

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**INVESTIGAÇÃO DE GENES RESPONSÁVEIS POR SÍNDROMES
ASSOCIADAS A MALFORMAÇÕES DE MEMBROS E SEU PAPEL NA
TERATOGENESE DA TALIDOMIDA**

Julia do Amaral Gomes

Dissertação de Mestrado submetida
ao Programa de Pós-Graduação em
Genética e Biologia Molecular da
UFRGS como requisito parcial para a
obtenção do grau de Mestre em
Genética e Biologia Molecular

Orientadora: Profa. Dra. Fernanda Sales Luiz Vianna

Co-Orientadora: Profa. Dra. Lavínia Schüler Faccini

Porto Alegre, março de 2017.

Este trabalho foi realizado no laboratório de Genética Médica e Evolução do Departamento de Genética da UFRGS, com apoio financeiro do Instituto Nacional de Genética Médica Populacional (INaGeMP), do Fundo de Incentivo à Pesquisa e Eventos do Hospital de Clínicas de Porto Alegre (FIPE/HCPA) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

A aluna recebeu bolsa de estudos concedida pelo CNPq, vinculada ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS (PPGBM/UFRGS).

Agradecimentos

Eu gostaria de agradecer algumas pessoas que foram muito importantes durante a realização desse trabalho e fazem parte dessa conquista.

Agradeço primeiramente à minha orientadora Fernanda por todos os ensinamentos e conhecimento passados ao longo destes dois anos, pelas críticas e sugestões ao trabalho, pelo incentivo, apoio e suporte em todas as etapas, pela amizade, gentileza e carinho sempre demonstrados. Toda essa dedicação me fez sempre ter confiança e segurança para trabalhar, fazendo com que agora eu me orgulhe dos resultados do nosso trabalho.

À professora Lavínia por ter me recebido de forma tão amável e receptiva no laboratório, ter aceitado me orientar e a partir daí me dar todo suporte para o desenvolvimento deste trabalho. Um exemplo de pessoa e de profissional que tenho orgulho de conhecer e conviver.

À Thayne por desde o primeiro encontro ter se mostrado tão disposta a contribuir com o meu trabalho, por toda a ajuda na execução de cada uma das etapas, pelos conselhos, incentivo e motivação sempre, por torcer pelo meu crescimento e pelo meu sucesso, e principalmente pela amizade que quero levar para sempre.

Ao Lucas por todas as contribuições a esse trabalho, desde o começo, mesmo distante. Todas as sugestões, críticas e suporte foram importantíssimos, principalmente na reta final.

À Bárbara e a Delva por terem me ensinado a trabalhar com os dados do sequenciamento de nova geração, e a Marina por ter realizado e tão gentilmente mostrado e explicado todas as etapas do sequenciamento.

Às ICs Bruna e Marilea por serem tão prestativas quando solicitadas, pelo bom humor e amizade.

À ABPST e a todos os indivíduos que aceitaram participar do nosso trabalho, pela confiança.

Ao CNPq, INaGeMP e HCPA pelo apoio financeiro.

Aos colegas do laboratório 113 e ao professor Nelson, pelas conversas e convívio tão prazeroso nesses dois anos. À Luiza e a Zuleide por terem me

auxiliado antes, ainda na prova de mestrado, por terem me apresentado o laboratório e me motivado a trabalhar com esse grupo tão especial, pela amizade e incentivo sempre.

Aos meus colegas das aulas pelo companheirismo e amizade, por compartilhamos nossos problemas e sempre tentarmos nos ajudar.

Ao Elmo por ser tão prestativo, gentil, bem-humorado e competente, sempre buscando a melhor forma de ajudar a todos.

Às minhas amigas Mari, May, Sassa, Débora e Géssica por torcerem por mim, me incentivarem a continuar, se orgulharem das minhas conquistas, me fazerem rir e renovarem minhas energias.

Agradeço imensamente aos meus familiares por sempre torcerem por mim, em especial às minhas avós Eroni e Rita e aos meus padrinhos Evanira e João que sempre fizeram o máximo para que eu chegasse até aqui e sempre rezaram para que tudo desse certo.

Aos meus pais, que nunca mediram esforços para me ajudar, por me incentivarem, torcerem por mim, se orgulharem de cada pequena conquista, pelo amor e dedicação.

Por fim, agradeço ao Julio, meu namorado, meu amor, por todo o incentivo, por ter um orgulho imenso de mim e do trabalho que desenvolvi, por me apoiar e entender minha ausência muitas vezes. Sem ele tudo seria muito mais difícil.

Muito obrigada!

Sumário

Lista de Abreviaturas	6
Resumo	9
Abstract	11
Capítulo I	13
1. Introdução	14
1.1. Histórico da Talidomida	14
<i>1.1.1. Potencial Teratogênico da Talidomida</i>	14
<i>1.1.2. Potencial Terapêutico da Talidomida</i>	15
1.2. Embriopatia por Talidomida	16
1.3. Mecanismos de Teratogênese da Talidomida	17
1.4. Fenocópias da Embriopatia por Talidomida	18
<i>1.4.1. Síndrome de Roberts</i>	19
<i>1.4.2. Síndrome de Holt-Oram</i>	21
<i>1.4.3. Síndrome de Duane de Raio Radial - Síndrome de Okihiro</i>	23
1.5. Interações de ESCO2, TBX5 e SALL4 com a talidomida	28
Capítulo II	30
2. Justificativa	31
Capítulo III	32
3. Objetivos	33
3.1. Objetivo Geral	33
3.2. Objetivos Específicos	33
Capítulo IV	34
Artigo I - The role of ESCO2, TBX5 and SALL4 genes in the susceptibility to thalidomide embryopathy	35
Capítulo V	68
5. Discussão	69
Capítulo VI	73
6. Conclusões e Perspectivas	74
Capítulo VII	76
7. Referências Bibliográficas	77
Capítulo VIII	85
8. Artigo Anexo - Genetic Susceptibility to Thalidomide Embryopathy in Humans: analyses in candidate genes	86

Lista de abreviaturas

ABPST	Associação Brasileira de Portadores da Síndrome da Talidomida
AIDS	<i>Acquired Immunodeficiency Syndrome</i> - Síndrome da Imunodeficiência Adquirida
ANVISA	Agência Nacional de Vigilância Sanitária
BMP4	<i>Bone morphogenic protein 4</i> - Proteína morfogênica óssea 4
Bp	<i>Base pairs</i> - Pares de base
CRBN	<i>Cereblon</i> - Cereblon
CYP2C19	<i>Cytochrome P450 2C19</i> - Citocromo P450 2C19
D'	Desequilíbrio de ligação
DL50	Dose letal aguda
DRRS	<i>Duane radial ray syndrome</i> - Síndrome de Duane de raio radial
ENH	Eritema Nodoso da Hanseníase
ESCO2	<i>Establishment of Sister Chromatid Cohesion N-acetyltransferase 2</i> - N-acetiltransferase 2 de estabelecimento de coesão entre cromátide-irmãs
ESE	<i>Exonic Splicing Enhancer</i> - Enhancer de splicing exônico
ESS	<i>Exonic Splicing Silencer</i> - Silenciador de splicing exônico
FGF10	<i>Fibroblast growth factor 10</i> - Fator de crescimento de fibroblastos 10
FGF8	<i>Fibroblast growth factor 8</i> - Fator de crescimento de fibroblastos 8
FGFs	<i>Fibroblast growth factors</i> - Fatores de crescimento de fibroblastos
F-SNP	<i>Functional Single Nucleotide Polymorphism</i>
c.	Sequência de referência de DNA codificante
HCPA	Hospital de Clínicas de Porto Alegre
HOS	<i>Holt-Oram syndrome</i> - Síndrome de Holt-Oram
hsa-miR	<i>human micro RNA of a mature sequence</i> - micro RNA humano de uma sequência madura

IBM	<i>International Business Machines Corp.</i>
INaGeMP	Instituto Nacional de Genética Médica Populacional
kb	Kilobase
LOD	<i>Logarithm of the odds</i>
mg	Miligrama
miRNA	<i>micro RNA</i> - micro RNA
mRNA	<i>messenger RNA</i> - RNA mensageiro
MSX1	<i>Muscle segment homeobox 1</i> - Segmento de músculo homeobox 1
NF-Kβ	<i>Nuclear factor - kappaβ</i> - Fator nuclear - kappa β
OMIM	<i>Online Mendelian Inheritance in Man</i>
p	Braço curto do cromossomo
PGM	<i>Personal Genome Machine</i>
PolyPhen	<i>Polymorphism Phenotyping</i>
q	Braço longo do cromossomo
RBS	<i>Roberts syndrome</i> - Síndrome de Roberts
RDC	Resolução da Diretoria Colegiada da Agência Nacional de Vigilância Sanitária
RING	<i>Really interesting new gene</i>
ROS	<i>Reactive oxygen species</i> - Espécies reativas de oxigênio
SALL4	<i>Spalt Like Transcription Factor 4</i> - Fator de transcrição Spalt 4
SHH	<i>Sonic hedgehog</i>
SIAT	Sistema Nacional de Informação sobre Agentes Teratogênicos
SIFT	<i>Sorting Intolerant From Tolerant</i>
SNPs	<i>Single Nucleotide Polymorphisms</i> - Polimorfismos de nucleotídeo único
sp.	<i>Specie</i> - Espécie
TBX5	<i>T-box transcription factor 5</i> - Fator de transcrição T-box 5
TE	<i>Thalidomide Embriopathy</i> - Embriopatia por Talidomida
TNF-α	<i>Tumoral necrosis factor alpha</i> - Fator de necrose tumoral alfa
TP53	<i>Tumor protein p53</i>
TP63	<i>Tumor protein p63</i>

TP73	<i>Tumor protein p73</i>
USA	<i>United States</i> - Estados Unidos
UTR	<i>Untranslated region</i> - Região não-traduzida
v.	Versão
VEGF	<i>Vascular endothelial growth factor</i> - Fator de crescimento do endotélio vascular

Resumo

A talidomida é um fármaco que foi sintetizado pela primeira vez em 1954. Poucos anos após sua síntese já era amplamente utilizada para uma série de indicações. No início da década de 1960, a alta incidência de bebês nascidos com defeitos de redução de membros foi correlacionada com o uso de talidomida por mulheres grávidas. Em torno de 20-50% dos embriões expostos à talidomida foram afetados e desenvolveram as malformações características da embriopatia por talidomida (TE). Atualmente sabe-se que o *background* genético dos indivíduos tem um papel importante nessa suscetibilidade, entretanto, os fatores associados à maior suscetibilidade e aqueles relacionados a diferentes fenótipos dentro da TE não são totalmente conhecidos. Algumas síndromes genéticas têm fenótipos muito semelhantes aos visualizados na TE, sendo a TE considerada uma fenocópia destas síndromes. As principais são: Síndrome de Roberts, Síndrome de Duane de raio radial – também chamada Síndrome de Okihiro –, e Síndrome de Holt-Oram, causadas por mutações nos genes *ESCO2*, *SALL4* e *TBX5*, respectivamente. A análise das bases moleculares e patogênese de síndromes genéticas de fenótipo similar a embriopatias causadas por fármacos (fenocópias) tem ajudado na compreensão de alguns mecanismos teratogênicos. O objetivo desse trabalho foi investigar a contribuição dos três genes acima citados na suscetibilidade à TE e seu papel nas diferentes anomalias congênitas e doenças observadas nos indivíduos afetados. Os éxons e regiões flangeadoras (UTRs e cerca de 50 pb dos íntrons adjacentes) desses três genes foram sequenciados em 29 indivíduos afetados pela TE e as variantes encontradas foram avaliadas através de preditores funcionais *in silico*. Quarenta e uma variantes foram identificadas, 11 (27%) no gene *ESCO2*, 14 (34%) no gene *SALL4* e 16 (39%) no gene *TBX5*. Inserções e deleções estavam presentes apenas no gene *ESCO2*. Duas das variantes identificadas, uma em *ESCO2* e uma em *TBX5*, não estavam descritas previamente em bancos de dados genômicos. Doze (29%) eram variantes genéticas raras (<0.01). As frequências alélicas e genotípicas das variantes foram determinadas na nossa amostra e comparadas com dados da população europeia do Projeto 1000 Genomas. Não

foram identificados alelos ou genótipos que conferissem risco ou proteção para a TE. As frequências foram significativamente diferentes entre os dois grupos apenas para uma variante 3'UTR (rs62498042) do gene *ESCO2* ($p < 0,001$), entretanto, não há relatos na literatura do papel funcional dessa variante ou de sua associação com aumento da susceptibilidade para alguma doença. A análise *in silico* que realizamos sugeriu possível alteração de *splicing* devido à presença dessa variante. Algumas variantes aqui identificadas foram previamente avaliadas em doenças cardiovasculares e malformações cardíacas. Nós comparamos a frequência dessas variantes em indivíduos com TE que apresentavam alguma anomalia cardíaca ou doença cardiovascular e aqueles que não apresentavam tais condições; entretanto, não houve associação entre a presença das variantes e risco aumentado para alguma anomalia cardíaca ou doença cardiovascular. Análises *in silico* foram realizadas a fim de avaliar o potencial de todas as variantes encontradas para alterações funcionais nas proteínas, alterações no *splicing* e em sítios de ligação de fatores de transcrição e miRNAs. Muitas variantes foram consideradas como tendo tal potencial, especialmente de afetar o *splicing*, entretanto, não foi possível estabelecer o papel dessas variantes no decaimento do mRNA desses genes, na formação de isoformas alternativas ou ainda na diminuição da expressão dos genes. Alguns resultados devem ser melhor explorados através de ensaios funcionais para determinar o papel das variantes na atividade das proteínas, no mecanismo molecular de desenvolvimento de anomalias congênitas e na teratogênese da talidomida.

Abstract

Thalidomide is a drug that was first synthesized in 1954. Following its release on to the market, thalidomide was already widely used for a diverse number of indications. The high incidence of babies being born with limb reduction defects in early 1960s was related with the use of thalidomide by pregnant women. Around 20-50% of the embryos exposed to thalidomide were affected and developed the characteristic malformations of thalidomide embryopathy (TE). Currently is known that the genetic background of the individuals plays an important role in this susceptibility; however, the factors associated with higher susceptibility and those related to different phenotypes are not fully known. Some phenotypes observed in TE are very similar to characteristics seen in some genetic syndromes. Because of that, TE is considered a phenocopy of these syndromes. The main syndromes that TE is phenocopy are: Roberts Syndrome, Duane Radial Ray Syndrome (also known as Okihiro Syndrome) and Holt-Oram Syndrome, caused by mutations in *ESCO2*, *SALL4* and *TBX5* genes, respectively. The analysis of the molecular bases of syndromes of similar phenotype to embryopathy cause by drugs (phenocopies) has helped in the understanding of teratogenic mechanisms. The aim of this study was to investigate the contribution of these three genes in TE susceptibility and theirs role in different congenital anomalies and diseases seen in affected individuals. The exons and flanking regions (UTRs and 50pb of adjacents introns) of these genes were sequenced by next generation sequencing in 29 individuals affected by TE and the variants found were evaluated through bioinformatic tools *in silico*. Forty-one variants were identified, 11 (27%) in *ESCO2*, 14 (34%) in *SALL4* and 16 (39%) in *TBX5*. Insertions and deletions were present only in *ESCO2* gene. Two of the identified variants, one in *ESCO2* and on in *TBX5*, were not previously described in genomic databases. Twelve variants (29%) were rare (<0.01). The allelic and genotypic frequencies of our sample were determined and compared with data from the European population of 1000 Genomes Project. No allele or genotype was related to risk or protection for TE. The frequencies were significantly different between the two groups for only one 3'UTR variant (rs62498042) found in *ESCO2* gene (p

<0.001). However, there are no reports on literature of the functional role of this variant or its association with increased susceptibility to some disease. *In silico* analyses here performed suggested splicing changes due this variant presence. Some variants here identified have been previously evaluated in cardiovascular diseases and cardiac malformations. We compared the frequency of these variants in individuals with TE that presented some cardiac anomaly or cardiovascular disease and those that did not present such conditions; however, there was no association between these variants and increased risk for these conditions. *In silico* analyses were performed in order to evaluate the potential of the variants for functional alterations in proteins, changes in splicing sites, transcription factors binding sites and miRNA interaction sites. Many variants were considered to have such potential, especially of splicing changes; however, it was not possible to establish the role of the variants in decay of these genes mRNA, or in alternative isoforms formation or in the decrease of these genes expression. Some variants should be further investigated. Functional analyses would be useful to investigate the role of the variants in the proteins activities, in the molecular mechanism of congenital anomalies development and in thalidomide teratogenesis.

CAPÍTULO I

INTRODUÇÃO

1. INTRODUÇÃO

1.1. Histórico da Talidomida

A talidomida foi sintetizada pela primeira vez em 1954, no laboratório *Chemie Grünenthal*, na Alemanha. Testes clínicos indicaram que a talidomida podia ser utilizada como um agente sedativo, devido ao seu efeito de depressor do sistema nervoso central (Lenz, 1988; Smithells & Newman, 1992). Testes de toxicidade foram realizados em roedores e mostraram que o medicamento apresentava baixo risco de intoxicação, não sendo nem mesmo possível determinar a dose letal (DL₅₀) (Shardein, 1993). A talidomida foi lançada no mercado alemão nos anos de 1956 e 1957 com os nomes comerciais *Grippex*® e *Contergan*® respectivamente, sendo indicada para o tratamento de irritabilidade, baixa concentração, ansiedade, tuberculose, enjôos, hipertireoidismo, doenças infecciosas, entre outras indicações (Lenz, 1988). Devido ao grande número de condições para as quais a talidomida se mostrava eficiente e ao seu baixo risco de toxicidade, a droga passou a ser vendida no mundo todo. A talidomida foi vendida como segura e sem a necessidade de prescrição médica, inclusive para mulheres grávidas (Lenz, 1988; Saldanha, 1994). No Brasil a comercialização teve início em março de 1958, e ao longo dos anos em que foi utilizada possuiu diversos nomes de marca (Oliveira et al., 1999).

1.1.1. Potencial Teratogênico da Talidomida

A partir de 1959 surgiram relatos de crianças nascidas com malformações de membros na Alemanha (Lenz, 1988), e em 1961, o Dr. Widukind Lenz sugeriu que essas malformações podiam ser resultado do uso de talidomida durante a gravidez. Do mesmo modo na Austrália, o Dr. William McBride relatou o aumento desses defeitos congênitos, correlacionando-os com o uso da talidomida (Miller & Strömmland, 1999). Foram relatados casos de embriopatia causada pelo uso de talidomida em cerca de 30 países (Saldanha, 1994). Em 1960 foram relatados os primeiros casos de malformações relacionadas ao uso de talidomida no Brasil, e

até 1965 o número registrado de vítimas foi de aproximadamente 300 indivíduos (Lenz, 1988), sendo este número provavelmente subestimado.

No final de 1961, a comercialização da talidomida foi suspensa na Alemanha e na Inglaterra, e mais tarde em outros países (Saldanha, 1994). Em agosto de 1962 o número de crianças relatadas com defeitos de membros já havia declinado significativamente (Shardein, 1993). No Brasil a talidomida continuou a ser vendida até 1962, sendo que segundo a Associação Brasileira de Portadores da Síndrome da Talidomida (ABPST) a medicação só foi de fato retirada do mercado em 1965 (Lenz, 1988; ABPST, 2017). Acredita-se que entre 10 e 15 mil crianças em todo o mundo tenham sido afetadas pela teratogênese da talidomida (Matthews & McCoy, 2003; Vargesson, 2009).

1.1.2. Potencial Terapêutico da Talidomida

Em 1964, o médico israelense Jacob Sheskin receitou talidomida como sedativo a um paciente com eritema nodoso da hanseníase (ENH), uma complicação inflamatória da hanseníase, e percebeu, após alguns dias, sua alta eficácia para o tratamento desta condição (Sheskin, 1965). A descoberta acidental do potencial terapêutico da talidomida gerou aumento do interesse de seu uso para outras condições inflamatórias, além de instigar estudos de investigação de outras propriedades terapêuticas do fármaco (Oliveira et al., 1999).

A talidomida voltou, a partir de então, a ser utilizada em diversos países ao redor do mundo, principalmente devido às suas propriedades antiinflamatória, imunomoduladora e antiangiogênica (Sampaio et al., 1991; D'Amato et al., 1994). No Brasil, seu uso foi aprovado pelo Ministério da Saúde a partir da segunda metade da década de 1960 para o tratamento do ENH (Oliveira et al. 1999; ABPST, 2017). Com a volta da circulação do fármaco, mesmo sendo criadas regulamentações para a sua distribuição, novos casos de embriopatia por talidomida vêm sendo registrados desde 1969 até os dias atuais (Castilla et al., 1996; Schuler-Faccini et al., 2007; Vianna et al., 2013).

A talidomida é utilizada atualmente no tratamento de diversas doenças, das quais destacam-se: doenças dermatológicas, crônico-degenerativas, relacionadas

à síndrome da imunodeficiência adquirida (AIDS) e diversos tipos de câncer (Miller & Strömmland, 1999; Oliveira et al., 1999; Chen et al., 2010). No Brasil, seu uso é aprovado para o tratamento de ENH, mieloma múltiplo, doença do enxerto versus hospedeiro, reações ulcerativas da AIDS e lúpus eritematoso sistêmico, sendo a dispensação do medicamento regulada pela Agência Nacional de Vigilância Sanitária sob regras rígidas (ANVISA) (RDC nº11, 22 de março de 2011).

