

MARILENE PAVAN

VARIANTES NO GENE RALDH2 E DOENÇAS CARDÍACAS
CONGÊNITAS

Dissertação apresentada ao
Instituto de Ciências Biomédicas da
Universidade de São Paulo, para obtenção do título de
Mestre em Ciências.

São Paulo
2007

MARILENE PAVAN

VARIANTES NO GENE RALDH2 E DOENÇAS CARDÍACAS
CONGÊNITAS

Dissertação apresentada ao
Instituto de Ciências Biomédicas da
Universidade de São Paulo, para obtenção do título de
Mestre em Ciências.

Área de Concentração:
Biologia Celular e Tecidual

Orientador:
Dr. José Xavier Neto

São Paulo
2007

Este trabalho foi realizado no Laboratório de Genética e Cardiologia Molecular do Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, pelo Programa de Pós-Graduação em Biologia Celular e Tecidual do Departamento de Biologia Celular e do Desenvolvimento do Instituto de Ciências Biomédicas da Universidade de São Paulo, com auxílio financeiro da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), bolsa de Mestrado, processo número 04/11606-8.

*Ao meu pai,
Aureo Antonio Rodrigues.
Graças à você eu sempre soube que seria bióloga.*

AGRADECIMENTOS

Ao meu orientador Dr. José Xavier Neto. Com certeza termino este mestrado uma profissional incomparavelmente melhor do que entrei. Obrigado por ter exigido de mim mais do que eu mesma pensei que pudesse realizar. Obrigada também por ter confiado a mim este projeto do qual hoje me orgulho tanto.

Ao Dr. José Eduardo Krieger, diretor do laboratório de Genética e Cardiologia Molecular do Instituto do Coração – InCor.

Ao Alexandrinho ao qual, por tudo que me ensinou, não posso simplesmente chamar de co-orientador por uma questão de justiça. Obrigada por compartilhar seus conhecimentos, pelo apoio e incentivo constantes.

Ao departamento de Biologia Celular e do Desenvolvimento do Instituto de Ciências Biomédicas da Universidade de São Paulo.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), cujo apoio financeiro possibilitou a realização deste trabalho.

Aos funcionários do laboratório: Maúde, Renata, Glória, Débora Bergara, Marcilene, Isabel Cristina, Silvana, Arruda e Janilton. Obrigada pela ajuda!

Ao meu namorado Vinícius, com certeza meu maior admirador e incentivador, pelo apoio, carinho, e principalmente, pela grande paciência.

À minha mãe Maria do Carmo e minhas irmãs Maisa e Márcia pelo amor incondicional.

Aos meus sobrinhos Gui e Léo, que fazem a vida de qualquer biólogo mais interessante com as perguntas mais difíceis de serem respondidas.

Ao *dream team*: Ana, Hozana, Marcos, Michelle, Priscila, Roberta e Rodrigo e também ao Angelo, Débora e Rafael. Tenham certeza de que vocês não serão lembrados apenas como companheiros de laboratório, mas amigos que levarei para toda a vida. Obrigada por tudo que me ensinaram profissional e pessoalmente e obrigada também pelos ótimos momentos que passamos (e que ainda vamos passar) juntos.

Agradecimento especial ao Allysson, que além de um ótimo profissional com quem aprendi muito, pessoalmente é como um irmão que eu não tive.

E, principalmente, agradeço às crianças portadoras de cardiopatias congênitas e seus pais. Sem sua contribuição este trabalho seria impossível de ser realizado.

The heart of creatures is the foundation of life, the Prince of all, the sun of their microcosm, from where all vigor and strength does flow.

- William Harvey, *De Motus Cordis*, 1628.

RESUMO

PAVAN, M. Variantes no gene RALDH2 e doenças cardíacas congênitas (Dissertação). São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo, 2007.

