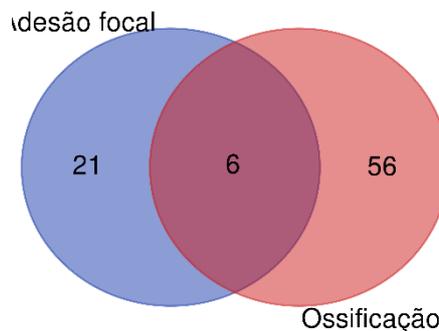


Figura 5: Diagrama de Venn dos genes relacionados com ossificação e adesão focal.

5.3 VERIFICAÇÃO DE EXPRESSÃO DIFERENCIAL

Com interesse de avaliar se os genes relacionados com ossificação e adesão focal sofriam alguma diferença de expressão significativa no osso em OI com relação a pacientes controles, escolhemos amostras de um estudo com duas linhagens de camundongos, *oim/oim* e *CrptaKO*, que representam dois tipos de OI com diferentes mutações genéticas resultantes de deficiência no colágeno tipo 1 e uma linhagem do

tipo selvagem (Zimmerman et al., 2019). Os dados foram gerados a partir de RNAseq de fêmures e tíbias corticais enriquecidos com osteócitos.

Após o tratamento no servidor use*Galaxy*, foram gerados os dados de expressão diferencial de *oim/oim* versus controle e *CrptaKO* versus controle, utilizando a lista de co-expressão dos genes já associados à OI. Os critérios de valor significativos aplicados para log foram ≥ 1 ou ≤ -1 e para valor ajustado de *p* de $\leq 0,05$. Assim, obtivemos 921 genes diferencialmente expressos em *oim/oim* versus controle e 1414 genes em *CrptaKO* versus controle.

Ao cruzarmos os genes diferencialmente expressos encontrados com os genes de ossificação e adesão focal chegamos ao diagrama de Venn apresentado na Figura 6. Na Tabela 4 estão descritos os resultados encontrados na expressão diferencial para os genes que foram vistos em coexistência em ossificação e/ou adesão focal.

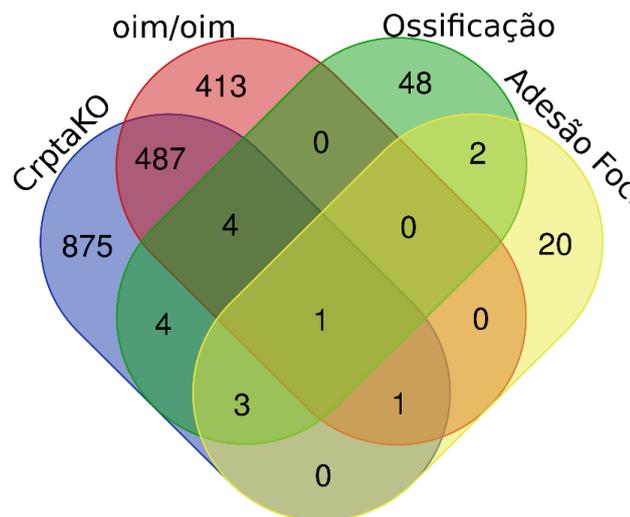


Figura 6: Diagrama de Venn de genes que apresentaram valor significativo na análise de expressão diferencial de *oim/oim* VS controle e *CrptaKO* VS controle com genes relacionados a ossificação e adesão focal.

Tabela 4: Resultado da expressão diferencial de *oim/oim* VS controle e *CrptaKO* VS controle dos genes encontrados em ossificação e/ou adesão focal.