1.2. Embriopatia por Talidomida

Análises de dados epidemiológicos mostraram que o período de sensibilidade à talidomida é entre o 34º e o 50º dia após a última menstruação, ou de 20 a 36 dias após a fertilização (Smithells & Newman, 1992; Miller & Strömmland, 1999). Estima-se que pequenas doses de talidomida (100mg) já são capazes de causar defeitos de membros (Penna et al., 1998; Vargesson, 2009). Cerca de 20 a 50% dos bebês de gestantes que fizeram uso de talidomida durante o período de sensibilidade desenvolveram a embriopatia (Newman, 1986). A mortalidade de bebês nascidos com embriopatia por talidomida é estimada em 40% antes do primeiro ano, devido principalmente a defeitos cardíacos e renais (Smithells & Newman, 1992; Vargesson, 2009).

Dentre as malformações causadas pelo efeito teratogênico da talidomida, as mais conhecidas são as de membros, sendo que estas afetam aproximadamente 90% dos sobreviventes (Kowalski et al., 2015). Os defeitos são bilaterais, mas não necessariamente simétricos e membros opostos são afetados de forma desigual (Newman, 1986; Pergament et al., 1997). Os membros inferiores e superiores têm alterações pré-axiais e intercalares, sendo as extremidades superiores geralmente acometidas por perda de dígitos e as inferiores por polidactilia e sindactilia (Newman, 1986).

Praticamente todos os órgãos do corpo podem ser afetados. Estima-se que malformações em outros órgãos ocorrem de forma simultânea às de membros, em cerca de 18% dos casos (Vargesson, 2009). Malformações de orelha, perda auditiva, anormalidades oculares e paralisia facial são frequentemente

identificadas (Smithells, 1973; Newman, 1986; Smithells & Newman, 1992). Hemangioma na linha média da face e lábio e/ou palato fendidos são mais relatados em vítimas da talidomida do que na população em geral (Smithells & Newman, 1992). Dentre as anormalidades de órgãos internos, malformações renais aparecem em cerca de 14% dos casos e malformações cardíacas em cerca de 8%, sendo a última a principal causa de morte (Smithells, 1973; Smithells & Newman, 1992; Miller & Strömland, 1999).

Manifestações tardias na TE são pouco relatadas, uma vez que a maioria dos indivíduos é avaliada apenas ao nascimento. Estudos apontam apenas para uma maior prevalência de distúrbios psicológicos e início precoce de doenças cardiovasculares nesses indivíduos (Imai et al., 2014; Kowalski et al., 2015).

1.3. Mecanismos de Teratogênese da Talidomida

Ao longo dos últimos 60 anos, diversos estudos têm investigado possíveis mecanismos pelos quais a talidomida desempenha sua propriedade teratogênica, entretanto, suas vias moleculares ainda não são totalmente entendidas. A identificação de tais mecanismos é extremamente difícil, visto que eles são resultado da interação entre diversos fatores ambientais e *background* genético da mãe e do embrião (Cassina et al., 2012).

Cerca de 30 hipóteses com o objetivo de explicar a ação teratogênica da talidomida foram geradas entre 1966 e 2003 (Hansen & Harris, 2004). As três hipóteses atualmente mais bem aceitas, estudadas a partir de modelos animais de coelhos, galinhas e *zebrafish*, postulam que a ação teratogênica da talidomida é devido à: (1) ligação com a proteína Cereblon (CRBN); (2) formação de espécies reativas de oxigênio (ROS) e estresse oxidativo; e (3) inibição da angiogênese (D'Amato et al., 1994; Parman et al., 1999; Hansen & Harris, 2004; Therapontos et al., 2009; Ito et al., 2010).

Ito et al. (2010) identificaram que a proteína CRBN é alvo de ligação da talidomida e que essa ligação impede a formação de um complexo de ubiquitinação do qual a proteína faz parte; devido a isso cria-se um desbalanço na expressão de genes importantes para o desenvolvimento de membros, como os

fatores de crescimento de fibroblastos 8 e 10 (*FGF8* e *FGF10*) por exemplo (Ito et al., 2010). Parman et al. (1999) demonstraram ainda que a talidomida induz oxidação no DNA de embriões de coelhos expostos intra-útero ao fármaco, e que essa oxidação não ocorre em embriões de camundongos, os quais são resistentes à teratogênese da talidomida. Foi sugerido que essa formação de ROS atuaria na teratogênese através da diminuição da atividade do fator nuclear *kappa-β* (NF- κ B) de se ligar ao promotor dos seus genes alvos: *FGF8*, *FGF10*, *msh homeobox 1 (MSX1)*, *sonic hedgehog (SHH)*, entre outros. Estes genes teriam então sua expressão diminuída e isso causaria, além de um distúrbio de regulação de outros genes, o comprometimento do desenvolvimento dos membros (Hansen & Harris, 2013). Estudos demonstraram, além disso, que a talidomida bloqueia a angiogênese nas córneas de coelhos e camundongos através da inibição de fatores de crescimento de fibroblastos (FGFs) e fator de crescimento do endotélio vascular (VEGF) (D'Amato et al., 1994; Kenyon et al., 1997). Therapontos et al. (2009) demonstraram que os membros são bastante suscetíveis à ação da talidomida, pois no período da janela teratogênica da talidomida há uma rede de vasos sanguíneos nessa região, como alta taxa de angiogênese (Therapontos et al., 2009).

1.4. Fenocópias da Embriopatia por Talidomida

Goldschmidt (1935) propôs o conceito de fenocópia a partir de estudos com larvas de *Drosophila sp.*, no qual demonstrou que muitos fenótipos visualizados em mutantes podiam ser produzidos através de exposição a agentes exógenos, sem modificações genéticas, como por exemplo através de tratamento térmico. Da mesma forma, Buchner (1958) após produzir, por hipóxia, malformações semelhantes a fenótipos mutantes em algumas espécies de anfíbios e aves, e Rapoport (1939) após produzir, por exposição a produtos químicos, anormalidades morfológicas em *Drosophila sp.*, sugeriram que cada malformação genética pode ser copiada por uma causa exógena, sendo essas malformações consideradas então fenocópias (Lenz, 1973).

Henke et al. (1941) propuseram a distinção entre verdadeiras fenocópias, nas quais o agente ambiental afeta a mesma fase de desenvolvimento do gene, e falsas fenocópias, nas quais apenas o resultado final é idêntico. Um ponto de vista semelhante ao sugerido por Hadorn (1955) (Lenz, 1973).

Algumas embriopatias e fetopatias provocadas por exposição a alguns fármacos apresentam consideráveis semelhanças com síndromes genéticas, e por isso, são consideradas fenocópias dessas condições. A investigação das bases moleculares dessas síndromes tem auxiliado no entendimento dos mecanismos teratogênicos de alguns fármacos, uma vez que estas síndromes são causadas geralmente por mutações em genes que codificam proteínas alvo de teratógenos ou então proteínas que participam das mesmas vias moleculares (Cassina et al., 2016).

Dentre as malformações que são visualizadas em indivíduos que foram expostos à talidomida, as mais comumente relatadas são os defeitos de redução de membros. Estes defeitos não são incomuns e são originados por diversas causas. Ainda que em animais sejam um indicador de exposição a teratógenos, são também muito frequentemente observados em síndromes genéticas (Gold et al., 2011). Existem algumas síndromes de herança monogênica cujo fenótipo apresentado se assemelha, em maior ou menor grau, com as malformações características da embriopatia por talidomida (Lenz, 1973; Smithells & Newman, 1992; Warren, 1999). Dentre as síndromes mais comumente citadas destacam-se a Síndrome de Roberts, a Síndrome de Holt-Oram e a Síndrome de Duane de raio radial (também chamada Síndrome de Okihiro) (Tabela 1).

1.4.1. Síndrome de Roberts

A síndrome de Roberts (RBS) é uma doença autossômica recessiva rara com grande variabilidade de expressão, tendo sido nomeada em 1919 por John B. Roberts a partir de um relato de irmãos afetados (Roberts, 1919). Devido aos defeitos dos membros na RBS serem semelhantes aos observados na embriopatia por talidomida, a síndrome é também conhecida como síndrome pseudotalidomida. Herrmann et al. (1969) relataram malformações semelhantes,

mas mais leves, sendo essas referidas como focomelia SC (Herrmann et al., 1969). Essas duas condições passaram a ser consideradas a mesma síndrome, com variação da expressão fenotípica (Van Den Berg & Francke, 1993).

Atualmente não há estimativas precisas de incidência da RBS, sendo que cerca de 150 a 200 casos de pessoas de diversas origens étnicas já foram relatados (Vega et al., 2010).

Aspectos clínicos característicos da RBS são restrição de crescimento pré e pós-natal, redução simétrica e bilateral dos membros, e anormalidades craniofaciais (Freeman et al., 1974; Van Den Berg & Francke, 1993). Pessoas afetadas têm diferentes graus de malformações envolvendo redução simétrica no número de dígitos, e encurtamento ou ausência dos ossos dos braços e pernas. Entre as malformações craniofaciais, microcefalia, hipertelorismo e fissura lábio-palatina são mais prevalentes (Van Den Berg & Francke, 1993). A gravidade das malformações faciais correlaciona-se com a gravidade das malformações de membros (Vega et al., 2010). A maioria dos casos de Síndrome de Roberts consiste na associação de tetrafocomelia com graves malformações faciais, geralmente levando a morte na primeira infância (Herrmann & Opitz, 1977).

Na RBS os cromossomos apresentam separação prematura do centrômero ou repulsão da heterocromatina perto dos centrômeros, especialmente nos cromossomos 1, 9 e 16, nos braços curtos dos acrocêntricos e no braço longo do cromossomo Y. A mitose nessas células é anormal em relação à duração da metáfase e progressão para anáfase, sendo isso associado a aneuploidias (German, 1979; Tomkins et al., 1979; Van Den Berg & Francke, 1993). A RBS é considerada uma síndrome de mutação mitótica que conduz a defeitos de desenvolvimento secundários (Jabs et al., 1991).

Vega et al. (2010) estabeleceram critérios clínicos para o diagnóstico da síndrome de Roberts baseados em uma coorte de 49 pacientes, incluindo 18 previamente confirmados com a doença através de análise genética. Os critérios clínicos incluem: retardo do crescimento, encurtamento simétrico dos membros, onde os membros superiores eram mais comumente afetados que os membros inferiores, e características faciais como microcefalia (Vega et al., 2010). O

diagnóstico confirmatório da RBS se baseia em testes de citogenética no sangue periférico de indivíduos com achados clínicos sugestivos (GeneReviews, 2017a).

O gene responsável pela RBS é chamado *N-acetiltransferase 2 de Estabelecimento de Coesão de Cromátides-Irmãs* (*Establishment of Sister Chromatid Cohesion N-acetyltransferase 2 - ESCO2*). Este gene está localizado no cromossomo 8p21.1, possui 11 éxons e abrange 30,3 kb. A proteína codificada por esse gene é uma acetiltransferase de 601 aminoácidos que possui dois domínios: um domínio acetiltransferase e um domínio dedo de zinco que se liga à cromatina. A função dessa proteína é o estabelecimento de coesão entre cromátides-irmãs durante a fase S do ciclo celular ou em caso de reparo de quebra de DNA dupla-fita (Hou & Zou, 2005; Schüle et al., 2005; Vega et al., 2005; Vega et al., 2010). Gordillo et al. (2008) demonstraram que a perda da atividade de acetiltransferase da proteína ESCO2 está associada à patogênese da RBS (Gordillo et al., 2008).

Atualmente, cerca de 28 mutações em *ESCO2* foram relatadas associadas à RBS, sendo a maioria *frameshift* e *nonsense*, resultando em uma proteína truncada ou em um mRNA instável (Schüle et al., 2005; Vega et al., 2005; Resta et al., 2006; Gordillo et al., 2008; Vega et al., 2010). Schüle et al (2005) sugerem que, dada a quase ausência de mutações *missense* em *ESCO2*, estas poderiam resultar em diferentes fenótipos clínicos ou estar sendo selecionadas negativamente em heterozigotos (Schüle et al., 2005). A correlação entre genótipo-fenótipo não é clara, sendo essa análise prejudicada pelo número limitado de pacientes com informação clínica disponível (Vega et al., 2010).

1.4.2. Síndrome de Holt-Oram

A síndrome de Holt-Oram (HOS), também chamada síndrome coração-mão, é uma síndrome autossômica dominante rara, descrita formalmente pela primeira vez por Mary Holt e Samuel Oram em 1960. Eles verificaram, a partir da análise de uma família de quatro gerações, a presença de polegar com três falanges como lesão esquelética predominante, e defeitos de septo-atrial

representando defeitos cardíacos (Holt & Oram, 1960; Temtamy & McKusick, 1978).

A HOS é caracterizada por alta penetrância e expressividade variável de malformações múltiplas, incluindo anormalidades nos membros superiores, defeitos cardíacos congênitos e/ou anormalidades de condução cardíaca (Holt & Oram, 1960). Existe alta variabilidade intra e interfamiliar de apresentação clínica da HOS (Newbury-Ecob et al., 1996).

A anomalia de membros superiores sempre é presente, podendo ser bilateral, simétrica ou assimétrica, ou unilateral, envolvendo estruturas de raio radial, como os ossos da região tenar, carpo e rádio. A ausência, desenvolvimento incompleto ou anormal, e ainda a fusão desses três ossos produzem uma ampla gama de fenótipos. Anomalias cardíacas, presentes em 75% dos indivíduos com HOS, incluem defeitos únicos, múltiplos ou mesmo ausência dos septos ventricular e atrial, além da ocorrência de arritmia cardíaca e vasos periféricos hipoplásicos (Holt & Oram, 1960; Holmes, 1965; Smith et al., 1979; Basson et al., 1994). A correlação entre a gravidade de defeitos cardíacos e de membros não foi estabelecida (Barisic et al., 2014).

O *locus* do gene associado à HOS foi mapeado no cromossomo 12q24.1 e mutações nesse gene, chamado *Fator de transcrição T-box 5 (T-box transcription factor 5 - TBX5)*, foram identificadas como causadoras da HOS (Basson et al., 1997; Li et al., 1997). O gene *TBX5* possui 9 éxons e abrange 54,5 kb. A proteína resultante possui 518 aminoácidos e é membro da família de fatores de transcrição T-box. *TBX5* desempenha um papel importante no desenvolvimento do coração e membros superiores durante a embriogênese (Gibson-Brown et al., 1996; Basson et al., 1997; Li et al., 1997; Basson et al., 1999; Rodriguez-Esteban et al., 1999; Koshiba-Takeuchi et al., 1999, Rallis et al., 2003).

Mutações em heterozigose em *TBX5* afetam a função de sua proteína e são associadas à HOS (Basson et al., 1997; Li et al., 1997). A maioria das mutações patogênicas até agora descritas estão localizadas dentro do domínio T-box de ligação ao DNA, o qual é altamente conservado nos diferentes genes da família (Cross et al., 2000; Ghosh et al., 2001). Já foram relatadas mutações *nonsense*, *missense*, deleções, inserções e rearranjos neste gene (Li et al., 1997;

Basson et al, 1997; Fan et al., 2003). A maioria dessas mutações resulta em proteínas truncadas e haploinsuficiência (Basson et al, 1997). Algumas acarretam em defeitos na ligação da proteína com o DNA, defeitos na ativação da transcrição mediada por TBX5, redução da interação da proteína com outros cofatores e alteração de sua localização nuclear (Basson et al, 1999; Fan et al., 2003). Não há correlação entre o tipo de mutação em *TBX5* ou o local da mutação no domínio T-box e a expressividade da HOS (Brassington et al., 2003). Algumas variações podem resultar em ganho de função causando um fenótipo semelhante. Uma vez que a baixa ou a alta concentração da proteína produz o mesmo fenótipo, a função de TBX5 é considerada sensível à dosagem (Postma et al., 2008).

O diagnóstico da HOS é feito com base em achados clínicos consistentes e teste genético. Achados clínicos sugestivos são: malformações de membros superiores envolvendo o carpo, sendo essa característica presente em todos os indivíduos com HOS e podendo ser a única evidência da doença; malformação cardíaca, em especial envolvendo septo; defeitos de condução cardíaca; e história familiar (Poznanski et al., 1970; Basson et al., 1994). A testagem genética avalia, primeiramente, a presença de variante patogênica em heterozigose no gene *TBX5*, sendo que a presença de tais variantes em indivíduos que cumprem critérios diagnósticos rigorosos de HOS chega a 74% dos casos (McDermott et al., 2005).

A prevalência de HOS é estimada em 0,95 a cada 100.000 nascimentos, com base em um estudo epidemiológico realizado na Hungria (Elek et al., 1991). A prevalência média de HOS baseada em registros europeus é de 0,7 a cada 100.000 nascimentos (Barisic et al., 2014). A maioria dos relatos clínicos refere-se a nascidos vivos, crianças ou adultos, enquanto há pouca informação sobre mortes fetais, interrupções de gravidez após a detecção pré-natal de anomalias graves e de pacientes diagnosticados no período neonatal (Barisic et al., 2014).

1.4.3. Síndrome de Duane de Raio Radial – Síndrome de Okihiro

A síndrome de Duane, descrita por Duane em 1905, é uma anomalia congênita de motilidade ocular, consistindo em limitação da abdução ocular associada com retração do globo ocular e estreitamento da fissura palpebral em adução. Um dos primeiros relatos da síndrome de Duane ocorrendo juntamente com malformações radiais de membros superiores é o de Ferrell et al. (1966). Outras famílias foram relatadas por Temtamy et al. (1975) e Okihiro et al. (1977). Temtamy & McKusick (1978) chamaram esse conjunto de anomalias de síndrome de Duane de raio radial (DRRS). O termo síndrome de Okihiro foi usado pela primeira vez por Hayes et al. (1985).

A DRRS é uma doença autossômica dominante, cuja penetrância é de aproximadamente 95%, podendo ser menor dependendo da variante patogênica presente (GeneReviews, 2017b). Caracteriza-se por anomalias unilaterais ou bilaterais de Duane e malformações de raio radial, podendo incluir hipoplasia de ossos da região tênar, hipoplasia ou aplasia dos polegares, polegares trifalangeais ou duplicação do polegar (polidactilia preaxial), hipoplasia ou aplasia dos raios e ainda encurtamento e desvio radial dos antebraços. Podem ser visualizadas ainda anormalidades na comunicação interatrial, estenose anal, anormalidades renais, perturbação pigmentar, deficiência auditiva, malformações de orelha externa e assimetria facial. A variabilidade intrafamiliar é significativa (Ferrell et al., 1966; Okihiro et al., 1977; Temtamy & McKusick, 1978; MacDermot & Winter, 1987).

O gene *Fator de Transcrição Spalt 4* (*Spalt Like Transcription Factor 4 - SALL4*), da família de genes SPALT (SAL) que codificam fatores de transcrição dedos de zinco, é o único que tem sido associado à DRRS (Al-Baradie et al., 2002; Kohlhase et al., 2002). Ele está localizado no cromossomo 20q13.13-q13.2, possui 4 éxons e abrange 18,4 kb. A proteína codificada por este gene possui três domínios dedos de zinco duplos, um domínio dedo de zinco único próximo à extremidade C-terminal e outro domínio único próximo a extremidade N-terminal, esse último típico dessa família de proteínas em vertebrados (Kohlhase et al., 1996). Estudos em modelos animais demonstraram que a proteína SALL4 tem um importante papel, correlacionando-se com TBX5 e outras proteínas, no desenvolvimento de membros e coração (Neff et al., 2005; Harvey & Logan, 2006;

Koshiba-Takeuchi et al., 2006; Neff et al., 2011). Além disso, *SALL4* também desempenha atividade de manutenção das propriedades pluripotentes e auto-renováveis de células-tronco embrionárias e adultas (Xiong, 2014).

Algumas variantes encontradas na sequência do gene *SALL4*, são associadas à DRRS. Mutações em homozigose não foram relatadas em humanos, possivelmente porque isso resultaria em letalidade embrionária precoce. Todas as mutações relatadas são heterozigotas e estão localizadas nos éxons 2 e 3. Em relação ao tipo, as mutações são geralmente *nonsense*, duplicações curtas ou deleções, e levam a haploinsuficiência da proteína funcional (Al-Baradie et al., 2002; Kohlhase et al., 2002; Borozdin et al., 2004; Kohlhase et al., 2005; Borozdin et al., 2007). Foram relatados dois casos de mutações *missense* em torno da região codificadora para o *cluster* de domínios dedos de zinco (Miertus et al., 2006; Diehl et al., 2015).

O diagnóstico da DRRS baseia-se em achados fenotípicos sugestivos, referentes à anomalia de Duane e malformações de raio radial, e testagem genética, onde a sequência do gene *SALL4* é avaliada em busca de variantes patogênicas (GeneReviews, 2017b).

A prevalência da DRRS é desconhecida, em parte porque em muitos países os distúrbios relacionados à *SALL4* não têm sido bem diferenciados da síndrome de Holt-Oram (GeneReviews, 2017b). Entretanto, Kohlhase (2009) estima, com base em referências de testagem genética de *SALL4* na Alemanha entre 2002 e 2007, que é pouco provável que a DRRS ocorra com uma frequência superior a 1 a cada 200.000 nascimentos (Kohlhase, 2009).

Figura 1: Características das três principais síndromes das quais a embriopatia por talidomida é fenocópia.