Estudos de privação de vitamina A indicam que o ácido retinóico (AR) é crítico para a morfogênese cardíaca. O recente desenvolvimento de ferramentas para investigar a via de síntese do AR mostraram novos papéis para o AR durante a formação do coração tais como na morfogênese sinoatrial, na assimetria direita – esquerda, na morfogênese da circulação coronariana, no crescimento da parede ventricular e dos septos e no desenvolvimento dos coxins. O objetivo do presente estudo é de investigar o papel da variação genética do gene da enzima sintetizadora do ácido retinóico – RALDH2 (ALDH1A2) – e as doenças cardíacas congênitas (DCCs). Seis diferentes SNPs foram estudados em 101 trios (pacientes e seus pais) em um estudo de TDT. Testes de associação foram desenvolvidos tanto para a informação genotípica de marcadores únicos como também para informação de estrutura de haplótipos. Ainda, foram conduzidos estudos funcionais para definir o papel potencial do único SNP localizado em um exon (A151G). Nenhum dos marcadores testados mostrou associação significativa com doenças cardíacas congênitas. Nenhum haplótipo estudado foi associado com um aumento de risco para o diagnóstico de doenças cardíacas congênitas. Análises posteriores do polimorfismo A151G indicaram que a variante produz mudanças substanciais na estrutura secundária do RNAm sugerindo um possível papel desta variante na estabilidade do RNAm. Nossas análises também mostraram que esta variante está localizada em um *exonic splicing enhancer* (ESE). Estudos funcionais de *splicing* não mostraram impacto significativo desta variante genética sobre a alteração do *splicing*. O mesmo estudo foi aplicado à outra mutação (G151T), encontrada no exon 4 durante o sequenciamento do gene RALDH2 em pacientes portadores de tetralogia de Fallot (TOF), e mostrou aumento no sinal de *splicing*. Sequenciando o gene RALDH2 nós encontramos ainda mais quatro mutações: rs34645259 (5'UTR), T157G (exon 4),

rs4646626 (exon 9) e rs35251510 (exon 11). Em resumo, não foi encontrada associação entre DCCs e variações genéticas no gene RALDH2 em humanos. Mutações com potencial de impacto na função do gene deverão ser estudadas de forma a definir seu real papel na perturbação da via do AR em humanos.

Palavras-Chave: Ácido Retinóico, Desenvolvimento Cardíaco, Estudos de Associação, Doenças Cardíacas Congênitas, TDT.

ABSTRACT

PAVAN, M. *RALDH2* genetic variants and congenital heart disease (Dissertation). São Paulo: Instituto de Ciências Biomédicas da Universidade de São Paulo, 2007.

Studies on Vitamin A deprivation indicate that retinoic acid (RA) is critical for cardiac morphogenesis. Recent development of tools to probe RA synthetic pathways disclosed previously unsuspected roles for RA in sinoatrial morphogenesis, in left right asymmetry, in coronary circulation morphogenesis, in ventricular wall and septum growth and in cushion development. The aim of the present study was to investigate the role of genetic variation in the *RALDH2* (ALDH1A2) locus and congenital heart disease. Six different SNPs were studied in 101 patient-parents trios in a TDT study design. Association tests were carried out both with single marker genotype information and with the haplotype structure information. In addition, functional studies to define the potential role of the only SNP located within an exonic region were conducted. None of the assayed markers displayed any significant association with CHD. No single haplotype was associated with an increased risk for the diagnosis of congenital heart disease. Further analysis of the A151G polymorphism indicated that the variant produced substantial changes in mRNA secondary structure suggesting a possible role of this variant in mRNA stability. Our analysis also showed that this variant is localized in a putative exonic splicing enhancer (ESE). Functional splicing studies failed to reveal a significant impact of this genetic variant and altered gene splicing. This methodology was applied to another mutation (G151T) found in exon 4 during the sequencing of *RALDH2* gene and an increase in splicing signal was observed. We found four mutations more: rs34645259, T157G (exon 4), rs4646626 and rs35251510. In summary, no

association between CHD and genetic variation at the *RALDH2* locus in humans was found. Potential functional genetic variants should be further studied in order to define their real role in RA pathway disturbances in humans.