Gene	Relacionado	Expressão diferencial	logFC	p-valor	FDR
COL1A1	Adesão Focal <i>CrptaKO</i> Ossificação <i>oim/oim</i>	<i>Oim/oim</i>	1.140730539	3.81e-21	1.67e-19
		<i>CrptaKo</i>	1.32463704	1.00e-27	7.10e-26
CREB3L1		<i>Oim/oim</i>	1.71201276	3.67e-32	5.64e-30

	<i>CrptaKO</i> Ossificação <i>oim/oim</i>	<i>CrptaKo</i>	1.714792193	2.93e-32	3.48e-30
<i>MMP2</i>		<i>Oim/oim</i>	1.28831419	1.38e-20	5.61e-19
		<i>CrptaKo</i>	1.265748428	6.02e-20	1.76e-18
<i>FBN2</i>		<i>Oim/oim</i>	1.046259351	2.49e-19	8.73e-18
		<i>CrptaKo</i>	1.380926091	3.10e-32	3.66e-30
<i>IHH</i>		<i>Oim/oim</i>	1.205585873	5.06e-10	5.90e-09
		<i>CrptaKo</i>	1.259488042	7.40e-11	7.96e-10
<i>COL11A2</i>	Adesão Focal	<i>Oim/oim</i>	1.002117144	2.75e-09	2.91e-08
	<i>CrptaKO</i> <i>oim/oim</i>	<i>CrptaKo</i>	1.474009731	5.08e-18	1.21e-16
<i>COL2A1</i>	Adesão Focal	<i>CrptaKo</i>	1.156025751	1.50e-06	9.71e-06
<i>COL11A1</i>	<i>CrptaKO</i> Ossificação	<i>CrptaKo</i>	1.56059926	6.51e-21	2.14e-19
<i>COL1A2</i>		<i>CrptaKo</i>	1.28304845	3.65e-137	9.91e-133
<i>SOX9</i>	<i>CrptaKO</i> Ossificação	<i>CrptaKo</i>	1.042196215	8.80e-10	8.44e-09
<i>SATB2</i>		<i>CrptaKo</i>	1.350703318	3.92e-19	1.05e-17
<i>PHEX</i>		<i>CrptaKo</i>	1.13904522	1.06e-27	7.46e-26
<i>GLI2</i>		<i>CrptaKo</i>	1.268534788	3.28e-17	7.08e-16

6 DISCUSSÃO

Os pacientes portadores de OI apresentam uma gama de sintomas clínicos, exibindo, também, uma heterogeneidade genética entre os indivíduos (Marini et al., 2007). Com o avanço da tecnologia e dos estudos em OI, cada vez mais tem se associado genes a doença. Acredita-se que novos genes e novas variantes patogênicas podem vir a ser identificadas e relacionadas à OI.

A primeira abordagem realizada para identificação de possíveis novos genes associados a OI foi utilizando análises a partir do fenótipo. A rede de interação inicial gerada apresentava quase 2 mil genes tornando difícil de ser analisada. Por ventura, decidimos manter na rede somente os genes que já estavam associados a três ou mais dos fenótipos selecionados (Figura 2). Assim, obtivemos uma rede que contemplava somente 7 dos genes já associados a OI, *COL1A1*, *COL1A2*, *CRTAP*, *PPIB*, *SP7*, *BMP1* e *WNT1*, estando todos fora da rede principal. Além disso, nenhum gene se destacou na rede.

A partir desses resultados, seguimos para uma segunda abordagem voltada para a co-expressão diferencial de genes em osteoblastos em comparação com células mesenquimais. Os osteoblastos são células que atuam na formação óssea, portanto estão relacionados a OI independente da causa genética (Marini et al., 2017). Ao realizarmos a co-expressão diferencial foi visto somente uma correlação significativa de *COL1A1*, *COL1A2*, *FAM46A*, *TMEM38B*, *CREB3L1* e *PSL3* dos genes já associados à OI. Além disso, ao cruzarmos as listas desses genes não é observado nenhum gene presente em todas as redes, como indica a Tabela 3.

No enriquecimento de GO e KEGG gerado neste trabalho, se destacam ossificação (ontologia) e via de adesão focal. Ossificação está diretamente relacionada a OI, uma vez que, é um grupo de doenças ósseas. Ao olharmos a via de adesão focal vimos a presença de associação à ECM (matriz extracelular), que está relacionada com diversos genes já associados a OI, como *COL1A1* e *COL1A2* (Figura 7).