Características/ Fenocópias	TE^a	RBS^b	DRRS/Síndrome de Okhiro^c	HOS^d
Etiologia	Teratogênica	Autossômica recessiva	Autossômica dominante	Autossômica dominante
Amelia	Sim, todos os membros	Não relatado	Não relatado	Não relatado
Focomelia	Sim, todos os membros	Sim, todos os membros	Não relatado	Membros superiores
Defeito de redução de membro pré-axial	Sim, todos os membros	Sim, todos os membros	Encurtamento e desvio radial dos antebraços; aplasia ou hipoplasia dos rádios; hipoplasia de ossos da região tênar	Ausência, hipoplasia ou malformação dos ossos da região tenar, carpo e rádio
Sindactilia	Membros superiores e inferiores	Membros superiores	Não relatado	Pré-axial
Polidactilia	Membros inferiores	Não relatado	Membros superiores	Pós-axial e central
Faciais	Lábio e/ou palato fendido; hemangioma da linha média	Lábio e/ou palato fendido; microcefalia; hipertelorismo	Hipertelorismo	Não são característicos
Oculares	Microftalmia; coloboma	Olhos proeminentes; opacificação da córnea	Anomalia de Duane unilateral ou bilateral	Não são característicos
Auriculares/Auditivos	Anotia; microtia; surdez	Lóbulos hipoplásicos	Surdez condutiva ou neurosensorial	Não são característicos

Figura 1: Características das três principais síndromes das quais a embriopatia por talidomida é fenocópia (Continuação).

Características/ Fenocópias	TE^a	RBS^b	DRRS/Síndrome de Okhiro^c	HOS^d
Cardíacos	Defeito de septo ventricular; coarctação da aorta; tetralogia de Fallot	Defeito de septo atrial ostium secundum; defeito de septo ventricular; defeitos de condução cardíaca	Defeito de septo ventricular; defeito de septo atrial; tetralogia de Fallot	Defeito de septo atrial; arritmias, defeito de septo ventricular
Genitália	Criptorquidia; utero bicórnio ou hipoplásico; pênis pequeno ou ausente; vagina ausente ou dupla	Criptorquidia; pênis grande em relação ao corpo; clitóris aumentado	Não relatado	Não relatado
Diagnóstico	Padrão de malformações e confirmação da exposição	Padrão de malformações e teste citogenético	Padrão de malformações e teste molecular	Padrão de malformações, teste molecular e história familiar de defeitos de septação ou condução cardíaca

TE - Embriopatia por talidomida; RBS - Síndrome de Roberts; DRRS - Síndrome de Duane de Raio Radial; HOS - Síndrome de Holt-Oram

^aNewman, 1986; Smithells & Newman, 1992; ^bGeneReviews (2017a); ^cGeneReviews (2017b); ^dGeneReviews (2017c)

1.5. Interações de ESCO2, TBX5 e SALL4 com a talidomida

Mönnich et al. (2011), a partir de um modelo de RBS em *zebrafish*, demonstraram que, apesar de ESCO2 não atuar diretamente no desenvolvimento de membros, a ausência da atividade dessa proteína está associada a defeitos de membros devido à falhas na proliferação e altos níveis de morte celular. Além disso, estudos têm demonstrado que ESCO2 tem papel regulatório na expressão de outros genes (Kim et al., 2008; Leem et al., 2011; Banerji et al., 2016), sendo possível considerar então que a ausência, o comprometimento ou o errôneo desempenho dessa atividade regulatória também poderiam levar ao fenótipo visualizado na RBS.

Uma das hipóteses mais bem aceitas atualmente sobre os mecanismos moleculares que levam a teratogênese da talidomida é a hipótese de ligação da talidomida à proteína CRBN. Ito et al. (2010) sugerem que a ausência da formação do complexo culina RING E3 ubiquitina-ligase - do qual CRBN faz parte - devido a ligação da talidomida, e a falta da atividade de ubiquitinação desse complexo, aumentariam substratos não ubiquitinados e isso causaria um desbalanço na expressão de alguns genes (Ito et al., 2010). Kim et al. (2011) demonstraram que ESCO2 é substrato de um complexo culina RING ubiquitina-ligase. Entretanto, não se sabe ainda qual complexo especificamente atua sobre a proteína, sendo possível, dessa forma, que o complexo previamente citado como afetado pela talidomida seja o que atue em ESCO2 (Kim et al., 2011).

TBX5 e SALL4 são duas proteínas que atuam no desenvolvimento de membros e também no desenvolvimento do coração (Li et al., 1997; Basson et al., 1999; Koshiba-Takeuchi et al., 1999; Rodriguez-Esteban et al., 1999; Rallis et al., 2003; Sakaki-Yumoto et al., 2006; Neff et al., 2011). TBX5 tem papel essencial na iniciação, especificação de identidade e crescimento dos membros superiores. Estudos em modelos animais demonstram que TBX5 regula a expressão de SALL4, e as duas proteínas interagem para regular o desenvolvimento dos membros e do coração (Harvey & Logan, 2006; Koshiba-Takeuchi et al., 2006).

Knobloch e Rüther (2008) demonstraram o envolvimento dos fatores de transcrição TBX5 e SALL4 na teratogênese da talidomida. A partir de

experimentos com embriões de galinha e fibroblastos embrionários humanos, observou-se que a talidomida reduz a transcrição dessas proteínas, sendo proposto que o *stress* oxidativo produzido pela talidomida afetaria sua expressão (Knobloch & Rüther, 2008).

CAPÍTULO II

JUSTIFICATIVA

2. JUSTIFICATIVA

A talidomida é um fármaco utilizado atualmente em todo o mundo para uma série de indicações, dentre as quais doenças dermatológicas, crônico-degenerativas, diversos tipos de câncer, entre outras. Entretanto, os mecanismos de ação da sua propriedade teratogênica, conhecida há mais de 60 anos, ainda não são completamente entendidos.

A exposição à talidomida durante a gravidez acarreta no nascimento de bebês com malformações em 20-50% dos casos. Sendo assim, uma grande parcela de embriões expostos não desenvolve fenótipo anormal. Estudos sugerem que alguns polimorfismos podem levar à suscetibilidade diferente em relação ao dano induzido por um teratígeno (Cassina et al., 2012). Com base nisso, torna-se evidente a necessidade de se buscar variantes genéticas que possam estar associadas à proteção ou aumento da suscetibilidade de certos indivíduos à teratogênese da talidomida.

Os defeitos de redução de membros, as principais malformações associadas à exposição talidomida, além de outras visualizadas na embriopatia por talidomida, são presentes também em indivíduos portadores de algumas síndromes genéticas, como a Síndrome de Roberts, de Holt-Oram e de Duane de raio radial (Síndrome de Okihiro). A análise das bases moleculares de síndromes genéticas das quais embriopatias produzidas por teratógenos são fenocópias têm se mostrado útil no entendimento dos mecanismos teratogênicos de outros fármacos (Cassina et al., 2016).

Dessa forma, a análise de genes importantes para o desenvolvimento do embrião, os quais são expressos durante a janela teratogênica da talidomida, por ela podem ser de alguma forma afetados e quando mutados são associados a fenótipos semelhantes aos observados na TE, pode auxiliar no entendimento do mecanismo teratogênico deste fármaco, bem como na compreensão da suscetibilidade diferencial no desenvolvimento das malformações e doenças observadas nos indivíduos afetados.

CAPÍTULO III

OBJETIVOS

3. OBJETIVOS

3.1. Objetivo Geral

Avaliar o papel de variantes genéticas em genes associados a síndromes genéticas de fenótipo semelhante à embriopatia por talidomida em indivíduos com embriopatia por talidomida.

3.2. Objetivos Específicos

- Investigar a presença de variantes genéticas (SNPs, deleções e inserções) nas sequências dos genes *ESCO2*, *TBX5* e *SALL4* em indivíduos com diagnóstico de embriopatia por talidomida;
- Estimar a frequência dessas variantes nesses indivíduos;
- Analisar o impacto das variantes identificadas nas proteínas através de preditores funcionais *in silico*;
- Tentar interpretar, a partir dos resultados encontrados, um possível papel destes genes e suas variantes na suscetibilidade genética para o desenvolvimento da embriopatia por talidomida;
- Avaliar um possível papel destes genes e suas variantes nas diferentes anomalias congênitas e doenças observadas nos indivíduos afetados com embriopatia por talidomida.

CAPÍTULO IV

ARTIGO I

The role of *ESCO2*, *TBX5* and *SALL4* genes in the susceptibility to thalidomide embryopathy

Artigo em preparação
Revista *Clinical Genetics*

The role of *ESCO2*, *TBX5* and *SALL4* genes in the susceptibility to thalidomide embryopathy

Julia do Amaral Gomes^{1,2}, Thayne Woycinck Kowalski^{1,2}, Lucas Rosa Fraga^{1,2}, Marilea Furtado Feira¹, Bruna Duarte Rengel¹, Maria Teresa Vieira Sanseverino^{1,2,3}, Lavínia Schuler-Faccini^{1,2,3} and Fernanda Sales Luiz Vianna^{1,2,3,4}

Authors' Affiliations

¹Postgraduate Program in Genetics and Molecular Biology, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.

²National Institute of Population Medical Genetics (INaGeMP), Porto Alegre, Brazil.

³Brazilian Teratogen Information Service (SIAT), Medical Genetics Service at the Porto Alegre Clinics Hospital, Porto Alegre, Brazil.

⁴Genomic Medicine Laboratory of Porto Alegre Clinics Hospital, Porto Alegre, Brazil.

Corresponding author:

Fernanda Sales Luiz Vianna

Brazilian Teratogen Information Service (SIAT) - Medical Genetics Service

Porto Alegre Clinics Hospital - Porto Alegre/RS - Brazil

Fax: 55 51 33598008

Telephone: 55 51 33598008

E-mail: fslvianna@gmail.com

The authors declare no conflicts of interest.

Acknowledgements

The authors acknowledge INAGEMP – National Institute of Population Medical Genetics (grant CNPq 573993/2008-4), FIPE/HCPA (10-0244) and Brazilian Association of People Affected by Thalidomide Syndrome - ABPST for the support provided for this project.

Abstract

Thalidomide is a drug largely used to treat several conditions; however, it causes thalidomide embryopathy (TE) in 20-50% of embryos exposed during pregnancy. Given the similarity of phenotypic characteristics, embryopathies caused by drugs exposure can present as phenocopies of genetic syndromes. The analysis of the molecular bases of these syndromes has helped in the understanding of some teratogenic mechanisms. In this study we sequenced coding exons, flanking introns, and untranslated regions of three genes associated with syndromes which TE is phenocopy - *ESCO2* gene (Roberts syndrome), *SALL4* gene (Duane radial ray syndrome/Okhiro syndrome) and *TBX5* gene (Holt-Oram syndrome) – in 29 individuals with TE, and performed *in silico* analyses of the variants found. The investigation of genetic variations with a slight impact on the expression or function of these genes could help to understand the genetic susceptibility to TE and to anomalies or diseases seen in affected individuals. Forty-one variants were identified, 11 (27%) in *ESCO2*, 14 (34%) in *SALL4* and 16 (39%) in *TBX5* genes. Two of them have not been previously described in genomic databases. Twelve (29%) were rare genetic variants. We compared the frequencies of our sample with data from the European population of 1000 Genomes Project. Allelic and genotypic frequencies were significantly different between the two groups only in one 3'UTR variant (rs62498042) of *ESCO2* gene ($p < 0.001$); however, there are no reports in literature of its functional role or association with increased susceptibility to any disease. Our *in silico* analysis suggested that splicing can be affected due this variant presence. Some variants here identified were previously associated with cardiovascular diseases or cardiac malformations. In this sample, there was no association between these variants and an increased risk for these conditions. *In silico* analyzes were performed to evaluate the potential of these variants for functional alterations in the protein or for changes in splicing sites and transcription factors and miRNAs binding sites. Many variants were considered to have such potential, however, it was not possible to correlate them with decay of mRNA, alternative isoforms or decrease of these genes expression. In this way, experimental assays and validation of these variants are needed in order to establish their actual role in the protein activities.

Key words: Duane radial ray syndrome, Holt-Oram syndrome, Okhiro syndrome, Phenocopy, Roberts syndrome, Susceptibility, Thalidomide embryopathy

1. Introduction

Thalidomide is a drug used worldwide for several diseases - dermatological, chronic-degenerative, AIDS-related and cancers (Miller and Strömmland, 1999; Oliveira et al., 1999; Chen et al., 2010) - due to its anti-inflammatory, immunomodulatory and antiangiogenic properties (Sampaio et al., 1991; D'Amato et al., 1994). However, the mechanisms of its teratogenicity are not yet fully elucidated.

It is known that 20-50% of the embryos exposed to thalidomide are affected and develop the characteristic malformations of thalidomide embryopathy (TE). Thus, a large number of exposed embryos did not develop an abnormal phenotype in the 1960's (Newman, 1986). The individual genetic background plays an important role in this susceptibility and it is suggested that some polymorphisms may lead to differential susceptibility to teratogen-induced damage (Cassina et al., 2012).

Phenocopies are phenotypes from conditions caused by non-genetic agents (i.e. environmental agents, chemicals, drugs, among others) that resemble those phenotypes caused by mutations (Lenz, 1973). Some embryopathies caused by exposure to drugs are phenocopies of genetic syndromes. The analysis of the molecular bases of some syndromes has been shown efficient in the understanding of teratogenic mechanisms. These phenocopies are usually caused by mutations associated with genes which encoding proteins that are teratogen targets or that participate of same molecular pathways (Cassina et al., 2017).

Due to phenotypic similarity, including limbs, craniofacial and cardiac anomalies, TE can be considered a phenocopy of three genetic syndromes: Roberts syndrome, Holt-Oram syndrome and Duane radial ray syndrome (Okiihiro syndrome). Roberts syndrome (RBS; OMIM #268300), also called Pseudothalidomide syndrome, is a rare autosomal recessive disease caused by mutations in *ESCO2* gene. *ESCO2* protein is an acetyltransferase responsible for cohesion of sister chromatids during the cell division (Schüle et al., 2005; Vega et al., 2005). The Holt-Oram syndrome (HOS; OMIM #142900) and Duane radial ray syndrome (DRRS; OMIM #607323) are both autosomal dominant conditions that are caused by mutations in *TBX5* and *SALL4* genes, respectively. These two

genes encode transcription factors that interact and act on limbs and heart development (Li et al., 1997; Al-Baradie et al., 2002; Harvey and Logan, 2006; Koshiba-Takeuchi et al., 2006).

Since individual genetic variability is suggested as important for thalidomide teratogenesis susceptibility, and given the phenotypic similarity between individuals with variations in the three genes cited and those individuals with TE, the investigation of genetic variants of slight impact in these genes seems to be promising and quite relevant, both to the TE susceptibility as to phenotypic variability in affected individuals. Thus, in order to evaluate the role of *ESCO2*, *TBX5* and *SALL4* genes in thalidomide teratogenesis, here we sequenced these three genes in individuals with TE and evaluated the variants found.

2. Materials and Methods

2.1. Ethical Aspects

This study was approved by the Research Ethics Committee of the Clinics Hospital of Porto Alegre (number 10-0244). The methods were carried out in accordance with the approved guidelines.

2.2. Sample

Twenty-nine subjects affected by TE were included in this study. The subjects were recruited through the Brazilian Association of People Affected by Thalidomide Syndrome (ABPST) and signed the informed consent. All were previously examined by experts in TE and differential diagnosis. Gender, age and date and place of birth were also collected from subjects. Some of the individuals here included were previously evaluated as congenital anomalies in order to confirm the compatible phenotype of TE (Kowalski et al., 2015). Telephone calls were used to evaluate the other individuals, given the distance of their residences from ABPST and difficulty to a follow up in person.

For comparison purposes, we use as reference group the European population (503 individuals) from 1000 Genomes Project.

2.3. Molecular Analysis

Saliva was collected through Oragene-DNA OG-500 (DNA Genotek®) and DNA was obtained according to the manufacturer's instructions.

A gene panel including *ESCO2*, *TBX5* and *SALL4* genes was designed through Ion Ampliseq™ Designer tool (Thermo Fisher Scientific, USA) covering the coding region and 50bp of adjacent introns of each gene. The targeted gene sequencing was conducted by semiconductor chip Ion Torrent™ technology (Thermo Fisher Scientific, USA) through Ion Personal Genome Machine™ (PGM™) sequencer (Thermo Fisher Scientific, USA) at Clinics Hospital of Porto Alegre. The sequences obtained were analyzed through Ion Reporter v.5.2 (Thermo Fisher Scientific, USA).

2.4. In silico analyses

In silico analyses of the variants found through the sequencing were performed with the following bioinformatics tools: (1) Sorting Intolerant From Tolerant (SIFT) (J. Craig Venter Institute, USA) that compiles a dataset of functionally related protein sequences and uses a sequence homology-based approach to classify aminoacid substitution (Kumar et al., 2009); (2) Polymorphism Phenotyping v.2 (PolyPhen-2) (Harvard Medical School, USA) that uses sequence-based and structure-based predictive features to predicts the impact of aminoacid changes in a human protein (Adzhubei et al., 2010); (3) Mutation Assessor (Memorial Sloan-Kettering Cancer Center, USA) that evaluates the functional impact of amino acid changes through multiple sequence alignment, basing on the evolutionary conservation of aminoacid residues in a protein family (Reva et al., 2017); (4) Functional Single Nucleotide Polymorphism (F-SNP) (Queen's University, Canada) that compiles functional information about SNPs from 16 bioinformatics tools and databases, and predicts both functional effects in protein as disruption of regions known to be splice sites and transcriptional regulation sites (Lee and Shatkay, 2008); (5) Mutation Taster (California Institute of Technology, USA) that evaluates the difference between the physico-chemical characteristics of original versus new amino acids to predict functional effects, and at looking to donor and acceptor splice sequences predicts possible splicing

changes (Mutation Taster, 2017); (6) Human Splicing Finder (HSF) v.3 (Inserm, France) that evaluates donor and acceptor splice sequences, branchpoint sequence and auxiliary sequences - such as exonic and intronic splicing enhancers and exonic and intronic splicing silencers -, to predict possible splice changes (Desmet et al., 2009); and (7) miRBase (Manchester, United Kingdom) that predicts possible gains or losses of microRNAs (miRNAs) interactions by sequence complementarity (Griffiths-Jones et al., 2006).

All the analyses were performed in reference transcripts: ENST00000305188 to *ESCO2* gene, ENST00000310346 to *TBX5* gene and ENST00000217086 to *SALL4* gene. Human default parameter settings were used in all analyses.

2.5. Statistical Analysis

Hardy-Weinberg Equilibrium test was carried out for all polymorphisms. Allelic and genotypic frequencies of all variants identified were compared between TE and reference groups by Fisher's Exact Test using SPSS® v.18 (SPSS, IBM, USA) software. The same test was used to compare the frequencies of rare variants between individuals of TE group with different types of congenital anomalies or diseases and those individuals without the same conditions (anomalies and diseases observed in the TE group are described in Table 1). Besides that, it was also used to compare the frequencies of variants previously evaluated in cardiac pathologies between individuals of TE group with some cardiac anomalies and/or cardiovascular diseases and those without such conditions. A two-tailed p-value <0.05 was considered significant.

Linkage disequilibrium between variants was estimated using the Haploview program v.4.2 (IBM, USA), and the haplotypes were obtained with Bayesian algorithm in the PHASE v.2.1.1 tool (University of Chicago, USA).

3. Results

Twenty-nine Brazilian individuals with TE were included in our study. The characterization of these individuals, regarding congenital anomalies and late onset diseases are described in Table 1.

Using a targeted sequencing approach, we sequenced exons, flanking introns, and untranslated regions of *ESCO2*, *TBX5* and *SALL4* genes. The average depth of coverage for the variants found was 491.5x. A coverage higher than 20x was obtained in 97.73% of the bases sequenced, with 92.44% of the bases having a coverage higher than 100x. Table 2 and Figure 1 summarize all the variants found and the position of each variant, respectively.

The distribution of the variants was in accordance with the Hardy-Weinberg equilibrium, except for rs1824449, rs533417099, rs144484866 and rs62498042 of *ESCO2* gene. We have not observed differences in the allelic and genotypic frequencies of the variants between TE and reference group (Table 3), with exception of the variant rs62498042 in *ESCO2* gene (presented below). In addition, to rare variants (frequency <0.01) also were not observed differences between the frequencies in individuals with different types of congenital anomalies or diseases and those without the same conditions, both of TE group (data not shown).

Variants present in the same gene had their linkage disequilibrium evaluated and demonstrated an unequal segregation between them ($D' > 0.9$ and $LOD > 2$) (Supplementary Table 1). The haplotypes was inferred (Supplementary Table 2) and the analyses did not demonstrate an association between the most prevalent haplotypes in *ESCO2*, *SALL4* and *TBX5* genes and some anomaly or disease (data not shown).

ESCO2 gene

A total of 11 variants were found in *ESCO2* gene. The mean of variants per individual was 4 variants. Four variants were rare (rs182074852, rs1824449, rs533417099 and rs114667641) and one variant (c.*78_*82insAATAA) had not been reported in genomic databases (Table 2, 3 and 4).

When compared to the reference group, the frequency of the GG genotype of rs62498042, located in 3'UTR, was higher in TE group ($p < 0.001$) and the GA genotypes of the same variant was lower ($p < 0.001$). There are no reports in the literature of this variant functional role or its association with increased

susceptibility to any disease. *In silico* analysis with Mutation Taster tool suggests splicing affected due this variant presence.

Regarding the other results from *in silico* analyses (Table 4), Mutation Taster, which evaluated possible gain and loss of donor and acceptor sites and/or increase or decrease of these sequences functions, considered more variants in this gene affecting splicing comparing to HSF tool, which evaluated possible gain or loss of donor and acceptor sites, branchpoint and auxiliary sequences sites. At looking at the gene sequence, none of the variants seems to be located in sequences responsible for alternative splicing that generates the isoforms 1 or 2 of the ESCO2 protein. The rs533417099 in the 3'UTR was predicted to affect the possible binding of two miRNAs: hsa-miR-606 and hsa-miR-3149.