Keywords: Retinoic Acid; Heart Development; Association Studies; Congenital Heart Disease; TDT.

LISTA DE ILUSTRAÇÕES

Figura 1. A via de síntese e degradação do ácido retinóico. O retinol é oxidado à retinaldeído pelas ADHs (álcool desidrogenases) em uma etapa reversível e oxidado a ácido retinóico em uma etapa irreversível pelas ALDHs (aldeído desidrogenases), sendo a ALDH1a2 (RALDH2) a enzima mais importante neste processo. A etapa de degradação é mediada pelas enzimas da família P450 de citocromo – oxidases, as CYPs.....**35**

Quadro 1. Marcadores e primers utilizados no estudo de TDT. Os marcadores foram escolhidos com base em sua frequência e localização.....**37**

Quadro 2. Primers utilizados para o sequenciamento do gene RALDH2.....**38**

Quadro 3. Temperatura dos primers utilizados no estudo de TDT e no sequenciamento do gene RALDH2.....**39**

Figura 2. Estrutura do gene RALDH2. O gene RALDH2 possui 13 exons e 113.000pb. A parte codificante é composta de 1557pb.....**42**

Figura 3. Mutação de troca T por C, encontrada na posição 157, no exon 4, durante o sequenciamento do gene RALDH2 (T157C). Essa mutação troca o aminoácido isoleucina pelo aminoácido treonina.....**43**

Figura 4. Mutação de troca G por T, encontrada na posição 151, no exon 4, durante o sequenciamento do gene RALDH2 (G151T). Essa mutação troca o aminoácido alanina pelo aminoácido serina.....**43**

Figura 5. Estrutura protéica da RALDH2 (vermelho). A e B: estrutura normal da proteína; aminoácido alanina (em destaque). C e D: proteína mutada; aminoácido serina (em destaque). A seta vermelha indica o segundo sítio de tetramerização da proteína (exon 13).....**45**

Figura 6. Estrutura protéica da RALDH2 (vermelho). A e B: estrutura normal da proteína; aminoácido isoleucina (em destaque). C e D: proteína mutada; aminoácido treonina (em destaque).....**46**

Figura 7. Heredogramas mostrando o padrão de herdabilidade das mutações G151T e T157C. À esquerda: o heredograma mostra que a mãe do paciente portador da mutação T157C transmitiu-a para o filho. À direita: o pai do paciente da mutação G151T também é portador da mutação e transmitiu-a para o filho.....**48**

Figura 8. Mutação de troca G por A (rs34645259) localizada na região 5´UTR do gene RALDH2. Essa mutação está localizada 6pb antes do sítio ATG e foi encontrada em aproximadamente 7% dos pacientes analisados.....**49**

Figura 9. Mutação de troca G por A (rs4646626), não sinônima, encontrada no exon 9 do gene RALDH2 em aproximadamente 67% dos pacientes analisados. Essa mutação troca um aminoácido valina por um aminoácido isoleucina e está localizada no domínio catalítico da enzima.....**49**

Figura 10. Mutação de troca C por A, sinônima (aminoácido isoleucina), encontrada no exon 11 do gene RALDH2. Essa mutação foi encontrada em aproximadamente 3.5% dos pacientes analisados.....**50**

Figura 11. Análise de *exonic splicing enhancers* através do programa ESEfinder para a seqüência referente ao exon 4 do gene RALDH2. A seta indica que o polimorfismo A151G está inserido em um sítio de ESE em um local de ligação para proteínas SF2 / ASF, pertencentes à família de proteínas do tipo SR.....**51**

Figura 12. Análise da estrutura secundária do RNA do gene RALDH2 sem o polimorfismo A151G (esquerda) e com o polimorfismo A151G (direita). Na presença do polimorfismo A151G observa-se grande variação na estrutura do RNAm, refletida na grande variação da energia livre da estrutura.....**52**

Quadro 4. Simulações de transições no primeiro e no segundo nucleotídeo do cistron GCA na posição 151 e da alteração A por G em todos os outros cistrons GCA, que codificam o aminoácido alanina.....**53**