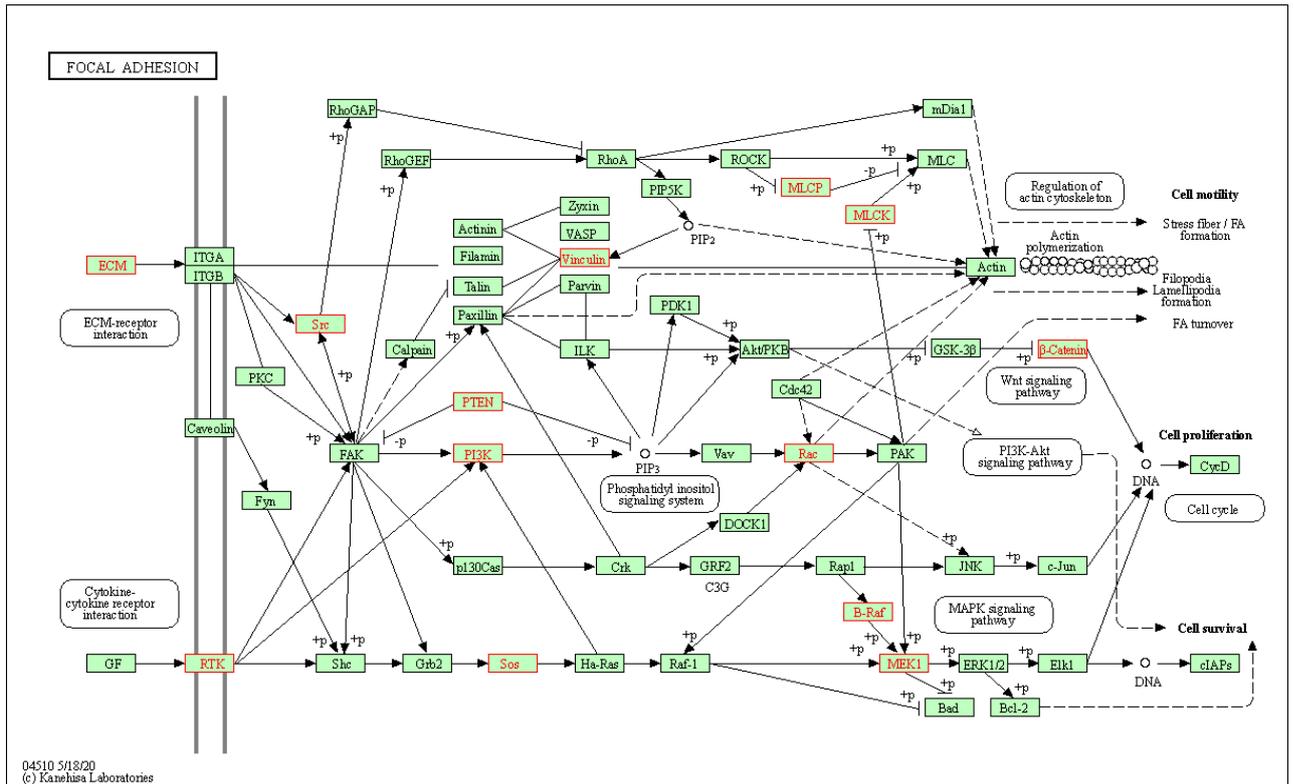


Figura 7: Via de KEGG de adesão focal.

Para verificarmos se os genes associados a ossificação e adesão focal sofriam alguma alteração, realizamos análise de expressão diferencial com dados de RNAseq a partir de osso (Zimmerman et. al, 2019). Ao cruzarmos as listas de genes de ossificação e adesão focal com os genes que apresentaram valor significativo na análise de expressão diferencial, encontramos 13 genes relacionados (Tabela 4). Desses, três já são associados à OI, *COL1A1*, *COL2A1* e *CREB3L1*. Os outros genes já foram associados a alguma doença como visto na (Tabela 5).

Tabela 5: Função e fenótipo dos genes encontrados na Tabela 4 que não são associados à OI.