SALL4 gene

We identified 14 variants in *SALL4* gene; the mean of variants per individual was 4 variants. Three variants were rare (rs6013281, rs140110863 and rs150300174) (Table 2, 3 and 4).

Regarding the results from *in silico* analyses (Table 4), twelve variants were considered with potential to affect splicing sites by Mutation Taster, HSF and/or F-SNP, which aggregated the evaluation of 4 other splicing prediction tools (ESEfinder, ESRSearch, PESX, RESCUE_ESE) to makes these suggestions. At looking at the gene sequence, none of all these variants appear to be located in the regions responsible for alternative splicing that generates the isoforms 1 or 2 of the *SALL4* protein. The rs150300174 in the 3'UTR was predicted to affect the possible binding of three miRNAs on *SALL4* mRNA: hsa-miR-5095, hsa-miR-1254 and hsa-miR-5689.

TBX5 gene

We identified 16 variants in *TBX5* gene; the mean of variants per individual was 5 variants. Five variants were rare (rs186960328, rs571924700, rs78344365, rs147405081 and rs12426660) and one (c.420G>A) had not been reported in genomic databases (Table 2, 3 and 4). One rare variant (rs147405081) has been previously associated with a cardiac malformation (bicuspid aortic valve)

(Bonachea et al., 2014). This variant was also considered here as deleterious by PolyPhen and Mutation Taster and as possibly affecting splicing by HSF and Mutation Taster. In TE sample, two individuals had this variant; one of them presents angina.

Regarding the other results from *in silico* analyses (Table 4), the synonymous variant c.420G>A was considered deleterious by Mutation Taster. This variant is not located in a protein domain and not appears to be located in sequences responsible for alternative splicing of isoforms 1, 2 and 3 of TBX5 protein. One intronic variant (rs185924249) was also considered deleterious by Mutation Taster. This variant also not appears to be located in sequences responsible for TBX5 isoforms and there are no reports in literature relating it with any abnormal phenotype or pathology. A total of 14 variants were considered by Mutation Taster and F-SNP with potential to change splicing, and as the others, none appear to be in sequences that generate the TBX5 isoforms. The rs12426660, located in the 3'UTR, was predicted to affect the interaction of hsa-miR-6748-3p, hsa-miR-300, hsa-miR-1185-2-3p and hsa-miR-1185-1-3p miRNAs on TBX5 mRNA. Besides this, rs883079 also was pointed out to possibly affect the interaction of hsa-miR-492 miRNA.

4. Discussion

Over the almost past 60 years, studies have attempt to explain the mechanisms by which thalidomide plays its teratogenic action, however, the complete understanding has not been yet obtained. Our group has studied the genetic susceptibility to TE by analyzing genetic variation of Cereblon gene and angiogenesis-related genes (Vianna et al., 2013a; Kowalski et al., 2016; Vianna et al., 2016; Kowalski et al., 2017). In this study, we investigated a new hypothesis to better understand the thalidomide teratogenesis. We sequenced exons, flanking introns, and untranslated regions of *ESCO2*, *TBX5* and *SALL4* genes - responsible by syndromes which TE is phenocopy - and evaluated the variants found as their possible roles in increased susceptibility to TE and to different anomalies and diseases seen in affected individuals.

Forty-one variants were identified in a sample of 29 individuals with TE. Each individual presented, in mean, 13 variants. Twelve (29%) of all variants found were rare variants. Some studies have shown that the evaluation of rare genetic variants may help in the understanding of the susceptibility to some diseases (Geschwind and Flint, 2015; Kang et al., 2016; Lee et al.; 2016). However, none of the rare variants found in the genes here investigated were associated with increased susceptibility to any specific anomaly or disease.

Comparing the frequencies of variants found in *TBX5* and *SALL4* genes between our sample and the European population, we did not find significant differences that could suggest their role in increased risk to TE development. Eight variants found in *SALL4* gene had their role evaluated in Okhiro syndrome (DRRS) and/or ventricular septal defects by other studies: the rare synonymous variant rs6013281, the synonymous rs61737139, rs13038893, rs6021437, rs13043248 and rs17802735, and the missense rs6126344 and rs6091375. All these variants were considered benign, not associated with these conditions (Kohlhase et al., 2005; Wang et al., 2009). The rs883079, located in the 3'UTR of *TBX5* gene, has been associated to ventricular depolarization (Holm et al, 2010; Sotoodehnia et al., 2010; Pazoki et al., 2013). The rare missense variant rs147405081, also in *TBX5*, has been previously associated with bicuspid aortic valve (Bonachea et al., 2014). It is estimated the frequency of 8% of individuals with cardiac abnormalities among those affected by TE (Miller and Ströland, 1999). In our TE group, although some carriers of these variants have arrhythmia, extrasystole, angina or history of heart attack, there was no association between the presence of these variants and increased risk for any cardiovascular disease and cardiac anomaly.

In *ESCO2* gene, one variant located in a non-coding region had a differentiated frequency between TE and reference groups: rs62498042. The importance of non-coding variants in increasing susceptibility to diseases has been reported in several studies, since these variants can act in transcription, splicing and translation processes (Knight, 2005; Deplancke et al., 2016; Khurana et al., 2016). In our study, however, it was not possible to find an association between the presence of this variant with increased susceptibility to TE. The

variant rs4732748 of this gene was previously reported as benign, including in people with RBS (ClinVar and GeneReviews). For the other variants, there are no reports of association or functional role in phenotypes or diseases. In our study was also not possible make correlation of them and TE or adverse phenotypes.

Most of changes predicted by the bioinformatics tools were in splicing. The analyses indicated missense variants as tolerated, except one in *TBX5*. This variant was previously report as associated with bicuspid aortic valve, however, the subjects who carrier it in TE sample did not present such cardiac malformation.

The bioinformatics tools here used have different approaches to predict the functional effects of the variants; they evaluate different characteristics such as: physico-chemical characteristics of aminoacids, sequence homology, evolutionary conservation of aminoacid in protein families, protein structure, etc. To predict the effect of the variants on splicing, some tools also have different approaches: while a tool considered only changes in donor and acceptor splice sequence, others evaluated changes in theses sequences and in branchpoint sequence, exonic and intronic splicing enhancers and exonic and intronic splicing silencers. Given these different approaches, there were cases of discordant predictions among the bioinformatics tools. The same situation have been reported in the literature in studies that investigated the role of variants in other diseases or in studies that compared the performance and characteristics of different bioinformatics tools (Castellana and Mazza, 2013; Dabhi and Mistry, 2014; Jian et al., 2014; Marín-Martín et al., 2014; Tang et al., 2016). Because of that, it is necessary a careful interpretation of this analyses. Moreover, experimental assays and validation of all variants is required in order to understand their actual biological role.

Although some miRNAs have been suggested as having their interactions with mRNAs affected due 3'UTR variants, they were never described or experimentally validated as regulators of the three genes here investigated. One miRNA, hsa-miR-1254, was suggested to have its possible interaction with the *SALL4* mRNA affected by the rs150300174. This miRNA was previously shown to be expressed in human embryonic stem cells, possibly regulating genes with a role in stem cells differentiation, development and transcription regulation (Morin et

al., 2008), therefore, it is possible that this miRNA act in the mRNA of SALL4. However, more studies are needed in order to evaluate and to confirm this.

None of the bioinformatics tools used here to *in silico* analyses were specific for transcription factors binding predictions. Perhaps, because of this, just a few variants were believed to change transcription factor binding sites. In addition, in cases where such predictions have occurred, it were not mentioned which transcription factors specifically would have their binding affected. Based on this, an expanded analysis, including specific tools for this type of prediction, needs to be performed in order to understand these results.

Although we have had an approach of sequenced these three genes, we did not cover extensively the promoter regions, thus, regulatory variants may have been lost, resulting in underestimation of the actual contribution of these genes to TE susceptibility. Furthermore, the sample size is other complicating factor, which decreases the probability of the identification of smaller risk effects and the use of the more robust statistical analyses. In addition, we did not have a control group, which should include individuals who were exposed to thalidomide and did not develop TE. Thus, the results obtained here should be interpreted and extrapolated taking this into account.

The genes here investigated to produce proteins that are important during the embryo development. TBX5 and SALL4 are direct involved on limbs and heart development (Harvey and Logan, 2006; Koshiba-Takeuchi et al., 2006) and were previously demonstrated to have their expression affected in chicken embryos and human embryonic fibroblasts by thalidomide exposure (Knobloch and Rütter, 2008). Nevertheless, there is no experimental evidence of involvement between these three genes and TE. Moreover, there are no so much studies in humans evaluating all these genes sequence, making difficult the comparison and speculation about genetic variability and their impact in the proteins. It is important to mention that genetic variations are not necessary or sufficient to lead to TE, but the accumulation of them, even benign, could increase the susceptibility to TE. Investigations about variability in these three genes may help to understand not only the susceptibility to TE, but also the development of specific anomaly and disease seen in affected individuals.

References

1. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010;7(4):248-249.
2. Al-Baradie R, Yamada K, St Hilaire C, Chan WM, Andrews C, McIntosh N, Nakano M, Martonyi EJ, Raymond WR, Okumura S, Okihira MM, Engle EC. Duane radial ray syndrome (Okihira syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *Am J Hum Genet*. 2002;71:1195–1199.
3. Bonachea EM, Zender G, White P, Corsmeier D, Newsom D, Fitzgerald-Butt D, Garg V, McBride KL. Use of a targeted, combinatorial next-generation sequencing approach for the study of bicuspid aortic valve. *BMC Medical Genomics*. 2014;7:56.
4. Cassina M, Salviati L, Di Gianantonio E, Clementi M. Genetic susceptibility to teratogens: state of the art. *Reprod Toxicol*. 2012;34:186-191.
5. Cassina M, Cagnoli GA, Zuccarello D, Di Gianantonio E, Clementi M. Human teratogens and genetic phenocopies. Understanding pathogenesis through human genes mutation. *Eur J Med Genet*. 2017;60(1):22-31.
6. Castellana S, Mazza T. Congruency in the prediction of pathogenic missense mutations: state-of-the-art web-based tools. *Brief Bioinform*. 2013;14(4):448-459.
7. Chen M, Doherty SD, Hsu S. Innovative uses of thalidomide. *Dermatol Clin*. 2010;28(3):577-586.
8. ClinVar. NCBI. <https://www.ncbi.nlm.nih.gov/clinvar/variation/21240/>. Published June 14, 2016. Accessed February 18, 2017.
9. D'Amato RJ, Loughnan MS, Flynn E, Folkman J. Thalidomide is an inhibitor of angiogenesis. *Proc Natl Acad Sci USA* 1994;91:4082-4085.
10. Dabhi B, Mistry KN. In silico analysis of single nucleotide polymorphism (SNP) in human TNF- α gene. *Meta Gene*. 2014;2:586-595.
11. Deplancke B, Alpern D, Gardeux V. The genetics of transcription factor DNA binding variation. *Cell*. 2016;166(3):538-554.
12. Desmet FO, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37(9):e67.
13. Gordillo M, Vega H, Jabs EW. Gene Reviews. NCBI. <https://www.ncbi.nlm.nih.gov/books/NBK1153/>. Published April 18, 2006. Accessed February 18, 2017.
14. Geschwind DH, Flint J. Genetics and genomics of psychiatric disease. *Science*. 2015;349(6255): 1489–1494.
15. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res*. 2006;34(Database issue):D140-144.
16. Hansen JM, Harris C. A novel hypothesis for thalidomide-induced limb teratogenesis: redox misregulation of the NF-kappaB pathway. *Antioxid Redox Signal*. 2004;6(1):1-14.
17. Harvey MP, Logan O. Sall4 acts downstream of tbx5 and is required for pectoral fin outgrowth. *Development*. 2006;133(6):1165-1173.

18. Holm H, et al. Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet.* 2010;42:117–122.
19. Ito T, Ando H, Suzuki T, Ogura T, Hotta K, Imamura Y, Yamaguchi Y, Handa H. Identification of a primary target of thalidomide teratogenicity. *Science.* 2010;327(5971): 1345-1350.
20. Jian X, Boerwinkle E, Liu X. In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res.* 2014;42(22):13534-13544.
21. Kang G, Bi W, Zhang H, Pounds S, Cheng C, Shete S, Zou F, Zhao Y, Zhang JF, Yue W. A Robust and Powerful Set-Valued Approach to Rare Variant Association Analyses of Secondary Traits in Case-Control Sequencing Studies. *Genetics.* 2016;pii: genetics.116.192377.
22. Khurana E, Fu Y, Chakravarty D, Demichelis F, Rubin MA, Gerstein M. Role of non-coding sequence variants in cancer. *Nat Rev Genet.* 2016;17(2):93-108.
23. Knight JC. Regulatory polymorphisms underlying complex disease traits. *J Mol Med (Berl).* 2005;83(2): 97–109.
24. Kohlhase J, Chitayat D, Kotzot D, Ceylaner S, Froster UG, Fuchs S, Montgomery T, Rösler B. SALL4 mutations in Okhiro syndrome (Duane-radial ray syndrome), acro-renal-ocular syndrome, and related disorders. *Hum Mutat.* 2005;26:176–183.
25. Koshiba-Takeuchi K, Arruda EP, Kathiriya IS, Mo R, Hui CC, Srivastava D, Bruneau BG. Cooperative and antagonistic interactions between Sall4 and Tbx5 pattern the mouse limb and heart. *Nat Genet.* 2006;38(2):175-183.
26. Kowalski TW, Sanseverino MT, Schuler-Faccini L, Vianna FS. Thalidomide embryopathy: Follow-up of cases born between 1959 and 2010. *Birth Defects Res A Clin Mol Teratol.* 2015;103(9):794-803.
27. Kowalski TW, Fraga LR, Tovo-Rodrigues L, Sanseverino MT, Hutz MH, Schuler-Faccini L, Vianna FS. New Findings in eNOS gene and Thalidomide Embryopathy Suggest pre-transcriptional effect variants as susceptibility factors. *Sci Rep.* 2016;6:23404.
28. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* 2009;4(7):1073-1081.
29. Lee PH, Shatkay H. F-SNP: computationally predicted functional SNPs for disease association studies. *Nucleic Acids Res.* 2008;36(Database issue):D820-824.
30. Lee S, Won S, Kim YJ, Kim Y, T2D-Genes Consortium, Kim BJ, Park T. Rare variant association test with multiple phenotypes. *Genet Epidemiol.* 2016;[Epub ahead of print].
31. Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis AR, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD. Holt-Oram syndrome is caused by mutations in TBX5, a member of the Brachyury (T) gene family. *Nat Genet.* 1997;15:21–29.
32. Marín-Martín FR, Soler-Rivas C, Martín-Hernández R, Rodríguez-Casado A. A Comprehensive In Silico Analysis of the Functional and Structural

- Impact of Nonsynonymous SNPs in the ABCA1 Transporter Gene. *Cholesterol*. 2014;2014:639751.
33. Miller MT, Strömland K. Teratogen update: Thalidomide: A review, with a focus on ocular findings and new potential uses. *Teratology*. 1999;60(5):306-321.
 34. Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA. Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res*. 2008;18:610-621.
 35. Mutation Taster. <http://www.mutationtaster.org/info/documentation.html>. Accessed January 15, 2017.
 36. Newman CG. The thalidomide syndrome: risks of exposure and spectrum of malformations. *Clin Perinatol*. 1986;13(3):555-573.
 37. Oliveira MA, Bermudez JAZ, de Souza ACM. Talidomida no Brasil: vigilância com responsabilidade compartilhada? *Caderno de Saúde Pública*. 1999;15(1):99-112.
 38. Parman T, Wiley MJ, Wells PG. Free radical-mediated oxidative DNA damage in the mechanism of thalidomide. *Nat Med*. 1999;5(5):582-585.
 39. Pazoki R, de Jong JS, Marsman RF, Bruinsma N, Dekker LR, Wilde AA, Bezzina CR, Tanck MW. SNPs Identified as Modulators of ECG Traits in the General Population Do Not Markedly Affect ECG Traits during Acute Myocardial Infarction nor Ventricular Fibrillation Risk in This Condition. *PLoS One*. 2013;8(2):e57216.
 40. Reva B, Antipin Y, Sander C. Determinants of protein function revealed by combinatorial entropy optimization. *Genome Biol*. 2007;8(11):R232.
 41. Sampaio EP, Sarno EN, Galilly R, Cohn ZA, Kaplan G. Thalidomide selectively inhibits tumor necrosis factor alpha production by stimulated human monocytes. *J Exp Med*. 1991;173(3):699-703.
 42. Schüle B, Oviedo A, Johnston K, Pai S, Francke U. Inactivating mutations in ESCO2 cause SC phocomelia and Roberts syndrome: no phenotype-genotype correlation. *Am J Hum Genet*. 2005;77:1117-1128.
 43. Sotoodehnia N, et al. Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet*. 2010;42(12):1068-1076.
 44. Tang R, Prosser DO, Love DR. Evaluation of Bioinformatic Programmes for the Analysis of Variants within Splice Site Consensus Regions. *Adv Bioinformatics*. 2016;2016:5614058.
 45. Therapontos C, Erskine L, Gardner E R, Figg W D, Vargesson, N. Thalidomide induces limb defects by preventing angiogenic outgrowth during early. *Proc Natl Acad Sci USA*. 2009;106(21):8573-8578.
 46. Vega H, Waisfisz Q, Gordillo M, Sakai N, Yanagihara I, Yamada M, van Gosliga D, Kayserili H, Xu C, Ozono K, Jabs EW, Inui K, Joenje H. Roberts syndrome is caused by mutations in ESCO2, a human homolog of yeast ECO1 that is essential for the establishment of sister chromatid cohesion. *Nat Genet*. 2005;37:468-470.
 47. Vianna FS, Fraga LR, Tovo-Rodrigues L, Tagliani-Ribeiro A, Biondi F, Maximino CM, Sanseverino MT, Hutz MH, Schuler-Faccini L.

- Polymorphisms in the endothelial nitric oxide synthase gene in thalidomide embryopathy. *Nitric Oxide*. 2013a;35:89-92.
48. Vianna FS, Schuler-Faccini L, Leite JC, de Sousa SH, da Costa LM, Dias MF, Morelo EF, Doriqui MJ, Maximino CM, Sanseverino MT. Recognition of the phenotype of thalidomide embryopathy in countries endemic for leprosy: new cases and review of the main dysmorphological findings. *Clin Dysmorphol*. 2013b;22:59–63.
49. Vianna FS, Kowalski TW, Tovo-Rodrigues L, Tagliani-Ribeiro A, Godoy BA, Fraga LR, Sanseverino MT, Hutz MH, Schuler-Faccini L. Genomic and in silico analyses of CRBN gene and thalidomide embryopathy in humans. *Reprod Toxicol*. 2016;66:99-106.
50. Wang B, Li L, Xie X, Wang J, Yan J, Mu Y, Ma X. Genetic variation of SALL4 (SALL4) in ventricular septal defect. *Int J Cardiol*. 2010;145(2):224-226.

Table 1. Clinical characterization of individuals with Thalidomide Embryopathy.

Characteristic	Number	%
Sex		
<i>Male</i>	18/29	62
Congenital Anomalies^a		
<i>Upper limbs</i>	17/18	94
<i>Lower limbs</i>	8/18	44
<i>Ears</i>	3/17	18
<i>Eyes</i>	5/20	25
<i>Heart</i>	2/20	10
<i>Genital tract</i>	1/20	5
<i>Urinary tract</i>	1/20	5
<i>Skeletal</i>	2/10	10
Late Diseases^a		
<i>Hearing deficiency</i>	9/17	53
<i>Visual deficiency</i>	13/17	76
<i>Cardiovascular disease</i>	6/17	35
<i>Psychological disturbance</i>	6/17	35

^aData not available for all 29 individuals of the TE sample.

Table 2. Variants indentified in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy^a.

Gene	Locus	Polymorphism	Class	Alleles	Position	Consequence	Protein	MAF
ESCO2	8:27631990	rs35430328	SNP	G/A	Exon 1 (UTR)			0.35 (A)
	8:27632143	rs2272730	SNP	C/T	Intron variant			0.42 (T)
	8:27634064	rs4732748	SNP	C/T	Exon 3	Missense	p.Ala80Val	0.11 (T)
	8:27634171	rs182074852	SNP	G/A	Exon 3	Missense	p.Asp116Asn	<0.01 (A)
	8:27641609	rs1824449	SNP	G/A	Intron variant			<0.01 (G)
	8:27649426	rs11402833	INS	-/T	Intron variant			0.43 (T)
	8:27661025	rs533417099	DEL	TATT/-	3'UTR			<0.01 (-)
	8:27661030	c.*78_*82insAATAA ^b	INS	-/AATAA	3'UTR			NA
	8:27661066	rs144484866	INS	-/AC	3'UTR			NA
	8:27661085	rs62498042	SNP	G/A	3'UTR			0.46 (G)
	8:27662444	rs114667641	SNP	A/T	3'UTR			<0.01 (T)
SALL4	20:50409207	rs11697572	SNP	G/A	Intron variant			0.04 (A)
	20:50409151	rs190593855	SNP	C/T	Intron variant			0.01 (T)
	20:50409117	rs146597375	SNP	A/G	Intron variant			0.02 (G)
	20:50408482	rs6013281	SNP	A/G	Exon 2	Synonymous		<0.01 (A)
	20:50408377	rs61737139	SNP	G/C	Exon 2	Synonymous		0.02(C)
	20:50407966	rs13038893	SNP	C/T	Exon 2	Synonymous		0.33 (T)
	20:50407669	rs140110863	SNP	G/A	Exon 2	Synonymous		<0.01 (A)
	20:50407502	rs6126344	SNP	A/C	Exon 2	Missense	p.Leu507Arg	0.33 (C)
	20:50407162	rs6021437	SNP	T/C	Exon 2	Synonymous		0.34 (C)
	20:50406985	rs13043248	SNP	G/A	Exon 2	Synonymous		0.12 (A)
	20:50406630	rs6091375	SNP	T/G	Exon 2	Missense	p.Ile798Leu	0.06 (G)
	20:50405502	rs17802735	SNP	C/G	Exon 3	Synonymous		0.13 (G)
	20:50400679	rs3171177	SNP	A/T	3'UTR			0.04 (T)
	20:50400307	rs150300174	SNP	A/G	Downstream			<0.01 (G)

Table 2. Variants identified in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy^a (Continued).