Figura 13. Estrutura do RNAm do gene RALDH2 normal (à esquerda) e em presença da mutação G151T (à direita). Observa-se, quando a mutação G151T está presente, mudança na estrutura do RNAm e na energia livre da estrutura.....**54**

Figura 14. Estudo de *splicing* empregando o vetor pSPL3. Este vetor contém dois exons separados por um longo intron onde está inserido o sítio múltiplo de clonagem. O inserto experimental (o exon flanqueado por curtos fragmentos de suas seqüências intrônicas) é subclonado neste vetor e transfectado em células COS-7. A maquinaria de *splicing* da célula pode: (I) remover os introns e reter os exons, inclusive o experimental, ou (II) se os sinais de *splicing* forem muito fracos, remover inclusive o exon experimental.....**56**

Figura 15. Resultado dos diferentes ciclos de amplificação por PCR do cDNA sintetizado a partir do RNA extraído de células COS – 7. Nota-se uma maior exclusão do exon 4 polimórfico (da ordem de aproximadamente 9%) em relação ao exon 4 selvagem.....**57**

Quadro 5. Ensaio de *splicing* para o polimorfismo A151G: medidas de densidade das bandas através do programa Image Pro Plus.....**57**

Figura 16. Ensaio de *splicing* para a mutação G151T. Observa-se, quando a construção contém o exon 4 selvagem do gene RALDH2, uma pequena taxa de exclusão já observada anteriormente, no experimento com o polimorfismo A151G. Esta exclusão, porém não é observada quando utiliza-se a construção contendo a mutação G151T.....**59**

Figura 17. Simulação do erro no quadro de leitura da proteína RALDH2 com a possível exclusão do exon 4. Se houver a exclusão do exon 4, muda-se o quadro de leitura a partir do início do exon 5 e, como consequência, gera-se um códon de parada no exon 6.....**60**

Figura 18. Desequilíbrio de ligação entre os marcadores: 1. rs2704188; 2. rs1441815; 3. rs3784259; 4. rs1530293; 5. rs1899430.....**61**

LISTA DE TABELAS

Tabela 1 – Razão (em porcentagem) das taxas de inclusão e exclusão dos exons experimentais.....**58**

Tabela 2 - Resumo dos resultados das frequências de transmissão dos alelos selecionados**63**

Tabela 3 - Resultado do estudo de haplótipos dos cinco blocos mais frequentes no gene RALDH2.....**63**