Gene	OMIM	Fenótipo
<i>MMP2</i>	*120360	Osteólise multicêntrica, nodulose e artropatia
<i>FBN2</i>	*612570	Aracnodactilia contratural congênita
		Degeneração macular
<i>IHH</i>	*600726	Displasia acrocapitofemoral
		Braquidactilia, tipo A1
<i>COL11A2</i>	*120290	Surdez autossômica dominante 13
		Surdez autossômica recessiva 53
		Fibrocondrogênese 2
		Displasia otopondilomegaepifisária, autossômica dominante

		Displasia otopondilomegaepifisária, autossômica recessiva
COL2A1	*120140	Acondrogênese, tipo II ou hipocondrogênese
		Necrose avascular da cabeça femoral
		Displasia checa
		Displasia de Kniest
		Doença de Legg-Calve-Perthes
		Osteoartrite com condrodisplasia leve
		Displasia esquelética platisspondílica, tipo Torrance
		SED congênita
		Tipo SMED Strudwick
		Displasia espondiloepifisária, tipo Stanescu
		Displasia espondiloperiférica
		Síndrome de Stickler, tipo I, ocular não síndrômica
		Síndrome de Stickler, tipo I
		Vitreo-retinopatia com displasia epifisária falangeana
COL11A1	*120280	Fibrocondrogênese 1
		Síndrome de Marshall
		Síndrome de Stickler, tipo II
SOX9	*608160	Displasia campomélica acampomélica
		Displasia campomélica
		Displasia campomélica com reversão sexual autossômica
SATB2	*608148	Síndrome de Glass
PHEX	*300550	Raquitismo hipofosfatêmico, dominante ligado ao X
GLI2	*165230	Síndrome de Culler-Jones
		Holoprosencefalia 9

Dados coletados do site OMIM (<https://www.omim.org/>)

Nos genes apresentados na Tabela 5 podemos ver uma presença de genes que estão envolvidos com o colágeno. Como a *MMP2*, que codifica uma proteína que degrada o colágeno tipo IV e V (Murphy et al., 1985). *COL2A1*, *COL11A1* e *COL11A2* estão envolvidos na formação de colágenos, e ainda os dois primeiros estão relacionados com ossificação e adesão focal (Cheah et al, 1985; Jun et al., 2001; Lui et al., 1996).

O gene *PHEX* presente no cromossomo X, codifica uma proteína responsável pela degradação da ligação de integrinas, sendo predominante expresso em osteoblastos (Barros et. al, 2013). Já o gene *IHH* tem como função a sinalização na regulação da morfogênese embrionária (Marigo et al., 1995). Os fatores de transcrição *SOX9* e *GLI2* atuam no desenvolvimento esquelético e na transdução do sinal SHH, respectivamente (Cox et al., 2011; Roessler et al., 2003). Enquanto, *SATB2* participa na regulação da transcrição e da remodelação da cromatina (Leoyklang et al., 2013).

E a fibrilina codificada por *FBN2*, atua na formação e remodelação de ECM (Olivieri et al., 2010).

Esses 10 genes listados na Tabela 5 seriam possíveis genes a serem associados à OI. Estudos *in vivo* e funcionais serão necessários para corroborar os nossos achados.