Gene	Locus	Polymorphism	Class	Alleles	Position	Consequence	Protein	MAF
	12:114846292	rs143563344	SNP	G/A	5'UTR			0.01 (A)
	12:114846244	rs186960328	SNP	C/T	Exon 1 (UTR)			<0.01 (T)
	12:114845506	rs1248046	SNP	T/C	Intron variant			0.42 (C)
	12:114843606	rs12423887	SNP	C/T	Intron variant			0.13 (T)
	12:114841450	rs571924700	SNP	G/T	Intron variant			<0.01 (T)
	12:114836468	c.420G>A ^b	SNP	G/A	Exon 5	Synonymous		NA
	12:114832754	rs185924249	SNP	A/G	Intron variant			0.01 (G)
<i>TBX5</i>	12:114832510	rs2236017	SNP	C/A	Intron variant			0.39 (C)
	12:114823187	rs2277377	SNP	G/T	Intron variant			0.31 (T)
	12:114804222	rs78344365	SNP	C/T	Intron variant			<0.01 (T)
	12:114804165	rs147405081	SNP	C/T	Exon 8	Missense	p.Val263Met	<0.01 (T)
	12:114793260	rs28730761	SNP	T/C	3'UTR			0.12 (C)
	12:114793240	rs883079	SNP	C/T	3'UTR			0.29 (C)
	12:114792953	rs12426660	SNP	A/G	3'UTR			<0.01 (G)
	12:114792525	rs143511878	SNP	G/A	3'UTR			0.01 (A)
	12:114792236	rs6489956	SNP	T/C	3'UTR			0.19 (T)

^aData obtained from Ensembl database (<http://ensembl.org>) – Genome Reference Consortium Human genome build 37 (also known as hg19).

^bNot described in genomic database.

MAF: minor allele frequency; NA: not available.

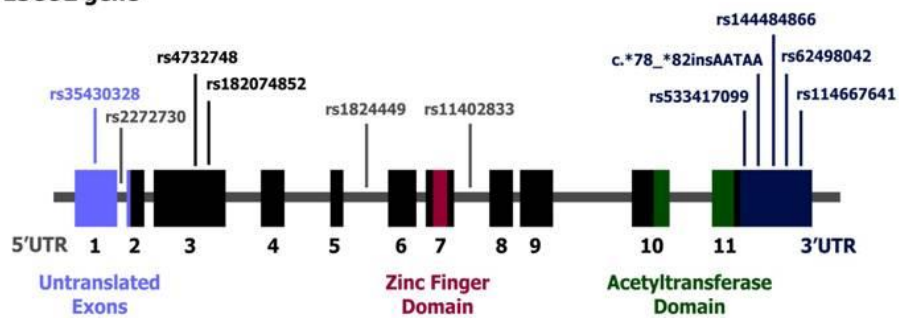
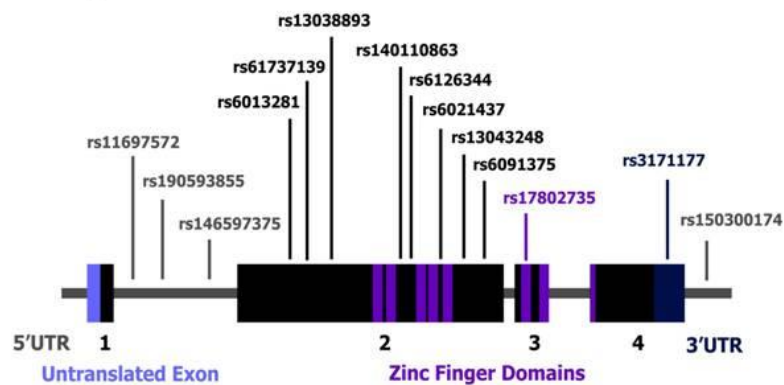
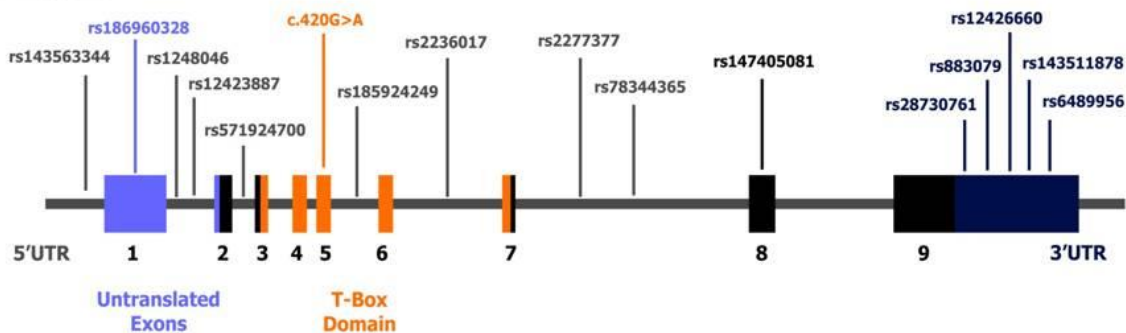
ESCO2 gene**SALL4 gene****TBX5 gene**

Figure 1: Position of variants found in *ESCO2*, *SALL4* and *TBX5* genes in people with thalidomide embryopathy.

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy.

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a	
			N	%	N	%
<i>ESCO2</i>	<i>rs35430328</i>	GG	16	55.2	214	42.5
		GA	11	37.9	222	44.2
		AA	2	6.9	133	13.3
		G	43	74.1	650	64.6
		A	15	25.9	350	35.4
	<i>rs2272730</i>	CC	5	17.2	175	34.8
		CT	16	55.2	229	45.5
		TT	8	27.6	99	19.7
		C	26	44.8	579	57.6
		T	32	55.2	427	42.4
	<i>rs4732748</i>	CC	26	89.7	401	79.7
		CT	3	10.3	96	19.1
		TT	0	0.0	6	1.2
		C	55	94.8	898	89.3
		T	3	5.2	108	10.7
	<i>rs182074852</i>	GG	28	96.6	503	100
		GA	1	3.4	0	0
		AA	0	0.0	0	0
		G	57	98.3	1006	100
		A	1	1.7	0	0
	<i>rs1824449^b</i>	GG	3	10.3	0	0
		AG	0	0.0	0	0
		AA	26	89.7	503	100
		G	6	10.3	0	0
		A	52	89.7	1006	100
	<i>rs11402833</i>	- -	7	24.1	174	34.6
		- T	15	51.8	230	45.7
		T T	7	24.1	99	19.7
		-	29	50.0	578	57.5
		T	29	50.0	428	42.5
	<i>rs533417099^b</i>	TATT TATT	27	93.2	503	100
		TATT -	1	3.4	0	0
- -		1	3.4	0	0	
TATT		55	94.8	1006	100	
-		3	5.2	0	0	
<i>c. *78_ *82insAATAA</i>	- -	26	89.7			
	- AATAA	3	10.3			
	AATAA AATAA	0	0.0			
	T	55	94.8			
	AATAA	3	5.2			

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a		
			N	%	N	%	
<i>ESCO2</i>	<i>rs144484866^b</i>	- -	18	62.1	NA		
		- AC	11	37.9			
		AC AC	0	0.0			
		T	47	81.0			
		AC	11	19.0			
	<i>rs62498042^{bc}</i>	GG	18	62.1	68	13.5	
		GA	8	27.6	325	64.6	
		AA	3	10.3	110	21.9	
		G	44	75.9	461	45.8	
		A	14	24.1	545	54.2	
	<i>rs114667641</i>	AA	28	96.6	503	100	
		AT	1	3.4	0	0	
		TT	0	0.0	0	0	
		A	57	98.3	1006	100	
		T	1	1.7	0	0	
	<i>SALL4</i>	<i>rs11697572</i>	GG	26	89.7	466	92.6
			GA	3	10.3	37	7.4
AA			0	0.0	0	0	
G			55	94.8	969	96.3	
A			3	5.2	37	3.7	
<i>rs190593855</i>		CC	28	96.6	496	98.6	
		CT	1	3.4	7	1.4	
		TT	0	0.0	0	0	
		C	57	98.3	999	99.3	
		T	1	1.7	7	0.7	
<i>rs146597375</i>		AA	28	96.6	486	96.6	
		AG	1	3.4	17	3.4	
		GG	0	0.0	0	0	
		A	57	98.3	989	98.3	
		G	1	1.7	17	1.7	
<i>rs6013281</i>		AA	0	0.0	0	0	
		AG	1	3.4	4	0.8	
		GG	28	96.6	499	99.2	
		A	1	1.7	4	0.4	
		G	57	98.3	1002	99.6	
<i>rs61737139</i>		GG	27	93.1	481	95.6	
		GC	2	6.9	22	4.4	
		CC	0	0.0	0	0	
		G	56	96.6	984	97.8	
		C	2	3.4	22	2.2	

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a	
			N	%	N	%
SALL4	rs13038893	CC	13	44.8	230	45.7
		CT	13	44.8	216	42.9
		TT	3	10.2	57	11.4
		C	39	67.2	676	67.2
		T	19	32.8	330	32.8
	rs140110863	GG	28	96.6	503	100
		GA	1	3.4	0	0
		AA	0	0.0	0	0
		G	57	98.3	1006	100
		A	1	1.7	0	0
	rs6126344	AA	14	48.3	223	44.3
		AC	11	37.9	223	44.3
		CC	4	13.8	57	11.4
		A	39	67.2	669	66.5
		C	19	32.8	337	33.5
	rs6021437	TT	14	48.3	222	44.1
		TC	11	37.9	224	44.5
		CC	4	13.8	57	11.4
		T	39	67.2	668	66.4
		C	19	32.8	338	33.6
	rs13043248	GG	21	72.4	394	78.3
		GA	8	27.6	102	20.3
		AA	0	0.0	7	1.4
		G	50	86.2	890	88.5
		A	8	13.8	116	11.5
	rs6091375	TT	25	86.2	440	87.5
		TG	4	13.8	61	12.1
GG		0	0.0	2	0.4	
T		54	93.1	941	93.5	
G		4	6.9	65	6.5	
rs17802735	CC	23	79.3	387	76.9	
	CG	6	20.7	104	20.7	
	GG	0	0.0	12	2.4	
	C	52	89.7	878	87.3	
	G	6	10.3	128	12.7	
rs3171177	AA	27	93.1	463	92	
	AT	2	6.9	39	7.8	
	TT	0	0.0	1	0.2	
	A	56	96.6	965	95.9	
	T	2	3.4	41	4.1	

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a	
			N	%	N	%
<i>SALL4</i>	<i>rs150300174</i>	AA	28	96.6	499	99.2
		AG	1	3.4	4	0.8
		GG	0	0.0	0	0
		A	1	1.7	1002	99.6
		G	57	98.3	4	0.4
<i>TBX5</i>	<i>rs143563344</i>	GG	28	96.6	489	97.2
		GA	1	3.4	14	2.8
		AA	0	0.0	0	0
		G	57	98.3	992	98.6
		A	1	1.7	14	1.4
	<i>rs186960328</i>	CC	27	93.1	503	100
		CT	2	6.9	0	0
		TT	0	0.0	0	0
		C	56	96.6	1006	100
		T	2	3.4	0	0
	<i>rs1248046</i>	TT	9	31.0	174	34.6
		CT	16	55.2	240	47.7
		CC	4	13.8	89	17.7
		T	34	58.6	588	58.4
		C	24	41.4	418	41.6
	<i>rs12423887</i>	CC	20	69.0	383	76.1
		CT	9	31.0	109	21.7
		TT	0	0.0	11	2.2
		C	49	84.5	875	87
		T	9	15.5	131	13
	<i>rs571924700</i>	GG	27	93.1	503	100
		GT	2	6.9	0	0
		TT	0	0.0	0	0
		G	56	96.6	1006	100
		T	2	3.4	0	0
	<i>c.420G>A</i>	GG	28	96.6		
		GA	1	3.4		
		AA	0	0.0		
		G	57	98.3		
A		1	1.7			
<i>rs185924249</i>	AA	28	96.6	495	98.4	
	AG	1	3.4	8	1.6	
	GG	0	0.0	0	0	
	A	57	98.3	998	99.2	
	G	1	1.7	8	0.8	

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a	
			N	%	N	%
<i>TBX5</i>	<i>rs2236017</i>	CC	4	13.8	76	15.1
		CA	12	41.4	245	48.7
		AA	13	44.8	182	36.2
		C	20	34.5	397	39.5
		A	38	65.5	609	60.5
	<i>rs2277377</i>	GG	17	58.6	239	47.5
		GT	10	34.5	219	43.5
		TT	2	6.9	45	8.9
		G	44	75.9	697	69.3
		T	14	24.1	309	30.7
	<i>rs78344365</i>	CC	28	96.6	NA	
		CT	1	3.4		
		TT	0	0.0		
		C	57	98.3		
		T	1	1.7		
	<i>rs147405081</i>	CC	27	93.1	503	100
		CT	2	6.9	0	0
		TT	0	0.0	0	0
		C	56	96.6	1006	100
		T	2	3.4	0	0
	<i>rs28730761</i>	TT	25	86.2	387	76.9
		TC	4	13.8	108	21.5
		CC	0	0.0	8	1.6
		T	54	93.1	882	87.7
		C	4	6.9	124	12.3
<i>rs883079</i>	CC	5	17.2	44	8.7	
	CT	12	41.4	199	39.6	
	TT	12	41.4	260	51.7	
	C	22	37.9	287	28.5	
	T	36	62.1	719	71.5	
<i>rs12426660</i>	AA	28	96.6	503	100	
	AG	1	3.4	0	0	
	GG	0	0.0	0	0	
	A	57	98.3	1006	100	
	G	1	1.7	0	0	
<i>rs143511878</i>	GG	28	96.6	495	98.4	
	GA	1	3.4	8	1.6	
	AA	0	0.0	0	0	
	G	57	98.3	998	99.2	
	A	1	1.7	8	0.8	

Table 3. Allelic and genotypic frequencies of variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Genotype/Allele	TE sample		EUR ^a	
			N	%	N	%
	<i>rs6489956</i>	TT	1	3.4	20	4
		TC	11	37.9	151	30
		CC	17	58.7	332	66
		T	13	22.4	191	18.9
		C	45	77.6	815	81.1

^a1000 Genomes database to Europeans

^bOut of Hardy-Weinberg equilibrium

^cStatistically significant

NA: not available.

Table 4. Functional prediction of the variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy.

Gene	Polymorphism	Position	Consequence	SIFT	PolyPhen	Mutation Assessor	F-SNP	Mutation Taster	HSF	miRBase
<i>ESCO2</i>	rs35430328	Exon 1					Transcriptional regulation affected	Splicing affected		
	rs2272730	Intron					Transcriptional regulation affected		ESE site broken	
	rs4732748	Exon 3	Missense	Tolerated	Benign	Medium		Splicing affected	New ESS site; ESE site broken	
	rs182074852	Exon 3	Missense	Tolerated	Benign	Low		Splicing affected	ESE site broken	
	rs1824449	Intron								
	rs11402833	Intron						Splicing affected		
	rs533417099	3'UTR						Splicing affected		Affects miRNA binding sites
	c.*78_*82insAATAA	3'UTR						Splicing affected		
	rs144484866	3'UTR						Splicing affected		
	rs62498042	3'UTR						Splicing affected		
	rs114667641	3'UTR						Splicing affected		

Table 4. Functional prediction of the variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Position	Consequence	SIFT	PolyPhen	Mutation Assessor	F-SNP	Mutation Taster	HSF	miRBase
	rs11697572	Intron						Splicing affected	ESE site broken	
	rs190593855	Intron						Splicing affected	ESE site broken	
	rs146597375	Intron						Splicing affected	ESE site broken	
	rs6013281	Exon 2	Synonymous				Splicing affected	Splicing affected	New donor site; New ESS site; ESE site broken	
	rs61737139	Exon 2	Synonymous							
	rs13038893	Exon 2	Synonymous				Splicing affected	Splicing affected		
SALL4	rs140110863	Exon 2	Synonymous							
	rs6126344	Exon 2	Missense	Tolerated	Benign	Medium	Splicing affected	Splicing affected		
	rs6021437	Exon 2	Synonymous				Splicing affected	Splicing affected		
	rs13043248	Exon 2	Synonymous					Splicing affected	New ESS site	
	rs6091375	Exon 2	Missense	Tolerated	Benign	Low	Splicing affected	Splicing affected	ESE site broken	
	rs17802735	Exon 3	Synonymous				Splicing affected	Splicing affected	New ESS site	
	rs3171177	3'UTR					Transcriptional regulation affected		ESE site broken	
	rs150300174	Downstream							New ESS site	Affects miRNA binding sites

Table 4. Functional prediction of the variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Position	Consequence	SIFT	PolyPhen	Mutation Assessor	F-SNP	Mutation Taster	HSF	miRBase
<i>TBX5</i>	rs143563344	5'UTR								
	rs186960328	Exon 1							New donor site; New ESS site; ESE site broken	
	rs1248046	Intron						Splicing affected		
	rs12423887	Intron						Splicing affected	ESE site broken	
	rs571924700	Intron						Splicing affected	New donor site	
	c.420G>A	Exon 5	Synonymous					Disease causing; Splicing affected	ESE site broken	
	rs185924249	Intron						Disease causing; Splicing affected		
	rs2236017	Intron						Transcriptional regulation affected	Splicing affected	
	rs2277377	Intron						Transcriptional regulation affected	Splicing affected	ESE site broken
	rs78344365	Intron								
	rs147405081	Exon 8	Missense	Tolerated	Possibly damaging	Low		Disease causing; Splicing affected	New ESS site	
rs28730761	3'UTR						Splicing affected			

Table 4. Functional prediction of the variants in *ESCO2*, *SALL4* and *TBX5* genes in individuals with Thalidomide Embryopathy (Continued).

Gene	Polymorphism	Position	Consequence	SIFT	PolyPhen	Mutation Assessor	F-SNP	Mutation Taster	HSF	miRBase
	rs883079	3'UTR							New ESS site	Affects miRNA binding sites
	rs12426660	3'UTR					Transcriptional regulation affected	Splicing affected	ESE site broken	Affects miRNA binding sites
	rs143511878	3'UTR						Splicing affected	ESE site broken	
	rs6489956	3'UTR						Splicing affected		

ESS site - Exonic Splicing Silencer; ESE site - Exonic Splicing Enhancers; donor site - identification of splice site at the 5' ends of exon-intron-exon.

Supplementary Table 1: Linkage disequilibrium of variants contained in the same gene (Continued).

	TBX5															
	rs6489956	rs143511878	rs12426660	rs883079	rs28730761	rs147405081	rs78344365	rs2277377	rs2236017	rs185924249	c.-420G>A	rs571924700	rs12423887	rs1248046	rs186960328	rs143563344
rs6489956	-															
rs143511878	1.0	-														
rs12426660	1.0	1.0	-													
rs883079	1.0*	1.0	1.0	-												
rs28730761	1.0	1.0	1.0	1.0	-											
rs147405081	0.099	1.0	1.0	1.0	1.0	-										
rs78344365	1.0	1.0	1.0	1.0	1.0	1.0	-									
rs2277377	1.0	1.0	1.0	1.0*	1.0	0.036	1.0	-								
rs2236017	0.613	1.0	1.0	0.531	0.275	1.0	1.0	1.0*	-							
rs185924249	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-						
rs01	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-					
rs571924700	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-				
rs12423887	1.0	1.0	1.0	0.289	0.297	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-			
rs1248046	0.565	1.0	1.0	0.407	1.0	1.0	1.0	0.523	0.218	1.0	1.0	1.0	1.0	-		
rs186960328	0.099	1.0	1.0	1.0	1.0	1.0*	1.0	0.036	1.0	1.0	1.0	1.0	1.0	1.0	-	
rs143563344	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	-

*High statistical significance (LOD>2)

Supplementary Table 2: Haplotypes of *ESCO2*, *SALL4* and *TBX5* genes.

<i>ESCO2</i>	N	%
GTCGAGA	26	90
ACCGAAA	13	45
GCCGAGA	7	24
GTCGGGA	4	14
GCTGAGA	2	7
ACCGAGA	2	7
others	4	3
<i>SALL4</i>	N	%
ACAAGCGATGTGCAA	18	62
ACAAGCACTGTGCAA	15	52
ACAAGTGATATGCAA	8	28
ACAGGCGATATGCAA	5	17
ACAACCGATATGCAG	3	10
ATAAGCGATAGGCAA	2	7
others	7	3
<i>TBX5</i>	N	%
GCTTGGGCTTCTCGGC	9	31
GCTTGGGAGTCTCGGC	6	21
ACTCGGGAGTCTCAGC	6	21
ACTCGGGAGTCTCGGC	5	17
GCTTGGGAGTCTTAGC	4	14
GCTCGGGAGTCTCAGC	4	14
GCTTGGGAGTCTCAGC	3	10
GCTTGGGATTCTCAGC	2	7
GCTTGAGAGTCTCAAC	2	7
GCTTTGGAGTCCTAGC	2	7
GCTTGGGCTTCTCGGT	2	7
ACTCGGGATTCTCAGC	2	7
GCTCGGGAGTCTCGGC	2	7
others	9	3

CAPÍTULO V

DISCUSSÃO

5. DISCUSSÃO

A talidomida tem sido foco de investigações desde a sua síntese, seja por sua propriedade teratogênica ou por suas propriedades terapêuticas. Apesar disso, o completo entendimento da propriedade teratogênica, conhecida há mais de 60 anos, ainda não foi obtido.