SUMÁRIO

1. INTRODUÇÃO	20
1.1 Desenvolvimento Cardíaco	20
1.2 Doenças Cardíacas Congênitas	21
1.3 O Paradigma da Haploinsuficiência	22
1.4 Doenças Complexas	25
1.5 Ácido Retinóico e Desenvolvimento	27
1.6 A Via de Síntese do Ácido Retinóico	30
1.7 Variantes no Gene RALDH2 e Doenças Cardíacas Congênitas	35
2. MATERIAIS E MÉTODOS.....	36
2.1 Desenho de Estudo	36
2.2 Estudo de Pacientes.....	36
2.3 Ferramentas de Bioinformática e Amplificação por PCR	37
2.4 PCR e Sequenciamento de DNA	37
2.5 Clonagem, Cultura de Células e Extração de RNA	39
2.6 Análises Estatísticas	40
3. RESULTADOS.....	41
3.1 Plano Inicial do Trabalho	41
3.2 Sequenciamento do Gene RALDH2.....	41
3.2.1 Estudo de Modelagem Molecular para as Mutações G151T e T157C	44
3.2.2 Sequenciamento dos Fragmentos Referentes ao Exon 4 do Gene RALDH2 dos Pais dos Pacientes Afetados pelas Mutações G151T e T157C	47
3.2.3 Outras Mutações Encontradas no Sequenciamento do Gene RALDH2.....	49
3.2.4 Análise do Dobramento do RNAm do Gene RALDH2	50
3.2.5 Estudos de <i>Splicing</i> para o Polimorfismo A151G.....	54
3.2.6 Estudos de <i>Splicing</i> para a Mutação G151T	58
3.2.7 Consequências para a Proteína RALDH2 com a Exclusão do Exon 4.....	60
3.3 Estudo de Associação.....	61
3.3.1 Análise da Estrutura de Desequilíbrio de Ligação	61
3.3.2 Teste de Desequilíbrio de Transmissão (TDT).....	62
3.3.3 Análise da Estrutura Haplotípica	63
3.3.4 Estudo de Associação Caso – Controle para a Variante A151G.....	63
4. DISCUSSÃO	64
4.1 Mutações Não-Conservativas Encontradas no Sequenciamento do Gene RALDH2	65
4.1.1 Estudo de Modelagem Molecular para as Mutações G151T e T157C	66
4.1.2 Sequenciamento dos Fragmentos Referentes ao Exon 4 do Gene RALDH2 dos Pais dos Pacientes Afetados pelas Mutações G151T e T157C	67
4.2 Variante Conservativa Encontrada no Sequenciamento do Gene RALDH2	67
4.3 Variantes Sinônimas Encontradas no Sequenciamento do Gene RALDH2	67
4.4 Consequências para a Proteína RALDH2 com a Exclusão do Exon 4 pela Maquinaria de <i>Splicing</i>	68
4.5 Ensaio de <i>splicing</i> para o Polimorfismo A151G	69
4.5.1 <i>Splicing</i> Alternativo e Doenças Humanas.....	69
4.6 Ensaio de <i>Splicing</i> para a Mutação G151T.....	72
4.6.1 Mutações em Sítios de <i>Splicing</i> que podem Aumentar a Especificidade pelo Sítio ESE	72
4.7 Mutação Localizada na Região 5'UTR do Gene RALDH2.....	73
4.8 Estudos de Associação	73
4.8.1 Estudo de Associação do Tipo Caso-Controle para a Variante A151G	74
4.8.3 Análise da Estrutura de Desequilíbrio de Ligação	74
4.8.4 Teste de Desequilíbrio de Transmissão	75
5. CONCLUSÃO	77
REFERÊNCIAS	78

1. INTRODUÇÃO

1.1 Desenvolvimento Cardíaco

O coração é o primeiro órgão a se formar no embrião humano e também o primeiro órgão a exercer suas funções mesmo antes de seu completo desenvolvimento. Todos os eventos subsequentes na vida deste organismo dependem de sua habilidade em suprir a demanda embrionária por oxigênio e nutrientes de forma apropriada.

Nos estágios iniciais do desenvolvimento do coração, os progenitores cardíacos podem ser reconhecidos como um crescente epitelial (o crescente cardíaco) nas porções cranial e crânio-lateral do embrião. Após a especificação dos precursores cardíacos, ou seja, quando as células precursoras cardíacas são capazes de se diferenciar de maneira autônoma mas ainda reversível, as células do músculo cardíaco convergem ao longo da linha média ventral do embrião para formar um tubo cardíaco linear pulsátil composto de camadas distintas de endocárdio e miocárdio separados por matriz extracelular. Nesta etapa a região de influxo do coração está localizada caudalmente e a região de efluxo está localizada cranialmente. Em todos os vertebrados este tubo cardíaco sofre um dobramento que é essencial para a orientação apropriada dos ventrículos e átrios, bem como para o alinhamento correto das câmaras cardíacas com os grandes vasos sanguíneos. O tubo cardíaco é segmentado ao longo do eixo ântero-posterior em precursores que darão origem ao saco aórtico, ao conotronco (trato de saída), aos ventrículos e aos átrios. Nas etapas subsequentes ocorre a formação das câmaras cardíacas bem como a maturação e septação do coração, dividindo átrios e ventrículos (SRIVASTAVA & OLSON, 2000; HARVEY, 2002). Dada a complexidade do desenvolvimento cardíaco, é amplo o espectro de defeitos congênitos observados durante as várias etapas da morfogenia do coração e que vão aumentando em complexidade conforme se avança nos estágios de seu desenvolvimento (BRUNEAU, 2003).