REFERÊNCIAS

- ALANAY, Y.; AVAYGAN, H.; CAMACHO, N.; UTINE, G. E. et al. Mutations in the gene encoding the RER protein FKBP65 cause autosomal-recessive osteogenesis imperfecta. *Am J Hum Genet*, 86, n. 4, p. 551-559, Apr 2010.
- AFGAN, E.; BAKER, D.; BATUT, B.; VAN DEN BEEK, M. et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*, 46, n. W1, p. W537-W544, Jul 2018.
- ASHBURNER, M.; BALL, C. A.; BLAKE, J. A.; BOTSTEIN, D. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*, 25, n. 1, p. 25-29, May 2000.
- BARBER, L. A.; ABBOTT, C.; NAKHATE, V.; DO, A. N. D. et al. Longitudinal growth curves for children with classical osteogenesis imperfecta (types III and IV) caused by structural pathogenic variants in type I collagen. *Genet Med*, 21, n. 5, p. 1233-1239, 05 2019.
- BARROS, N. M.; HOAC, B.; NEVES, R. L.; ADDISON, W. N. et al. Proteolytic processing of osteopontin by PHEX and accumulation of osteopontin fragments in Hyp mouse bone, the murine model of X-linked hypophosphatemia. *J Bone Miner Res*, 28, n. 3, p. 688-699, Mar 2013.
- BREGOU BOURGEOIS, A.; AUBRY-ROZIER, B.; BONAFÉ, L.; LAURENT-APPLEGATE, L. et al. Osteogenesis imperfecta: from diagnosis and multidisciplinary treatment to future perspectives. *Swiss Med Wkly*, 146, p. w14322, 2016.
- CABRAL, W. A.; CHANG, W.; BARNES, A. M.; WEIS, M. et al. Prolyl 3-hydroxylase 1 deficiency causes a recessive metabolic bone disorder resembling lethal/severe osteogenesis imperfecta. *Nat Genet*, 39, n. 3, p. 359-365, Mar 2007.
- CAO, Y. J.; ZHANG, H.; ZHANG, Z. L. NOVEL MUTATIONS IN THE WNT1, TMEM38B, P4HB, AND PLS3 GENES IN FOUR UNRELATED CHINESE FAMILIES WITH OSTEOGENESIS IMPERFECTA. *Endocr Pract*, 25, n. 3, p. 230-241, Mar 2019.
- CHEAH, K. S.; STOKER, N. G.; GRIFFIN, J. R.; GROSVELD, F. G. et al. Identification and characterization of the human type II collagen gene (COL2A1). *Proc Natl Acad Sci U S A*, 82, n. 9, p. 2555-2559, May 1985.
- CHRISTIANSEN, H. E.; SCHWARZE, U.; PYOTT, S. M.; ALSWAID, A. et al. 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, n. 3, p. 389-398, Mar 2010.
- COX, J. J.; WILLATT, L.; HOMFRAY, T.; WOODS, C. G. A SOX9 duplication and familial 46,XX developmental testicular disorder. *N Engl J Med*, 364, n. 1, p. 91-93, Jan 2011.

DOYARD, M.; BACROT, S.; HUBER, C.; DI ROCCO, M. et al. mutations are responsible for autosomal recessive osteogenesis imperfecta. *J Med Genet*, 55, n. 4, p. 278-284, 04 2018.

ENGELBERT, R. H.; UITERWAAL, C. S.; GERVER, W. J.; VAN DER NET, J. J. et al. Osteogenesis imperfecta in childhood: impairment and disability. A prospective study with 4-year follow-up. *Arch Phys Med Rehabil*, 85, n. 5, p. 772-778, May 2004.

EYRE, D. R.; WEIS, M. A. Bone collagen: new clues to its mineralization mechanism from recessive osteogenesis imperfecta. *Calcif Tissue Int*, 93, n. 4, p. 338-347, Oct 2013.

JAIN, M.; TAM, A.; SHAPIRO, J. R.; STEINER, R. D. et al. Growth characteristics in individuals with osteogenesis imperfecta in North America: results from a multicenter study. *Genet Med*, 21, n. 2, p. 275-283, 02 2019.

JUN, A. S.; LIU, S. H.; KOO, E. H.; DO, D. V. et al. Microarray analysis of gene expression in human donor corneas. *Arch Ophthalmol*, 119, n. 11, p. 1629-1634, Nov 2001.

KANEHISA, M.; FURUMICHI, M.; TANABE, M.; SATO, Y. et al. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res*, 45, n. D1, p. D353-D361, 01 2017.

KEUPP, K.; BELEGGIA, F.; KAYSERILI, H.; BARNES, A. M. et al. Mutations in WNT1 cause different forms of bone fragility. *Am J Hum Genet*, 92, n. 4, p. 565-574, Apr 2013.

KOWALSKI, T. W.; DUPONT, Á.; RENGEL, B. D.; SGARIONI, E. et al. Assembling systems biology, embryo development and teratogenesis: What do we know so far and where to go next? *Reprod Toxicol*, 88, p. 67-75, 09 2019.