Sabe-se que somente uma parcela (de 20 a 50%) dos casos de exposição à talidomida levam ao desenvolvimento de malformações (McBride, 1961; Lenz, 1988) e que o *background* genético dos indivíduos tem um papel importante nessa suscetibilidade (Hansen & Harris, 2013). Entretanto, quais fatores levam a maior suscetibilidade para o desenvolvimento da embriopatia ou são associados à proteção, e quais estão associados a fenótipos clínicos mais leves e mais graves, são questões que ainda precisam ser respondidas.

Neste trabalho, nós investigamos uma nova hipótese na busca pelo melhor entendimento da teratogênese da talidomida. Visto que indivíduos com TE apresentam fenótipos muito semelhantes aos observados em pessoas com mutações nos genes *ESCO2*, *TBX5* e *SALL4*, e com base no fato de que a variabilidade de cada indivíduo afeta a susceptibilidade genética a danos causados por teratógenos, nós nos perguntamos se esses genes poderiam ter algum papel no aumento da suscetibilidade para o desenvolvimento da TE ou estariam relacionados à alguma anomalia congênita ou doença observada nos indivíduos afetados. Para responder a estas perguntas, foram sequenciados éxons, íntrons adjacentes e regiões não traduzidas destes três genes e, posteriormente, avaliou-se o papel das variantes encontradas através de preditores funcionais *in silico*.

Quarenta e uma variantes foram identificadas na nossa amostra, sendo doze destas raras. Apesar de estudos sugerirem que variantes raras podem aumentar a suscetibilidade para algumas doenças (Geschwind & Flint, 2015, Kang et al., 2016, Lee et al., 2016), neste estudo nenhuma das variantes raras foi associada com aumento do risco para TE ou para doenças e malformações presentes nos indivíduos afetados aqui avaliados.

A amostra avaliada neste estudo contou com 29 indivíduos afetados pela TE que foram caracterizados, quanto às suas anomalias congênitas e doenças de aparecimento tardio, nesse e em um estudo anterior do nosso grupo (Kowalski et al., 2015). No estudo citado, foi verificado que as doenças cardiovasculares, quando presentes, tiveram um início precoce em indivíduos com TE, comparando com a população brasileira sem anomalias congênitas. A frequência de anomalias cardíacas em indivíduos afetados por TE tem sido estimada em aproximadamente 8% em outros estudos (Miller & Ströland, 1999). Dois dos genes avaliados neste trabalho – *SALL4* e *TBX5* – estão diretamente envolvidos no desenvolvimento do coração (Harvey & Logan, 2006; Koshiba-Takeuchi et al., 2006). Ainda não há nenhuma evidência experimental demonstrando o envolvimento desses genes com o desenvolvimento das malformações ou doenças cardíacas, ou mesmo outros fenótipos, em indivíduos com TE. Entretanto, foi demonstrado que a talidomida reduz a expressão dessas duas proteínas (Knobloch & Rüther, 2008). Portanto, uma avaliação molecular de variantes nesses genes poderia levar a descoberta de maior risco de malformações cardíacas e desenvolvimento de doenças cardiovasculares em indivíduos expostos à talidomida. No presente estudo, algumas das variantes encontradas nesses dois genes foram previamente estudadas em doenças cardiovasculares e anomalias cardíacas (Wang et al., 2009; Holm et al., 2010; Sotoodehnia et al., 2010; Pazoki et al., 2013; Bonachea et al., 2014), sendo que algumas foram consideradas benignas e outras foram associadas a maior suscetibilidade para o desenvolvimento de tais condições. Em nossa amostra, contudo, não houve associação entre a presença dessas variantes e o aumento do risco para essas ou outras doenças e malformações.

Apenas uma das variantes aqui identificadas no gene *ESCO2* (rs4732748) teve descrição funcional na literatura. Esta variante foi considerada benigna (ClinVar, 2017; GeneReviews, 2017a), inclusive em indivíduos com RBS. As demais variantes deste gene não possuem qualquer relato de associação com algum fenótipo ou doença e, em nossos resultados, também não demonstraram correlação com anomalias ou doenças presentes nos indivíduos com TE.

Comparando a frequência das variantes encontradas em *TBX5* e *SALL4* entre a nossa amostra de indivíduos com TE e dados da população europeia do

Projeto 1000 Genomas, usada neste trabalho como grupo de comparação, não foram encontradas diferenças significativas que pudessem sugerir o papel dessas variantes em risco aumentado para a TE. No gene *ESCO2*, foi verificada diferença entre os grupos com TE e de referência quanto aos genótipos GG e GA da rs62498042, localizada na 3'UTR. O grupo com TE apresentou maior frequência de GG e menor frequência de GA ($p < 0,001$) em relação ao grupo de referência. Apesar da importância que variantes não codificantes têm no aumento da suscetibilidade de algumas doenças, dado seu potencial de alterar a transcrição, o splicing ou a tradução das proteínas (Knight, 2005; Deplancke et al., 2016, Khurana et al., 2016), neste estudo não foi possível correlacionar a variante rs62498042 com maior susceptibilidade à TE.

As análises *in silico* realizadas sugeriram que variantes *missense* nestes genes foram bem toleradas, exceto uma em *TBX5* (rs147405081), a qual foi previamente relatada na literatura como associada com a malformação cardíaca - válvula aorta bicúspide (Bonachea et al., 2014). Dois indivíduos eram portadores dessa variante na nossa amostra, entretanto, eles não apresentaram tal condição. Algumas variantes na 3'UTR demonstraram potencial de afetar, inibindo ou permitindo, a ligação de miRNAs nos mRNAs dos genes, entretanto, nenhum dos miRNAs citados foi descrito na literatura ou experimentalmente validado, segundo dados do miRTarBase (Chou et al., 2016), como regulador desses genes. Poucas variantes foram consideradas com potencial de alterar a ligação de fatores de transcrição nos genes aqui avaliados, entretanto, esta informação pode estar subestimada, uma vez que nenhuma das ferramentas de bioinformática utilizadas foi específica para este tipo de predição. Dessa forma, uma análise ampliada, incluindo ferramentas que avaliem possíveis alterações nesses sítios de ligação de fatores de transcrição, é necessária.

As ferramentas de bioinformática que foram utilizadas nas análises *in silico* se utilizaram de diferentes características para realizar as predições de impacto funcional e devido a estas diferentes abordagens, foram observadas predições discordantes entre as ferramentas. Este tipo de situação não é incomum e já foram relatados casos semelhantes a este em outros estudos que investigaram o papel funcional de variantes em doenças, e também em estudos que compararam

o desempenho e as características de diferentes ferramentas utilizadas para predições funcionais (Castellana & Mazza, 2013; Dabhi & Mistry, 2014; Jian et al., 2014; Marín-Martín et al., 2014; Tang et al., 2016).

Uma limitação deste estudo foi a região gênica avaliada. Não foram sequenciadas as regiões promotoras dos genes aqui avaliados, assim, variantes regulatórias podem ter sido perdidas, resultando em subestimação da real contribuição desses genes na susceptibilidade à TE. Além disso, o pequeno tamanho da amostra impede a identificação de efeitos de risco menores e a realização de análises estatísticas mais robustas. E, finalmente, não foi possível utilizar um grupo controle adequado, incluindo indivíduos que foram expostos à talidomida e não desenvolveram TE, uma informação considerada perdida devido ao tempo transcorrido desde a exposição.

Apesar das três síndromes genéticas aqui citadas serem conhecidas há mais tempo que a TE, os genes causadores dessas condições, bem como seus fenótipos, são pouco estudados. A abordagem utilizada aqui pôde trazer maior conhecimento destes genes, mas ainda assim, ensaios experimentais e a validação das variantes encontradas são necessários para o melhor entendimento de seus papéis biológicos. Os resultados aqui obtidos devem ser interpretados e extrapolados levando em consideração estas limitações. Independente disso, estas informações trarão maiores perspectivas para a compreensão da etiologia da TE e compreensão dessas síndromes.

CAPÍTULO VI

CONCLUSÕES E PERSPECTIVAS

6. CONCLUSÕES E PERSPECTIVAS

Os resultados obtidos nesse trabalho cumpriram os objetivos que haviam sido propostos:

*** Investigar a presença de variantes genéticas (SNPs, deleções e inserções) nas sequências dos genes *ESCO2*, *TBX5* e *SALL4* em indivíduos com diagnóstico de embriopatia por talidomida;**

*** Estimar a frequência dessas variantes nesses indivíduos;**

*** Tentar interpretar, a partir dos resultados encontrados, um possível papel destes genes e suas variantes na suscetibilidade genética para o desenvolvimento da embriopatia por talidomida;**

*** Avaliar um possível papel destes genes e suas variantes nas diferentes anomalias congênitas e doenças observadas nos indivíduos afetados com embriopatia por talidomida;**

Através do sequenciamento de éxons e regiões flanqueadoras dos genes citados, nós identificamos 39 variantes já previamente descritas e duas variantes novas nos 29 indivíduos com TE aqui avaliados, sendo elas SNPs, inserções e deleções, tanto comuns quanto raras. As frequências dessas variantes foram estimadas e comparadas com a população europeia de 503 indivíduos do projeto 1000 genomas. Entretanto, não foi possível demonstrar o envolvimento dessas variantes com maior suscetibilidade para o desenvolvimento da TE ou de doenças e malformações visualizadas em indivíduos afetados.

Nós ainda iremos concluir o sequenciamento, pelo método de Sanger, do éxon 5 e da região 3'UTR do gene *ESCO2* que tiveram uma cobertura inferior a 20x no sequenciamento de nova geração. Posteriormente, nós pretendemos avaliar as regiões promotoras desses genes.

*** Analisar o impacto das variantes identificadas nas proteínas através de preditores funcionais *in silico*;**

As variantes identificadas nos genes *ESCO2*, *SALL4* e *TBX5* foram avaliadas através de preditores funcionais *in silico*, e consideradas com potencial de afetar sítios de ligação de fatores de transcrição, sítios de interação de miRNAs e sítios de *splicing*. O impacto na função da proteína e provável associação com doenças também foi considerado para algumas das variantes identificadas. Entretanto, não foi possível aqui estabelecer o papel dessas variantes no decaimento do mRNA desses genes, ou na formação de isoformas alternativas ou ainda na diminuição da expressão dos genes que, juntamente com à exposição a talidomida, poderiam levar ao desenvolvimento da TE ou de anomalias e doenças em indivíduos afetados, como doenças cardiovasculares, por exemplo.

Nós iremos incluir novos preditores funcionais para a continuação das análises *in silico*, especialmente para análise de sítios de ligação de fatores de transcrição e miRNAs. Além disso, é possível que a análise dessas variantes, através de ensaios experimentais em modelos de galinhas e *zebrafish*, seja futuramente realizada por laboratórios colaboradores do nosso grupo, a fim de se obter a melhor compreensão de seus papéis biológicos no desenvolvimento de algum tipo de anomalia.

CAPÍTULO VII

REFERÊNCIAS BIBLIOGRÁFICAS

7. REFERÊNCIAS BIBLIOGRÁFICAS

- ABPST - Associação Brasileira de Portadores da Síndrome da talidomida, <http://www.talidomida.org.br/oque.asp> (Janeiro, 2017).
- Al-Baradie R Yamada K, St Hilaire C, Chan WM, Andrews C, McIntosh N, Nakano M, Martonyi EJ, Raymond WR, Okumura S, Okihiro MM, Engle EC (2002) Duane radial ray syndrome (Okihiro syndrome) maps to 20q13 and results from mutations in SALL4, a new member of the SAL family. *Am J Hum Genet* 71(5):1195-9.
- Banerji R, Eble DM, Lovine MK, Skibbens RV (2016) Esco2 Regulates cx43 Expression During Skeletal Regeneration in the Zebrafish Fin. *Dev Dyn* 245(1):7-21.
- Barisic I, Boban L, Greenlees R, Garne E, Wellesley D, Calzolari E, Addor MC, Arriola L, Bergman JE, Braz P, Budd JL, Gatt M, Haeusler M, Khoshnood B, Klungsoyr K, McDonnell B, Nelen V, Pierini A, Queisser-Wahrendorf A, Rankin J, Rissmann A, Rounding C, Tucker D, Verellen-Dumoulin C, Dolk H (2014) Holt Oram syndrome: a registry-based study in Europe. *Orphanet J Rare Dis* 25(9):156.
- Basson CT, Cowley GS, Solomon SD, Weissman B, Poznanski AK, Traill TA, Seidman JG, Seidman CE (1994) The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). *N Engl J Med* 330:885–891
- Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soultis J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE (1997) Mutations in human TBX5 [corrected] cause limb and cardiac malformation in Holt-Oram syndrome. *Nat Genet* 15:30–35.
- Basson CT, Huang T, Lin RC, Bachinsky DR, Weremowicz S, Vaglio A, Bruzzone R, Quadrelli R, Lerone M, Romeo G, Silengo M, Pereira A, Krieger J, Mesquita SF, Kamisago M, Morton CC, Pierpont ME, Müller CW, Seidman JG, Seidman CE (1999) Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc Natl Acad Sci USA* 96:2919–2924.
- Bonachea EM, Zender G, White P, Corsmeier D, Newsom D, Fitzgerald-Butt D, Garg V, McBride KL (2014) Use of a targeted, combinatorial next-generation sequencing approach for the study of bicuspid aortic valve. *BMC Medical Genomics* 7:56
- Borozdin W, Boehm D, Leipoldt M, Wilhelm C, Reardon W, Clayton-Smith J, Becker K, Mühlendyck H, Winter R, Giray O, Silan F, Kohlhase J (2004) SALL4 deletions are a common cause of Okihiro and acro-renal-ocular syndromes and confirm haploinsufficiency as the pathogenic mechanism. *J Med Genet* 41(9):e113.
- Borozdin W, Graham JM Jr, Böhm D, Bamshad MJ, Spranger S, Burke L, Leipoldt M, Kohlhase J (2007) Multigene deletions on chromosome 20q13.13-q13.2 including SALL4 result in an expanded phenotype of Okihiro syndrome plus developmental delay. *Hum Mutat.* 28(8):830.
- Brassington AM, Sung SS, Toydemir RM, Le T, Roeder AD, Rutherford AE, Whitby FG, Jorde LB, Bamshad MJ (2003) Expressivity of Holt-Oram syndrome is not predicted by TBX5 genotype. *Am J Hum Genet* 73:74–85.

- Cassina M, Salviati L, Di Gianantonio E, Clementi M (2012) Genetic susceptibility to teratogens: state of the art. *Reprod Toxicol* 34:186-91.
- Cassina M, Cagnoli GA, Zuccarello D, Di Gianantonio E, Clementi M (2017) Human teratogens and genetic phenocopies. Understanding pathogenesis through human genes mutation. *Eur J Med Genet* 60(1):22-31.
- Castellana S, Mazza T (2013) Congruency in the prediction of pathogenic missense mutations: state-of-the-art web-based tools. *Brief Bioinform* 14(4):448-459.
- Castilla EE, Ashton-Prolla P, Barreda-Mejia E, Brunoni D, Cavalcanti DP, Correa-Neto J, Degaldillo JL, Dutra MG, Felix T, Giraldo A, Juarez N, Lopez-Camelo JS, Nazer J, Orioli IM, Paz JE, Pessoto MA, Pina-Neto JM, Quadrelli R, Rittler M, Rueda S, Saltos M, Sánchez O, Schuler-Faccini L (1996) Thalidomide, a Current Teratogen in South America. *Teratology* 54:273-277.
- Chen M, Doherty SD, Hsu S (2010) Innovative uses of thalidomide. *Dermatol Clin*. 28(3):577-86.
- Chou CH, et al. (2016) miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. *Nucleic Acids Res*. 44(D1):D239-47.
- ClinVar, <https://www.ncbi.nlm.nih.gov/clinvar/variation/21240/> (Janeiro, 2017).
- Cross SJ, Ching YH, Li QY, Armstrong-Buisseret L, Spranger S, Lyonnet S, Bonnet D, Penttinen M, Jonveaux P, Mortier G, Leheup B, Van Ravenswaaij C, Gardiner CA (2000) The mutation spectrum in Holt-Oram syndrome. *J Med Genet* 37:785-7.
- D'Amato RJ, Loughnan MS, Flynn E, Folkman J (1994) Thalidomide is an inhibitor of angiogenesis. *Proc Natl Acad Sci USA* 91:4082-4085.
- Dabhi B, Mistry KN (2014) In silico analysis of single nucleotide polymorphism (SNP) in human TNF- α gene. *Meta Gene* 2:586-595.
- Diehl A, Mu W, Batista D, Gunay-Aygun M (2015) An atypical 0.73 MB microduplication of 22q11.21 and a novel SALL4 missense mutation associated with thumb agenesis and radioulnar synostosis. *Am J Med Genet A* 167, 1644-1649.
- Elek C, Vitéz M, Czeizel E (1991) Holt-Oram syndrome. *Orv Hetil* 132:73–74.
- Fan C, Duhagon MA, Oberti C, Chen S, Hiroi Y, Komuro I, Duhagon PI, Canessa R, Wang Q (2003) Novel TBX5 mutations and molecular mechanism for Holt-Oram syndrome. *J Med Genet* 40:e29.
- Freeman MV, Williams DW, Schimke RN, Temtamy SA, Vachier E, German J (1974) The Roberts syndrome. *Clin Genet* 5:1–16.
- GeneReviews, <https://www.ncbi.nlm.nih.gov/books/NBK1153/> (Janeiro, 2017a).
- GeneReviews, <http://www.ncbi.nlm.nih.gov/books/NBK1373/> (Janeiro, 2017b).
- GeneReviews, <https://www.ncbi.nlm.nih.gov/books/NBK1111/> (Janeiro, 2017c).
- German J (1979) Roberts' syndrome I. Cytological evidence for a disturbance in chromatid pairing. *Clin Genet* 16:441-7.
- Geschwind DH, Flint J (2015) Genetics and genomics of psychiatric disease, *Science* 349(6255): 1489–1494.
- Ghosh TK, Packham EA, Bonser AJ, Robinson TE, Cross SJ, Brook JD (2001) Characterization of the TBX5 binding site and analysis of mutations that cause Holt–Oram syndrome. *Hum Mol Genet* 10(18):1983-94.
- Gibson-Brown JJ, Agulnik SI, Chapman DL, Alexiou M, Garvey N, Silver LM, Papaioannou VE (1996) Evidence of a role for Tbox genes in the evolution of

- limb morphogenesis and the specification of forelimb/hindlimb identity. *Mech Dev* 56, 93-101.
- Gold NB, Westgate MN, Holmes LB (2011) Anatomic and etiological classification of congenital limb deficiencies. *Am J Med Genet A* 155A(6):1225-1235.
- Gordillo M, Vega H, Trainer AH, Hou F, Sakai N, Luque R, Kayserili H, Basaran S, Skovby F, Hennekam RC, Uzielli ML, Schnur RE, Manouvrier S, Chang S, Blair E, Hurst JA, Forzano F, Meins M, Simola KO, Raas-Rothschild A, Schultz RA, McDaniel LD, Ozono K, Inui K, Zou H, Jabs EW (2008) The molecular mechanism underlying Roberts syndrome involves loss of ESCO2 acetyltransferase activity. *Hum Mol Genet* 17:2172-80.
- Hansen JM, Harris C (2004) A novel hypothesis for thalidomide-induced limb teratogenesis: redox misregulation of the NF-kappaB pathway. *Antioxid Redox Signal* 6(1):1-14.
- Hansen JM, Harris C (2013) Redox control of teratogenesis. *Reprod Toxicol* 35:165-79.
- Harvey SA, Logan MP (2006) Sall4 acts downstream of tbx5 and is required for pectoral fin outgrowth. *Development* 133, 1165-1173.
- Hayes A, Costa T, Polomeno RC (1985) The Okhiro syndrome of Duane anomaly, radial ray abnormalities, and deafness. *Am J Med Genet* 22(2):273-80.
- Herrmann J, Feingold M, Tuffli GA, Opitz JM (1969) A familial dysmorphogenetic syndrome of limb deformities, characteristic facial appearance and associated anomalies: the "pseudothalidomide" or "SC-syndrome". *Birth Defects Orig Artic* 5:81-9.
- Herrmann J, Opitz JM (1977) The SC phocomelia and the Roberts syndrome: nosologic aspects. *Eur J Pediatr* 125:117-34.
- Holm H, et al. (2010) Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* 42:117-122.
- Holmes LB (1965) Congenital heart disease and upper-extremity deformities: A report of two families. *N Engl J Med* 272:437-44.
- Holt M, Oram S (1960) Familial heart disease with skeletal malformations. *Br Heart J* 22:236-242.
- Hou F, Zou H (2005) Two human orthologues of Eco1/Ctf7 acetyltransferases are both required for proper sister-chromatid cohesion. *Mol Biol Cell* 16:3908e18.
- Imai K, Iida T, Yamamoto M, Komatsu K, Nukui Y, Yoshizawa A (2014) Psychological and mental health problems in patients with thalidomide embryopathy in Japan. *Psychiatry Clin Neurosci* 68:479-486.
- Ito T, Ando H, Suzuki T, Ogura T, Hotta K, Imamura Y, Yamaguchi Y, Handa H. (2010) Identification of a primary target of thalidomide teratogenicity. *Science* 327(5971): 1345-1350.
- Ivanov D, Schleiffer A, Eisenhaber F, Mechtler K, Haering CH, Nasmyth K (2002) Eco1 is a novel acetyltransferase that can acetylate proteins involved in cohesion. *Curr Biol* 12:323e8.
- Jian X, Boerwinkle E, Liu X (2014) In silico prediction of splice-altering single nucleotide variants in the human genome. *Nucleic Acids Res* 42(22):13534-13544.
- Kang G, Bi W, Zhang H, Pounds S, Cheng C, Shete S, Zou F, Zhao Y, Zhang JF, Yue W (2016) A Robust and Powerful Set-Valued Approach to Rare