1.2 Doenças Cardíacas Congênitas

Malformações cardíacas constituem o grupo mais freqüente de patologias congênitas em humanos, com uma incidência de aproximadamente 10 a cada 1000 nascimentos e são responsáveis pela vasta maioria das perdas pré – natais (PERLOFF, 1994; HOFFMAN, 1995). Além disto, 3 em cada 1000 nascidos vivos necessitarão de intervenção cirúrgica durante o primeiro ano de vida (GRUBER & EPSTEIN, 2004). Esta elevada freqüência representa um custo significativo à sociedade devido à necessidade de tratamento cirúrgico precoce e supervisão intensiva que acompanha estas patologias.

A última década se caracterizou pela transição dos estudos funcionais e fisiológicos para estudos mais profundos que se utilizam de abordagens genéticas e moleculares, possibilitando análises mais profundas a respeito das funções e disfunções do coração, principalmente no que diz respeito aos aspectos relacionados ao desenvolvimento e, conseqüentemente, à doenças congênitas (OLSON, 2004). Esses avanços incluem: (I) a descrição de um número cada vez maior de genes–alvo de doenças cardíacas humanas em modelos animais; (II) a disponibilidade da seqüência genômica completa de múltiplos organismos e (III) ferramentas cada vez mais sofisticadas de bioinformática (GRUBER & EPSTEIN, 2004) – e trouxeram novas oportunidades para os estudos de prevenção e tratamento destes tipos de anomalia. Nessa mesma época, genes que foram sendo descobertos e cuja função estava relacionada ao controle dos processos embrionários básicos em modelos animais, levantaram a possibilidade de que formas mutantes destes genes poderiam ser a causa de desordens congênitas humanas importantes.

No passado, as doenças cardíacas congênitas eram entendidas como doenças de base multifatorial e poligênica, quando sob influência genética. No entanto, trabalhando-se com famílias com vários membros afetados foi possível identificar um grupo de genes responsáveis por DCCs. Esse grupo, representado por genes como NKX-2.5 (LYONS *et al.*, 1995; HARVEY, 1996; SCHOTT *et al.*, 1998; REAMON-BUETTNER &

BORLAK, 2004), TBX-5 (BRUNEAU *et al.*, 2001; GARG *et al.*, 2003), GATA-4 (GARG *et al.*, 2003) e JAGGED (YUAN *et al.*, 1998), por exemplo, são causadores de formas típicas de DCCs tais como tetralogia de Fallot (TOF), defeitos de septação atrial (ASD) e defeitos de septação ventricular (VSD). Estas DCCs podem ser causadas por mutações em alelos únicos de genes “cardíacos” codificando fatores de transcrição, proteínas sarcoméricas, proteínas de matriz extracelular e canais iônicos. Provavelmente, o caso mais representativo é o do gene NKX2-5, que está envolvido em aspectos precoces do desenvolvimento cardíaco e que é expresso em todos os cardiomiócitos durante o desenvolvimento. Além disso, mutações no gene NKX2-5 levam à letalidade precoce de embriões de camundongos (HARVEY, 1996). O papel de NKX2-5 no desenvolvimento cardíaco parece ser conservado entre as espécies. Duas constatações reforçam essa hipótese: o homólogo deste gene em *Drosophila* – o gene *tinman* – é absolutamente necessário para o desenvolvimento da estrutura que desempenha o papel do coração em *Drosophila* e, além disso, experimentos onde o gene *tinman* foi substituído pelo gene NKX2-5 de camundongos mostram restauro no desenvolvimento do “coração” e das células progenitoras formadoras de musculatura visceral, como no organismo selvagem (HARVEY, 1996). Mais tarde, com estudos em famílias com membros que apresentavam defeitos de septação atrial, por exemplo, foi possível mapear os fatores genéticos que poderiam estar envolvidos na alteração do fenótipo, o que levou à identificação de mutações de base única no gene NKX2-5 como causadoras desta anomalia (SCHOTT *et al.*, 1998; BRUNEAU, 2003; XAVIER-NETO, 2005).