LAINE, C. M.; JOENG, K. S.; CAMPEAU, P. M.; KIVIRANTA, R. et al. WNT1 mutations in early-onset osteoporosis and osteogenesis imperfecta. *N Engl J Med*, 368, n. 19, p. 1809-1816, May 2013.

LAPUNZINA, P.; AGLAN, M.; TEMTAMY, S.; CAPARRÓS-MARTÍN, J. A. et al. Identification of a frameshift mutation in Osterix in a patient with recessive osteogenesis imperfecta. *Am J Hum Genet*, 87, n. 1, p. 110-114, Jul 2010.

LE NOVÈRE, N. Quantitative and logic modelling of molecular and gene networks. *Nat Rev Genet*, 16, n. 3, p. 146-158, Mar 2015.

LEOYKLANG, P.; SUPHAPEETIPORN, K.; SRICHOMTHONG, C.; TONGKOBPETCH, S. et al. Disorders with similar clinical phenotypes reveal underlying genetic interaction: SATB2 acts as an activator of the UPF3B gene. *Hum Genet*, 132, n. 12, p. 1383-1393, Dec 2013.

LIN, H. Y.; CHUANG, C. K.; SU, Y. N.; CHEN, M. R. et al. Genotype and phenotype analysis of Taiwanese patients with osteogenesis imperfecta. *Orphanet J Rare Dis*, 10, p. 152, Dec 2015.

LINDERT, U.; CABRAL, W. A.; AUSAVARAT, S.; TONGKOBPETCH, S. et al. MBTPS2 mutations cause defective regulated intramembrane proteolysis in X-linked osteogenesis imperfecta. *Nat Commun*, 7, p. 11920, 07 2016.

LOMAURO, A.; FRASCHINI, P.; POCHINTESTA, S.; ROMEI, M. et al. Ribcage deformity and the altered breathing pattern in children with osteogenesis imperfecta. *Pediatr Pulmonol*, 53, n. 7, p. 964-972, 07 2018.

LU, Y.; ZHANG, S.; WANG, Y.; REN, X. et al. Molecular mechanisms and clinical manifestations of rare genetic disorders associated with type I collagen. *Intractable Rare Dis Res*, 8, n. 2, p. 98-107, May 2019.

LUI, V. C.; NG, L. J.; SAT, E. W.; CHEAH, K. S. The human alpha 2(XI) collagen gene (COL11A2): completion of coding information, identification of the promoter sequence, and precise localization within the major histocompatibility complex reveal overlap with the KE5 gene. *Genomics*, 32, n. 3, p. 401-412, Mar 1996.

MAKAREEVA, E.; AVILES, N. A.; LEIKIN, S. Chaperoning osteogenesis: new protein-folding disease paradigms. *Trends Cell Biol*, 21, n. 3, p. 168-176, Mar 2011.

MARIGO, V.; ROBERTS, D. J.; LEE, S. M.; TSUKUROV, O. et al. Cloning, expression, and chromosomal location of SHH and IHH: two human homologues of the Drosophila segment polarity gene hedgehog. *Genomics*, 28, n. 1, p. 44-51, Jul 1995.

MARINI, J. C.; CABRAL, W. A.; BARNES, A. M. Null mutations in LEPRE1 and CRTAP cause severe recessive osteogenesis imperfecta. *Cell Tissue Res*, 339, n. 1, p. 59-70, Jan 2010.

MARINI, J. C.; FORLINO, A.; BÄCHINGER, H. P.; BISHOP, N. J. et al. Osteogenesis imperfecta. *Nat Rev Dis Primers*, 3, p. 17052, Aug 2017.

MARINI, J. C.; FORLINO, A.; CABRAL, W. A.; BARNES, A. M. et al. 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, n. 3, p. 209-221, Mar 2007.

MAROM, R.; LEE, Y. C.; GRAFE, I.; LEE, B. Pharmacological and biological therapeutic strategies for osteogenesis imperfecta. *Am J Med Genet C Semin Med Genet*, 172, n. 4, p. 367-383, 12 2016.