- Variant Association Analyses of Secondary Traits in Case-Control Sequencing Studies. *Genetics* pii: genetics.116.192377.
- Kenyon BM, Browne F, D'Amato RJ (1997) Effects of thalidomide and related metabolites in a mouse corneal model of neovascularization. *Exp Eye Res* 64(6):971-8.
- Kim W, Bennett EJ, Huttlin EL, Guo A, Li J, Possemato A, Sowa ME, Rad R, Rush J, Comb MJ, Harper JW, Gygi SP (2011) Systematic and quantitative assessment of the ubiquitin modified proteome *Mol Cell* 44(2): 325–340.
- Knobloch J, Rüther U (2008) Shedding light on an old mystery: thalidomide suppresses survival pathways to induce limb defects. *Cell Cycle* 7(9):1121-7.
- Kohlhase J, Schuh R, Dowe G, Kühnlein RP, Jäckle H, Schroeder B, Schulz-Schaeffer W, Kretzschmar HA, Köhler A, Müller U, Raab-Vetter M, Burkhardt E, Engel W, Stick R (1996) Isolation, characterization, and organ-specific expression of two novel human zinc finger genes related to the *Drosophila* gene *spalt*. *Genomics* 38:291–298.
- Kohlhase J, Heinrich M, Schubert L, Liebers M, Kispert A, Laccone F, Turnpenny P, Winter RM, Reardon W (2002) Okihiro syndrome is caused by *SALL4* mutations. *Hum Mol Genet* 11:2979–87.
- Kohlhase J, Chitayat D, Kotzot D, Ceylaner S, Froster UG, Fuchs S, Montgomery T, Rösler B (2005) *SALL4* mutations in Okihiro syndrome (Duane-radial ray syndrome), acro-renal-ocular syndrome, and related disorders. *Hum Mutat* 26:176–83.
- Kohlhase J (2007) Okihiro Syndrome. *Encyclopedia of Molecular Mechanisms of Disease*. 2nd ed. Springer, Berlin, p 1517-1519.
- Koshiba-Takeuchi K, Matsumoto K, Vogel-Hopker A, Naitoh-Matsuo M, Ogura K, Takahashi N, Yasuda K, Ogura T (1999) *Tbx5* and *Tbx4* genes determine the wing/leg identity of limb buds. *Nature* 398:810-4.
- Koshiba-Takeuchi K, Arruda EP, Kathiriya IS, Mo R, Hui CC, Srivastava D, Bruneau BG (2006) Cooperative and antagonistic interactions between *Sall4* and *Tbx5* pattern the mouse limb and heart. *Nat Genet* 38(2):175-83.
- Kowalski TW, Sanseverino MT, Schuler-Faccini L, Vianna FS (2015) Thalidomide embryopathy: Follow-up of cases born between 1959 and 2010. *Birth Defects Res A Clin Mol Teratol* 103(9):794-803.
- Lee S, Won S, Kim YJ, Kim Y, T2D-Genes Consortium, Kim BJ, Park T (2016) Rare variant association test with multiple phenotypes. *Genet Epidemiol* [Epub ahead of print].
- Leem YE, Choi HK, Jung SY, Kim BJ, Lee KY, Yoon K, Qin J, Kang JS, Kim ST (2011) *Esco2* promotes neuronal differentiation by repressing Notch signaling. *Cell Signal* 23(11):1876-84.
- Lenz W (1973) Phenocopies. *J Med Genet* 10(1):34-49.
- Lenz W (1988) A short history of Thalidomide Embryopathy. *Teratology* 38(3): 203-215.
- Li QY, Newbury-Ecob RA, Terrett JA, Wilson DI, Curtis AR, Yi CH, Gebuhr T, Bullen PJ, Robson SC, Strachan T, Bonnet D, Lyonnet S, Young ID, Raeburn JA, Buckler AJ, Law DJ, Brook JD (1997) Holt-Oram syndrome is caused by mutations in *TBX5*, a member of the Brachyury (T) gene family. *Nat Genet* 15:21–29.

- MacDermot K, Winter R (1987) Radial ray defect and duane anomaly: report of a family with autosomal dominant transmission. *Am J Med Genet* 27:313–319.
- Marín-Martín FR, Soler-Rivas C, Martín-Hernández R, Rodríguez-Casado A (2014) A Comprehensive In Silico Analysis of the Functional and Structural Impact of Nonsynonymous SNPs in the ABCA1 Transporter Gene. *Cholesterol*. 2014:639751.
- Matthews SJ, McCoy C (2003) Thalidomide: A review of approved and investigational uses. *Clin Ther* 25(2):342-395.
- McBride WG (1961) Thalidomide and congenital abnormalities. *Lancet* 2:1358.
- McDermott DA, Bressan MC, He J, Lee JS, Aftimos S, Brueckner M, Gilbert F, Graham GE, Hannibal MC, Innis JW, Pierpont ME, Raas-Rothschild A, Shanske AL, Smith WE, Spencer RH, St John-Sutton MG, van Maldergem L, Waggoner DJ, Weber M, Basson CT (2005) TBX5 genetic testing validates strict clinical criteria for Holt-Oram syndrome. *Pediatr Res* 58(5):981-6.
- Miertus J, Borozdin W, Frecer V, Tonini G, Bertok S, Amoroso A, Miertus S, Kohlhase J (2006) A SALL4 zinc finger missense mutation predicted to result in increased DNA binding affinity is associated with cranial midline defects and mild features of Okihiro syndrome. *Hum Genet* 119(1-2):154-61.
- Miller MT, Strömmland K (1999) Teratogen update: Thalidomide: A review, with a focus on ocular findings and new potential uses. *Teratology* 60(5):306-321.
- Mönnich M, Kuriger Z, Print CG, Horsfield JA (2011) A Zebrafish Model of Roberts Syndrome Reveals That Esco2 Depletion Interferes with Development by Disrupting the Cell Cycle. *PLoS One* 6(5):e20051.
- Morin RD, O'Connor MD, Griffith M, Kuchenbauer F, Delaney A, Prabhu AL, Zhao Y, McDonald H, Zeng T, Hirst M, Eaves CJ, Marra MA (2008) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res* 18:610-621.
- Neff AW, King MW, Harty MW, Nguyen T, Calley J, Smith RC, Mescher AL (2005) Expression of Xenopus XISALL4 During Limb Development and Regeneration. *Dev Dyn* 233(2):356-67.
- Neff AW, King MW, Mescher AL (2011) Dedifferentiation and the Role of Sall4 in Reprogramming and Patterning During Amphibian Limb Regeneration. *Dev Dyn* 240(5):979-89.
- Newbury-Ecob R, Leanage R, Raeburn JA, Young ID (1996) The Holt-Oram syndrome: a clinical genetic study. *J Med Genet* 33:300–307.
- Newman CG (1986) The thalidomide syndrome: risks of exposure and spectrum of malformations. *Clin Perinatol* 13(3):555-573.
- Okihiro MM, Tasaki T, Nakano KK, Bennett BK (1977) Duane syndrome and congenital upper-limb anomalies: A familial occurrence. *Arch Neurol* 34:174–179.
- Oliveira MA, Bermudez JAZ, de Souza ACM (1999) Talidomida no Brasil: vigilância com responsabilidade compartilhada? *Caderno de Saúde Pública* 15(1):99-112.
- Parman T, Wiley MJ, Wells PG (1999) Free radical-mediated oxidative DNA damage in the mechanism of thalidomide. *Nat Med* 5(5): 582-585.
- Pazoki R, de Jong JS, Marsman RF, Bruinsma N, Dekker LR, Wilde AA, Bezzina CR, Tanck MW (2013) SNPs Identified as Modulators of ECG Traits in the General Population Do Not Markedly Affect ECG Traits during Acute

- Myocardial Infarction nor Ventricular Fibrillation Risk in This Condition. *PLoS One* 8(2):e57216.
- Penna GO, Pinheiro AMC, Hajjar LA (1998) Talidomida: mecanismo de ação, efeitos colaterais e uso terapêutico. *An Bras Dermatol* 73(6):501-514.
- Postma AV, van de Meerakker JB, Mathijssen IB, Barnett P, Christoffels VM, Ilgun A, Lam J, Wilde AA, Lekanne Deprez RH, Moorman AF (2008) A gain-of-function TBX5 mutation is associated with atypical Holt-Oram syndrome and paroxysmal atrial fibrillation. *Circ Res* 102:1433–1442.
- Poznanski AK, Gall JC Jr, Stern AM (1970) Skeletal manifestations of the Holt-Oram syndrome. *Radiology* 94:45–53.
- Rallis C, Bruneau BG, Del Buono J, Seidman CE, Seidman JG, Nissim S, Tabin CJ, Logan MP (2003) Tbx5 is required for forelimb bud formation and continued outgrowth. *Development* 130(12):2741-51.
- Resta N, Susca FC, Di Giacomo MC, Stella A, Bukvic N, Bagnulo R, Simone C, Guanti G (2006) A homozygous frameshift mutation in the ESCO2 gene: evidence of intertissue and interindividual variation in Nmd efficiency. *J Cell Physiol* 209(1):67-73.
- Roberts JB (1919) A child with double cleft lip and palate, protrusion of the intermaxillary portion of the upper jaw and imperfect development of the bones of the four extremities. *Ann Surg* 70:252-3
- Rodriguez-Esteban C, Tsukui T, Yonei S, Magallon J, Tamura K, Belmonte JCL (1999) The T-box genes Tbx4 and Tbx5 regulate limb outgrowth and identity. *Nature* 398:814-8.
- Russell M, Williams M, May E, Stewart C (2010) The conundrum of detecting stable angina pectoris in the community setting. *Nat Rev Cardiol* 7, 106–113.
- Sakaki-Yumoto M, Kobayashi C, Sato A, Fujimura S, Matsumoto Y, Takasato M, Kodama T, Aburatani H, Asashima M, Yoshida N, Nishinakamura R (2006) The murine homolog of SALL4, a causative gene in Okihiro syndrome, is essential for embryonic stem cell proliferation, and cooperates with Sall1 in anorectal, heart, brain and kidney development. *Development* 133:3005–3013.
- Saldanha PH (1994) A tragédia da Talidomida e o Advento da Teratologia Experimental. *Revista Brasileira de Genética* 17:449-464.
- Sampaio EP, Sarno EN, Galilly R, Cohn ZA, Kaplan G (1991) Thalidomide selectively inhibits tumor necrosis factor alpha production by stimulated human monocytes. *J Exp Med* 173(3):699-703.
- Schüle B, Oviedo A, Johnston K, Pai S, Francke U (2005) Inactivating mutations in ESCO2 cause SC phocomelia and Roberts syndrome: no phenotype-genotype correlation. *Am J Hum Genet* 77:1117-28.
- Schuler-Faccini L, Soares RCF, Sousa ACM, Maximino C, Luna E, Schwartz IVD, Waldman C, Castilla EE (2007) New cases of thalidomide embriopathy in Brazil. *Birth Defects Res A* 79:671-672.
- Shardein JL (1993) Psychotropic Drugs. *Chemically Induced Birth Defects*. 2nd ed. Marcel Dekker, New York, p 208-270.
- Sheskin J (1965) Thalidomide in the treatment of lepra reactions. *Clin Pharmacol Ther* 6:303-306.
- Smith AT, Sack GH, Taylor GJ (1979) Holt-Oram syndrome. *J Pediatr* 95:538-43.

- Smithells RW (1973) Defects and disabilities of thalidomide children. *Br Med J* 1:269-272.
- Smithells RW, Newman CGH (1992) Recognition of thalidomide defects. *J Med Genet* 29(10): 716-723.
- Sotoodehnia N, et al. (2010) Common variants in 22 loci are associated with QRS duration and cardiac ventricular conduction. *Nat Genet* 42(12):1068.
- Strömmland K, Nordin V, Miller M, Akerström B, Gillberg C (1994) Autism in thalidomide embryopathy: a population study. *Dev Med Child Neurol* 36(4):351-6.
- Sweetman D, Smith T, Farrell ER, Chantry A, Munsterberg A (2003) The conserved glutamine-rich region of chick *csal1* and *csal3* mediates protein interactions with other spalt family members. Implications for Townes-Brocks syndrome. *The Journal of biological chemistry* 278:6560-6566.
- Tang R, Prosser DO, Love DR (2016) Evaluation of Bioinformatic Programmes for the Analysis of Variants within Splice Site Consensus Regions. *Adv Bioinformatics* 2016:5614058.
- Temtamy SA, Shoukry AS, Fayad I, El-Meligy MR (1975) Limb malformations in the cloverleaf skull anomaly. *Birth Defects Orig Artic Ser* 11(2):247-51.
- Temtamy SA, McKusick VA (1978) The genetics of hand malformations. *Birth Defects Orig Artic Ser* 14(3):i-xviii, 1-619.
- Therapontos C, Erskine L, Gardner ER, Figg WD, Vargesson N (2009) Thalidomide induces limb defects by preventing angiogenic outgrowth during early. *Proc Natl Acad Sci USA* 106(21): 8573-8578.
- Tomkins D, Hunter A, Roberts M (1979) Cytogenetic findings in Roberts-SC phocomelia syndrome(s). *Am J Med Genet* 4:17-26.
- Van den Berg DJ, Francke U (1993) Roberts syndrome: a review of 100 cases and a new rating system for severity. *Am J Med Genet* 47:1104-23.
- Vargesson N (2009) Thalidomide-induced limb defects: resolving a 50-year-old puzzle. *Bioessays* 31:1327-36.
- Vega H, Waisfisz Q, Gordillo M, Sakai N, Yanagihara I, Yamada M, van Gosliga D, Kayserili H, Xu C, Ozono K, Jabs EW, Inui K, Joenje H (2005) Roberts syndrome is caused by mutations in *ESCO2*, a human homolog of yeast *ECO1* that is essential for the establishment of sister chromatid cohesion. *Nat Genet* 37:468-70.
- Vega H, Trainer AH, Gordillo M, Crosier M, Kayserili H, Skovby F, Uzielli MLG, Schnur RE, Manouvrier S, Blair E, Hurst JA, Forzano F, Meins M, Simola KOJ, Raas-Rothschild A, Hennekam RCM, Jabs EW (2010) Phenotypic variability in 49 cases of *ESCO2* mutations, including novel missense and codon deletion in the acetyltransferase domain, correlates with *ESCO2* expression and establishes the clinical criteria for Roberts syndrome. *J Med Genet* 47(1):30-7.
- Vianna FS, Schüler-Faccini L, Leite JC, de Sousa SH, da Costa LM, Dias MF, Morelo EF, Doriqui MJ, Maximino CM, Sanseverino MT (2013) Recognition of the phenotype of thalidomide embryopathy in countries endemic for leprosy: new cases and review of the main dysmorphological findings. *Clin Dysmorphol* 22:59-63.
- Wang B, Li L, Xie X, Wang J, Yan J, Mu Y, Ma X (2010) Genetic variation of *SAL-Like 4* (*SALL4*) in ventricular septal defect. *Int J Cardiol* 145(2):224-6.

Xiong J (2014) SALL4: engine of cell stemness. *Curr Gene Ther* 14(5):400-11.

CAPÍTULO VIII

ANEXOS

8. ARTIGO ANEXO

Genetic Susceptibility to Thalidomide Embryopathy in Humans: analyses in candidate genes

Artigo em preparação
Revista *Birth Defects Part A*

De forma a contribuir com um dos objetivos almejados neste trabalho, sobre o entendimento da suscetibilidade para o desenvolvimento de TE, uma avaliação de variantes genéticas em genes de desenvolvimento foi conduzida paralelamente (Anexo 1). No estudo citado, nós investigamos variantes nos genes *FGF8*, *FGF10*, *SHH*, *BMP4* – importantes no desenvolvimento de membros e diferencialmente expressos após exposição à talidomida – e *TP53*, *TP63* e *TP73* – importantes no desenvolvimento de membros ou na teratogênese – em uma amostra de 38 indivíduos afetados pela TE, comparando com um grupo controle, não afetado, de 136 indivíduos. Nós não identificamos diferenças nas frequências alélicas e genótípicas dos polimorfismos avaliados entre os dois grupos. As limitações foram as mesmas deste presente trabalho. Nós consideramos que outras variantes, em regiões de ligação de fatores de transcrição e promotores desses genes, devem ser investigadas para auxiliar no entendimento do papel desses genes na TE.

Genetic Susceptibility to Thalidomide Embryopathy in Humans: analyses in candidate genes

Julia do Amaral Gomes^{1,2*}, Thayne Woycinck Kowalski^{1,2*}, Luciana Tovo-Rodrigues¹, Lucas Rosa Fraga^{1,2}, Flavia Costa Biondi^{1,2}, Maria Teresa Vieira Sanseverino^{1,2,3}, Lavínia Schuler-Faccini^{1,2,3} and Fernanda Sales Luiz Vianna^{1,2,3,4}

Authors' Affiliations

¹Postgraduate Program in Genetics and Molecular Biology, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.

²National Institute of Population Medical Genetics (INAGEMP), Porto Alegre, Brazil.

³Brazilian Teratogen Information Service (SIAT), Medical Genetics Service at the Porto Alegre Clinics Hospital, Porto Alegre, Brazil.

⁴Genomic Medicine Laboratory of Porto Alegre Clinics Hospital, Porto Alegre, Brazil.

*The authors contributed equally to this work

Contract grant sponsor: National Institute of Population Medical Genetics (INAGEMP);
Contract grant number: (CNPq 573993/2008-4).

Corresponding author:

Fernanda Sales Luiz Vianna

Brazilian Teratogen Information Service (SIAT)

Medical Genetics Service

Hospital de Clínicas

90035-903 – Porto Alegre – RS – Brazil

Fax 55 51 33598008

E-mail: fslvianna@gmail.com

Abstract

Thalidomide is a drug used worldwide for several indications, but the molecular mechanisms of its teratogenic property are not fully understood. Studies in animal models suggest the oxidative stress, the inhibition of angiogenesis, and the binding to E3-ubiquitin ligase complex as mechanisms by which thalidomide can alter the expression of genes important to embryonic development. In this study, 15 polymorphisms in genes related to development and to pharmacogenetic response (*FGF8*, *FGF10*, *BMP4*, *SHH*, *TP53*, *TP63*, *TP73*, *CYP2C19* and *TNF- α*) were analyzed in people with thalidomide embryopathy (TE) and compared with those without malformations. The sample consisted of 38 people with TE and 136 unrelated and non-syndromic people who had their DNA genotyped by PCR real-time. No allelic or genotypic differences were observed among the groups. We hypothesized that other regions in these and related genes may play more important role in thalidomide teratogenesis, which is known to have a genetic contribution. Identifying such molecular mechanisms is essential for the development of a molecule that will be analogous to thalidomide but safe enough to avoid the emergence of new cases of embryopathy.

Key words: teratogenesis; development genes; pharmacogenetic genes; limb reduction defects; thalidomide embryopathy

1. Introduction

Thalidomide was found to cause congenital defects in humans few years following its release to the market (Lenz, 1988; Smithells and Newman, 1992). The range of damage caused by the exposure to thalidomide in humans is known as thalidomide embryopathy (TE) and the most well described defects are the limb reduction defects (LRD), which occur in 90% of the affected individuals. Other defects include eye, ear and cardiac malformations (Lenz, 1988; Smithells and Newman, 1992; Kowalski et al., 2015). It is estimated that 10,000 people were affected by the teratogenic effects of thalidomide worldwide. Although thalidomide was withdrawn from the market in early 1960s (Lenz, 1988; Smithells and Newman, 1992), the medication returned to the market in the same decade, since it was found to have immunomodulatory, anti-inflammatory and antiangiogenic properties (Sampaio et al., 1991; D'Amato et al., 1994). Those properties make the drug powerful for the treatment of several conditions such as erythema nodosum leprosum (ENL) and multiple myeloma (MM) (Matthews and McCoy, 2003).

Several hypotheses have been postulated to try to understand the molecular mechanisms that underlie TE, including antiangiogenesis, formation of reactive oxygen species (ROS), and inhibition of ubiquitination activity of an E3 ubiquitin ligase complex by direct connection of thalidomide to Cereblon (CRBN) protein, part of this complex (D'Amato et al., 1994; Hansen and Harris, 2004; Ito et al., 2010). Our group have previously published studies in TE people who were assessed to *CRBN*-thalidomide binding region and angiogenesis-related genes functional polymorphisms (Vianna et al., 2013; Kowalski et al., 2016; Vianna et al., 2016; Kowalski et al., 2017), suggesting variants that could help to explain the role of molecular mechanisms in TE in humans.