MARTIN, E.; SHAPIRO, J. R. Osteogenesis imperfecta: epidemiology and pathophysiology. *Curr Osteoporos Rep*, 5, n. 3, p. 91-97, Sep 2007.

MARTÍNEZ-GLEZ, V.; VALENCIA, M.; CAPARRÓS-MARTÍN, J. A.; AGLAN, M. et al. Identification of a mutation causing deficient BMP1/mTLD proteolytic activity in

autosomal recessive osteogenesis imperfecta. *Hum Mutat*, 33, n. 2, p. 343-350, Feb 2012.

MATSUO, K.; IRIE, N. Osteoclast-osteoblast communication. *Arch Biochem Biophys*, 473, n. 2, p. 201-209, May 2008.

MENDOZA-LONDONO, R.; FAHIMINIYA, S.; MAJEWSKI, J.; TÉTREAULT, M. et al. Recessive osteogenesis imperfecta caused by missense mutations in SPARC. *Am J Hum Genet*, 96, n. 6, p. 979-985, Jun 2015.

MOOSA, S.; YAMAMOTO, G. L.; GARBES, L.; KEUPP, K. et al. Autosomal-Recessive Mutations in MESD Cause Osteogenesis Imperfecta. *Am J Hum Genet*, 105, n. 4, p. 836-843, 10 2019.

MORELLO, R.; BERTIN, T. K.; CHEN, Y.; HICKS, J. et al. CRTAP is required for prolyl 3-hydroxylation and mutations cause recessive osteogenesis imperfecta. *Cell*, 127, n. 2, p. 291-304, Oct 2006.

MORTIER, G. R.; COHN, D. H.; CORMIER-DAIRE, V.; HALL, C. et al. Nosology and classification of genetic skeletal disorders: 2019 revision. *Am J Med Genet A*, 179, n. 12, p. 2393-2419, 12 2019.

MROSK, J.; BHAVANI, G. S.; SHAH, H.; HECHT, J. et al. Diagnostic strategies and genotype-phenotype correlation in a large Indian cohort of osteogenesis imperfecta. *Bone*, 110, p. 368-377, 05 2018.

MURPHY, G.; MCALPINE, C. G.; POLL, C. T.; REYNOLDS, J. J. Purification and characterization of a bone metalloproteinase that degrades gelatin and types IV and V collagen. *Biochim Biophys Acta*, 831, n. 1, p. 49-58, Sep 1985.

NAJIRAD, M.; MA, M. S.; RAUCH, F.; SUTTON, V. R. et al. Oral health-related quality of life in children and adolescents with osteogenesis imperfecta: cross-sectional study. *Orphanet J Rare Dis*, 13, n. 1, p. 187, 10 2018.

OLIVIERI, J.; SMALDONE, S.; RAMIREZ, F. Fibrillin assemblies: extracellular determinants of tissue formation and fibrosis. *Fibrogenesis Tissue Repair*, 3, p. 24, Dec 2010.

OMIM. Online Mendelian Inheritance in Man, 2021. Disponível em: <<https://www.omim.org/>>. Acesso em: 22 de abril de 2021.

PALOMO, T.; AL-JALLAD, H.; MOFFATT, P.; GLORIEUX, F. H. et al. Skeletal characteristics associated with homozygous and heterozygous WNT1 mutations. *Bone*, 67, p. 63-70, Oct 2014.

PILLION, J. P.; VERNICK, D.; SHAPIRO, J. Hearing loss in osteogenesis imperfecta: characteristics and treatment considerations. *Genet Res Int*, 2011, p. 983942, 2011.

RAUCH, A.; HAAKONSSON, A. K.; MADSEN, J. G. S.; LARSEN, M. et al. Author Correction: Osteogenesis depends on commissioning of a network of stem cell

transcription factors that act as repressors of adipogenesis. *Nat Genet*, 51, n. 4, p. 766, 04 2019.

RAUCH, F.; LALIC, L.; ROUGHLEY, P.; GLORIEUX, F. H. Relationship between genotype and skeletal phenotype in children and adolescents with osteogenesis imperfecta. *J Bone Miner Res*, 25, n. 6, p. 1367-1374, Jun 2010.

ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26, n. 1, p. 139-140, Jan 2010.

ROESSLER, E.; DU, Y. Z.; MULLOR, J. L.; CASAS, E. et al. Loss-of-function mutations in the human *GLI2* gene are associated with pituitary anomalies and holoprosencephaly-like features. *Proc Natl Acad Sci U S A*, 100, n. 23, p. 13424-13429, Nov 2003.

ROSSI, V.; LEE, B.; MAROM, R. Osteogenesis imperfecta: advancements in genetics and treatment. *Curr Opin Pediatr*, 31, n. 6, p. 708-715, 12 2019.

SHAHEEN, R.; ALAZAMI, A. M.; ALSHAMMARI, M. J.; FAQEIH, E. et al. Study of autosomal recessive osteogenesis imperfecta in Arabia reveals a novel locus defined by *TMEM38B* mutation. *J Med Genet*, 49, n. 10, p. 630-635, Oct 2012.

SILLENCE, D. O.; RIMOIN, D. L.; DANKS, D. M. Clinical variability in osteogenesis imperfecta-variable expressivity or genetic heterogeneity. *Birth Defects Orig Artic Ser*, 15, n. 5B, p. 113-129, 1979.

SILLENCE, D. O.; BARLOW, K. K.; GARBER, A. P.; HALL, J. G. et al. Osteogenesis imperfecta type II delineation of the phenotype with reference to genetic heterogeneity. *Am J Med Genet*, 17, n. 2, p. 407-423, Feb 1984.

TAM, A.; CHEN, S.; SCHAUER, E.; GRAFE, I. et al. A multicenter study to evaluate pulmonary function in osteogenesis imperfecta. *Clin Genet*, 94, n. 6, p. 502-511, 12 2018.

TAUER, J. T.; ROBINSON, M. E.; RAUCH, F. Osteogenesis Imperfecta: New Perspectives From Clinical and Translational Research. *JBMR Plus*, 3, n. 8, p. e10174, Aug 2019.

VAN DIJK, F. S.; NESBITT, I. M.; ZWIKSTRA, E. H.; NIKKELS, P. G. et al. *PPIB* mutations cause severe osteogenesis imperfecta. *Am J Hum Genet*, 85, n. 4, p. 521-527, Oct 2009.

VAN DIJK, F. S.; ZILLIKENS, M. C.; MICHA, D.; RIESSLAND, M. et al. *PLS3* mutations in X-linked osteoporosis with fractures. *N Engl J Med*, 369, n. 16, p. 1529-1536, Oct 2013.

VAN DIJK, F. S.; SEMLER, O.; ETICH, J.; KÖHLER, A. et al. Interaction between *KDELR2* and *HSP47* as a Key Determinant in Osteogenesis Imperfecta Caused by Bi-allelic Variants in *KDELR2*. *Am J Hum Genet*, 107, n. 5, p. 989-999, Nov 2020.

WILLING, M. C.; DESCHENES, S. P.; SLAYTON, R. L.; ROBERTS, E. J. Premature chain termination is a unifying mechanism for COL1A1 null alleles in osteogenesis imperfecta type I cell strains. *Am J Hum Genet*, 59, n. 4, p. 799-809, Oct 1996.

YANG, J.; YU, H.; LIU, B. H.; ZHAO, Z. et al. DCGL v2.0: An R Package for Unveiling Differential Regulation from Differential Co-expression. *PLoS ONE*, 8, n. 11, e79729, Nov 2013.

YU, G.; WANG, L. G.; HAN, Y.; HE, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*, 16, n. 5, p. 284-287, Mai 2012.

ZIMMERMAN, S. M.; DIMORI, M.; HEARD-LIPSMAYER, M. E.; MORELLO, R. The Osteocyte Transcriptome Is Extensively Dysregulated in Mouse Models of Osteogenesis Imperfecta. *JBMR Plus*, 3, n. 7, p. e10171, Jul 2019.