Besides these three hypotheses, the analysis of genes related to pharmacogenetic response of thalidomide and development genes seems promising and deserves attention in the understanding of thalidomide teratogenesis. The cytochrome P450 enzyme complex (CYP2C19) is responsible for 5'-hydroxylation of thalidomide metabolites in humans (Ando et al., 2002) and the tumor necrosis factor alpha (TNF- α) is a cytokine selectively inhibited by

treatment with thalidomide (Sampaio et al., 1991). Some development genes, such as the p53 family (*TP53*, *TP63* and *TP73*), are important during the embryo development and have already been associated to teratogenesis outcomes (Rinne et al., 2007; Torchinsky and Toder, 2010). *Fibroblast growth factor-8* and *-10* (*FGF8* and *FGF10*), *Sonic hedgehog* (*SHH*) and *Bone morphogenetic protein 4* (*BMP4*) genes, important to limb development, were reported to have their expression changed in embryos of animal models exposed to thalidomide (Hansen et al., 2002; Therapontos et al., 2009; Kawamura et al., 2014).

Here, we show results from the progress of our study in genetic susceptibility study of TE, which we have analysed several genetic variants in genes involved on developmental processes and metabolism and inflammatory properties of thalidomide in people having TE comparing to people without malformations.

2. Materials and Methods

2.1 Sample

Thirty-eight individuals with TE were recruited from the Brazilian Association of People Affected by Thalidomide Syndrome (ABPST), all subjects signed the informed consent and all of whom were previously examined by experts in TE and differential diagnosis. A clinical evaluation of congenital anomalies was performed in 28 of these individuals in order to assess details the compatible phenotype of TE (Kowalski et al., 2015). Telephone calls were used to evaluate the other individuals, given the distance of their residences from ABPST and difficulty to follow up in person. Saliva samples were collected and stored in DNA Oragene® Kits (Genotek). DNA extraction was performed according to the manufacturer's instructions. A sample of 136 individuals without congenital anomalies from the Brazilian population was used as a control group. These subjects were selected according to similar time and place of birth of the individuals with TE. This study was approved by the Ethics Committee in Research of Porto Alegre Clinics Hospital (number 10-0422).

2.2 Genomic Analysis

Single nucleotide polymorphisms (SNPs) in candidate genes were analysed. The genes analyzed were *FGF8*, *FGF10*, *BMP4*, *SHH*, *TP53*, *TP63*, *TP73*, *TNF- α* and *CYP2C19*. The SNPs of these genes were selected based on the following criteria: frequency in the population; reported association with any medical condition; and, in negative cases, the greatest functionality value based on the Functional Single Nucleotide Polymorphism (F-SNP) database. Supplementary table summarizes all the polymorphisms evaluated in the present study. In order to genotype them, DNA was extracted from saliva samples using the OrageneDNA extraction kit (DNA Genotek®), in accordance with the manufacturer's instructions.

The genotypic determination of the polymorphisms was performed by allelic discrimination using specific probes that had been previously designed and validated (TaqMan® Genotyping Assay, Applied Biosystems, USA). All the assays were genotyped in accordance with the protocol recommended by the manufacturer.

2.3 Statistical Analysis

For all the polymorphisms, the Hardy-Weinberg equilibrium was tested. The differences for the allelic and genotypic frequencies between the groups were compared using Fisher's exact test and a two-tailed p-value <0.05 was considered significant. The Bonferroni correction was applied in all the sample groups. The tests were performed using the SPSS® version 18 (IBM, USA).

The linkage disequilibrium for SNPs contained in the same gene was calculated using the Haploview version 4.2. The haplotypes were inferred using the Bayesian algorithm implemented in the Phase 2.1.1.

3. Results

A sample of 38 individuals with TE and 136 members of the general population, unaffected by TE, were included in our study. The characterization of individuals with TE is described in Table 1.

The allelic and genotypic frequencies of the polymorphisms are shown in Table 2. The distribution of all polymorphisms in both sample groups was in accordance with the Hardy-Weinberg equilibrium. The *loci* rs28936675 (*SHH* gene) and rs4986893 (*CYP2C19*3*) were monomorphic in our sample and then were excluded from further analyses. We did not find significant differences in the frequencies of all analyzed polymorphisms comparing TE subjects and controls. Further dominant and recessive models analyses were conducted to verify the contribution of alleles for TE and no statistical differences was found between the groups.

The linkage disequilibrium (LD) analyses for the SNPs contained in the *TNF- α* and *CYP2C19* genes showed that most of the *loci* are in strong linkage disequilibrium despite the analysis not having the ability to achieve statistical significance for some values (data not shown). Finally, there was no difference in the distribution of the haplotypes in these genes between the assessed groups.

4. Discussion

Studies attempting to understand the molecular aspects of TE have advanced greatly over the past 20 years, mainly in the following three areas: oxidative stress, antiangiogenic mechanism, and binding of thalidomide to the Cereblon protein, however, more research is needed to the full understanding of the molecular mechanisms responsible for the teratogenic propriety of thalidomide.

The susceptibility of an individual a teratogen exposure is determined by pre-existing genetic factors, resulting from the interaction between genetic factors of the mother and of the embryo. Some polymorphisms in humans can lead to higher or lower susceptibility to teratogen-induced damage (Cassina et al., 2012). Our group have previously conducted exploratory studies in CRBN and

angiogenesis-related genes, and demonstrated that variants in these targets can help to understand TE scenario in humans (Vianna et al., 2013; Kowalski et al., 2016; Vianna et al., 2016; Kowalski et al., 2017).

In the present study, we evaluated whether a genetic predisposition to thalidomide embryopathy might be detected by analysing polymorphisms in genes of pathways that are not binding targets of thalidomide (i.e. development, metabolism and inflammatory pathways), but affected by the drug. This non-excluding hypothesis is based on assays with model organisms of thalidomide teratogenicity, which showed that genes important for limb development are affected by thalidomide, having their expression altered after exposure to thalidomide (Hansen et al., 2002; Therapontos et al., 2009; Kawamura et al., 2014). To address this, we analysed genetic variants in some genes that possibly could be involved in genetic susceptibility to TE, both development genes and genes involved on metabolism and inflammatory properties of thalidomide. Some of these variants evaluated were previously reported as associated to impairment of their proteins functions or increased risk to developmental defects, cancers or adult onset conditions (Roessler et al., 1996; Dumont et al., 2003; Cui et al., 2012; Uppugunduri et al., 2012; Zhang et al., 2014; Hu et al., 2015).

It was not possible to establish an association between the 15 polymorphisms evaluated and TE susceptibility. The lack of association between the genotypes studied and the appearance of malformations can be attributed to limitations such as the small sample size and power, lack of information about the day of exposure and dose, and the selected polymorphisms not playing a major role in this outcome. Moreover, we do not know exactly what degree of genetic contribution and what kind of changes could allow a higher susceptibility to this phenotype.

Thalidomide embryopathy is a condition of complex origin, in which genes and characteristics of drug exposure as well as other factors contribute to the occurrence of birth defects and phenotype. The genes analyzed in this work are extremely important for the metabolism of the drug in the body and limb development in the embryonic period. However, the evaluated polymorphisms do

not appear to be associated with susceptibility to TE in the analyzed sample. We hypothesized that other genetic alterations, such as changes in regulatory binding sites or in the promoter region of these and other genes, should be investigated to help to clarify their genetic role in TE.

Acknowledgements

The authors acknowledge INAGEMP – National Institute of Population Medical Genetics (grant CNPq 573993/2008-4), FIPE/HCPA (10-0244) and Brazilian Association of People Affected by Thalidomide Syndrome - ABPST for the support provided for this project.

The authors declare no conflicts of interest.

References

- Ando Y, Price DK, Dahut WL, Cox MC, Reed E, Figg WD. 2002. Pharmacogenetic associations of CYP2C19 genotype with in vivo metabolisms and pharmacological effects of thalidomide. *Cancer Biol Ther* 1: 669-673.
- Cassina M, Salviati L, Di Gianantonio E, Clementi M. 2012. Genetic susceptibility to teratogens: state of the art. *Reprod Toxicol*. 34(2):186-191.
- Cui G, Wang H, Li R, Zhang L, Li Z, Wang Y, Hui R, Ding H, Wang DW. 2012. Polymorphism of tumor necrosis factor alpha (TNF-alpha) gene promoter, circulating TNF-alpha level, and cardiovascular risk factor for ischemic stroke. *J Neuroinflammation*. 9:235.
- D'Amato RJ, Loughnan MS, Flynn E, Folkman J. 1994. Thalidomide is an inhibitor of angiogenesis. *Proc Natl Acad Sci U S A* 91: 4082-4085.
- Dumont P, Leu JI, Della Pietra AC III, George DL, Murphy M. 2003. The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet*. 33(3):357-365.
- Hansen JM, Gong SG, Philbert M, Harris C. 2002. Misregulation of gene expression in the redox-sensitive NF-kappaB-dependent limb. *Dev Dyn* 225: 186-194.
- Hansen JM, Harris C. 2004. A novel hypothesis for thalidomide induced limb teratogenesis: redox misregulation of the NFkappaB pathway. *Antioxid Redox Signal* 6:1-14.
- Hu Y, Qin C, Deng M, Niu Y, Long X. 2015. Association between BMP4 rs17563 Polymorphism and NSCL/P Risk: A Meta-Analysis. *Dis Markers*. 2015:763090.
- Ito T, Ando H, Suzuki T, Ogura T, Hotta K, Imamura Y, Yamaguchi Y, Handa H. 2010. Identification of a primary target of thalidomide teratogenicity. *Science* 327: 1345-1350.
- Kawamura Y, Yamashita T, Yamauchi T, Matsumoto K, Sato K. 2014. Effects of thalidomide on Fgf8, Bmp4 and Hoxa11 expression in the limb bud in Kbl:JW rabbit embryos. *Congenit Anom (Kyoto)* 54(1):54-62.
- Kowalski TW, Sanseverino MT, Schuler-Faccini L, Vianna FS. 2015. Thalidomide embryopathy: Follow-up of cases born between 1959 and 2010. *Birth Defects Res A Clin Mol Teratol* 103: 794-803.
- Kowalski TW, Fraga LR, Tovo-Rodrigues L, Sanseverino MT, Hutz MH, Schuler-Faccini L, Vianna FS. 2016. New Findings in eNOS gene and Thalidomide Embryopathy Suggest pre-transcriptional effect variants as susceptibility factors. *Sci Rep*. 6:23404.
- Kowalski TW, Fraga LR, Tovo-Rodrigues L, Sanseverino MT, Hutz MH, Schuler-Faccini L, Vianna FS. 2017. Angiogenesis-related Genes and Thalidomide

Teratogenesis in Humans: An Approach on Genetic Variation and Review of Past In Vitro Studies. *Reprod Toxicol.* . pii: S0890-6238(17)30054-0.

Lenz W. 1988. A short history of thalidomide embryopathy. *Teratology* 38: 203-215.

Matthews SJ, McCoy C. 2003. Thalidomide: a review of approved and investigational uses. *Clin Ther* 25: 342-395.

Rinne T, Brunner HG, van Bokhoven H. 2007. p63-Associated Disorders. *Cell Cycle* 6(3):262-268.

Roessler E, Belloni E, Gaudenz K, Jay P, Berta P, Scherer SW, Tsui LC, Muenke M. 1996. Mutations in the human Sonic Hedgehog gene cause holoprosencephaly. *Nat Genet.* 14(3):357-360.

Sampaio EP, Sarno EN, Galilly R, Cohn ZA, Kaplan G. 1991. Thalidomide selectively inhibits tumor necrosis factor alpha production by stimulated human monocytes. *J Exp Med* 173: 699-703.

Smithells RW, Newman CG. 1992. Recognition of thalidomide defects. *J Med Genet* 29: 716-723.

Therapontos C, Erskine L, Gardner ER, Figg WD, Vargesson N. 2009. Thalidomide induces limb defects by preventing angiogenic outgrowth during early. *Proc Natl Acad Sci U S A* 106: 8573-8578.

Torchinsky A, Toder V. 2010. Mechanisms of the embryo's response to embryopathic stressors: a focus on p53. *J Reprod Immunol* 85: 76-80.

Uppugunduri CR, Daali Y, Desmeules J, Dayer P, Krajinovic M, Ansari M. 2012. Transcriptional regulation of CYP2C19 and its role in altered enzyme activity. *Curr Drug Metab.* 13(8):1196-1204.

Vianna FS, Fraga LR, Tovo-Rodrigues L, Tagliani-Ribeiro A, Biondi F, Maximino CM, Sanseverino MT, Hutz MH, Schuler-Faccini L. 2013. Polymorphisms in the endothelial nitric oxide synthase gene in thalidomide embryopathy. *Nitric Oxide* 35:89-92.

Vianna FS, Kowalski TW, Tovo-Rodrigues L, Tagliani-Ribeiro A, Godoy BA, Fraga LR, Sanseverino MT, Hutz MH, Schuler-Faccini L. 2016. Genomic and in silico analyses of CRBN gene and thalidomide embryopathy in humans. *Reprod Toxicol* 66:99-106.

Zhang N, Huo Q, Wang X, Chen X, Long L, Guan X, Jiang L, Ma T, Hu W, Yang Q. 2014. A genetic variant in p63 (rs17506395) is associated with breast cancer susceptibility and prognosis. *Gene.* 535(2):170-176.

Table 1. Social and clinical characterization of individuals with TE.

Characteristic	Number	%
Sex		
<i>Male</i>	16/28	57
Time of Birth		
<i>1959-1964</i>	21/28	75
<i>1965-2010</i>	7/28	25
Congenital Anomalies		
<i>Upper limbs</i>	27/28	96
<i>Lower limbs</i>	13/28	46
<i>Ears</i>	4/28	14
<i>Eyes</i>	7/28	25
<i>Oral cavity</i>	2/28	7
<i>Neurologic</i>	6/28	21
<i>Heart</i>	4/28	14
<i>Genital tract</i>	2/28	7
<i>Urinary tract</i>	3/28	11
<i>Skeletal</i>	3/28	11

Table 2. Allelic and genotypic frequencies of the studied variants in affected by thalidomide embryopathy and unaffected groups.

Gene	Genotype/ Allele	Affected n	Affected (%)	Unaffected N	Unaffected (%)	p^a
SHH (rs28936675)	CC	38	(100.0)	135	(100.0)	-
	CT	0	(0.0)	0	(0.0)	
	TT	0	(0.0)	0	(0.0)	
	C	76	(100.0)	170	(100.0)	
	T	0	(0.0)	0	(0.0)	
FGF8 (rs1348870)	AA	11	(28.9)	41	(30.1)	0.940
	AG	20	(52.6)	67	(49.3)	
	GG	7	(18.4)	28	(20.6)	
	A	42	(55.3)	149	(54.8)	
	G	34	(44.7)	123	(45.2)	
FGF10 (rs900379)	CC	11	(28.9)	46	(34.1)	0.366
	CT	19	(50.0)	50	(37.0)	
	TT	8	(21.1)	39	(28.9)	
	C	41	(53.9)	142	(52.6)	
	T	35	(46.1)	128	(47.4)	
BMP4 (rs17563)	AA	12	(31.6)	49	(36.3)	0.882
	AG	19	(50.0)	62	(45.9)	
	GG	7	(18.4)	24	(17.8)	
	A	43	(56.6)	160	(59.3)	
	G	33	(43.4)	110	(40.7)	
TP53 (rs1042522)	GG	16	(42.1)	62	(45.9)	0.477
	GC	20	(52.6)	58	(43.0)	
	CC	2	(5.3)	15	(11.1)	
	G	52	(68.4)	182	(67.4)	
	C	24	(31.6)	88	(32.6)	
TP63 (rs17506395)	TT	16	(42.1)	81	(60.0)	0.131
	TG	18	(47.4)	42	(31.1)	
	GG	4	-12	12	(8.9)	
	T	50	-204	204	(75.6)	
	G	26	-66	66	(24.4)	
TP73 (rs2273953)	GG	22	(57.9)	86	(63.2)	0.153
	GA	12	(31.6)	46	(33.8)	
	AA	4	(10.5)	4	(3.0)	
	G	56	(73.7)	218	(80.1)	
	A	20	(26.3)	54	(19.9)	

Table 2. Allelic and genotypic frequencies of the studied variants in affected by thalidomide embryopathy and unaffected groups (Continued).

Gene	Genotype/ Allele	Affected		Unaffected		<i>p</i> ^a
		n	(%)	n	(%)	
<i>TNF-α (rs361525)</i>	GG	36	(94.7)	121	(89.0)	0.370
	GA	2	(5.3)	15	(11.0)	
	AA	0	(0.0)	0	(0.0)	
	G	74	(97.4)	257	(94.5)	0.383
	A	2	(2.6)	15	(5.5)	
<i>TNF- α (rs1800629)</i>	GG	27	(71.1)	103	(75.7)	0.705
	GA	11	(28.9)	31	(22.8)	
	AA	0	(0.0)	2	(1.5)	
	G	65	(85.5)	237	(87.1)	0.704
	A	11	(14.5)	35	(12.9)	
<i>TNF- α (rs1799724)</i>	CC	33	(86.8)	107	(79.3)	0.686
	CT	5	(13.2)	26	(19.3)	
	TT	0	(0.0)	2	(1.5)	
	C	71	(93.4)	240	(88.9)	0.289
	T	5	(6.6)	30	(11.1)	
<i>TNF- α (rs1800630)</i>	AA	2	(5.3)	6	(4.5)	0.476
	AC	5	(13.2)	30	(22.4)	
	CC	31	(81.6)	98	(73.1)	
	A	9	(11.8)	42	(15.7)	0.469
	C	67	(88.2)	226	(84.3)	
<i>TNF- α (rs1799964)</i>	TT	26	(68.4)	87	(64.0)	0.632
	TC	10	(26.3)	44	(32.4)	
	CC	2	(5.3)	5	(3.6)	
	T	62	(81.6)	218	(80.1)	0.871
	C	14	(18.4)	54	(19.9)	
<i>CYP2C19*2 (rs4244285)</i>	GG	32	(84.2)	107	(79.3)	0.729
	GA	6	(15.8)	27	(20.0)	
	AA	0	(0.0)	1	(0.7)	
	G	70	(92.1)	241	(89.3)	0.666
	A	6	(7.9)	29	(10.7)	
<i>CYP2C19*3 (rs4244285)</i>	GG	38	(100.0)	136	(100.0)	-
	GA	0	(0.0)	0	(0.0)	
	AA	0	(0.0)	0	(0.0)	
	G	76	(100.0)	172	(100.0)	-
	A	0	(0.0)	0	(0.0)	

Table 2. Allelic and genotypic frequencies of the studied variants in affected by thalidomide embryopathy and unaffected groups (Continued).

Gene	Genotype/ Allele	Affected		Unaffected		<i>p</i> ^a
		n	(%)	n	(%)	
<i>CYP2C19*4 (rs28399504)</i>	AA	38	(100.0)	135	(99.3)	1.000
	AG	0	(0.0)	1	(0.7)	
	GG	0	(0.0)	0	(0.0)	
	A	76	(100.0)	271	(99.6)	1.000
	G	0	(0.0)	1	(0.5)	

^aChi-Square Test

Supplementary Table. Gene loci and position of polymorphisms evaluated^a.

Gene	Role^b	Locus	Polymorphism	rs Number	Allelic Frequency	Region	Type
<i>FGF8</i>	Limb Development – Altered expression in embryos of animal models exposed to thalidomide	10q24.32	c.-160-2012C>T	rs1348870	C=0.47	Intron	SNP
<i>FGF10</i>	Limb Development – Altered expression in embryos of animal models exposed to thalidomide	5p12	c.325+18804G>A	rs900379	A=0.47	Intron	SNP
<i>BMP4</i>	Limb Development – Altered expression in embryos of animal models exposed to thalidomide	14q22.2	c.455T>C	rs17563	T=0.33	Exon 4	Missense
<i>SHH</i>	Limb Development – Altered expression in embryos of animal models exposed to thalidomide	7q36.3	c.91G>A	rs28936675	NA	Exon 1	Missense
<i>TP53</i>	Monitors the early development process; Anti- teratogenic function	17p13.1	c.215G>C	rs1042522	G=0.46	Exon 4	Missense
<i>TP63</i>	Limb Development	3q28	c.325-4742T>G	rs17506395	G=0.11	Intron	SNP
<i>TP73</i>	Importance on development	1p36.32	c.-30G>A	rs2273953	A=0.18	5'UTR	SNP

Supplementary Table. Gene loci and position of polymorphisms evaluated^a (Continued).

Gene	Role^b	Locus	Polymorphism	rs Number	Allelic Frequency	Region	Type
TNF-α	Proinflammatory cytokine – Cytokine selectively inhibited by thalidomide	6p21.33	c.-418G>A	rs361525	A=0.06	Promoter region	SNP
		6p21.33	c.-488G>A	rs1800629	A=0.09	Promoter region	SNP
		6p21.33	c.-1037C>T	rs1799724	T=0.10	Promoter region	SNP
		6p21.33	c.-1043A>C	rs1800630	A=0.15	Promoter region	SNP
		6p21.33	c.-1211T>C	rs1799964	C=0.22	Promoter region	SNP
CYP2C19	Drugs metabolization – Biotransformation of thalidomide to 5'-OH- thalidomide						
CYP2C19*2	Poor metabolizer	10q23.33	c.681G>A	rs4244285	A=0.22	Exon 5	Synonymous
CYP2C19*3	Poor metabolizer	10q23.33	c.636G>A	rs4986893	A=0.01	Exon 4	Stop gained
CYP2C19*4	Poor metabolizer	10q23.33	c.1A>G	rs28399504	G=<0.01	Exon 1	Start lost

^aData obtained from Ensembl database^bData obtained from OMIM database.

NA: not available.