1.3 O Paradigma Da Haploinsuficiência

Anomalias resultantes de mutações que são encontradas em heterozigose são geralmente atribuídas à haploinsuficiência, um termo que sugere que a redução pela metade da expressão do gene é responsável pelo fenótipo observado. Isto indica que parece haver um nível ótimo, para cada uma das proteínas codificadas por esses genes, abaixo do qual o desenvolvimento normal do embrião não é possível. Um aspecto importante e que pode relacionar as DCCs ao paradigma da haploinsuficiência é o fato

de que, sintomaticamente, é virtualmente impossível encontrar-se pacientes com alelos defeituosos dos genes supracitados em homozigose. Trabalhos com os animais *knockout* para genes como NKX2.5 e TBX5, por exemplo, mostram que o inativação do gene alvo leva à interrupção do desenvolvimento cardíaco. Nos animais mutantes para o gene NKX2-5, por exemplo, o coração permanece como um tubo linear primitivo, que não se desenvolve em câmaras, levando à morte embrionária precoce, *in utero* (LYONS *et al.*, 1995; HARVEY, 1996; SCHOTT *et al.*, 1998; BRUNEAU *et al.*, 2001). Isto indica que perdas substanciais da função desses genes não são compatíveis com a sobrevivência na fase intra-uterina.

Mutações descobertas em estudos familiares por uma combinação de estudos de ligação, clonagem posicional e abordagem de genes candidatos ofereceram esperanças de que as DCCs poderiam ser eventualmente explicadas pela redução da dosagem (haploinsuficiência) de genes com papéis mais importantes no desenvolvimento cardíaco. Assim, as DCCs têm sido recentemente consideradas como o resultado de haploinsuficiência (insuficiência alélica) em genes fundamentais para múltiplos eventos do desenvolvimento cardíaco, cuja remoção do genoma está associada à morte embrionária em animais-modelo (GRUBER *et al.*, 1996; KASTNER *et al.*, 1997; SEIDMAN & SEIDMAN, 2002).

O acúmulo de informações tem levado à sugestão de que expressão gênica é controlada de maneira probabilística – mais do que determinística. Deste modo, as conseqüências moleculares da haploinsuficiência provavelmente precisarão ser revistas. Assim, proteínas – como fatores de transcrição, por exemplo – expressas em heterozigotos podem, em algumas células, não serem expressas pois não ocorrerá ativação do gene, ou essa ativação será esporádica, devido a instabilidade inerente causada por dosagem insuficiente de genes que são críticos em todo o desenvolvimento cardíaco (ROSENTHAL & HARVEY, 1999). Assim, em heterozigotos, a ativação de genes – alvos nas células, durante o desenvolvimento, pode ser determinada de forma casual, variando sua distribuição temporal e espacialmente entre indivíduos afetados por

mutações nestes genes. Analisando desta forma, os mecanismos probabilísticos da expressão gênica teriam implicações bastante importantes nas consequências morfológicas destas mutações. Em outras palavras, a haploinsuficiência de genes críticos aumenta as chances de desenvolvimento de malformações cardíacas por causa dos níveis de proteína ativa durante as múltiplas fases do desenvolvimento, que pode estar perigosamente perto do limiar abaixo do qual a morfogênese normal é impossível. Então, a haploinsuficiência pode comprometer a morfogênese cardíaca e predispor, em múltiplos períodos críticos, a uma gama de malformações disparadas por combinações não favoráveis de variantes alélicas de genes de desenvolvimento cardíaco, por distúrbios ambientais tais como as mudanças no estado nutricional (como por exemplo na disponibilidade de ácido fólico), ou por exposição à componentes químicos tóxicos (álcool, por exemplo). Desta forma, é importante desenvolver estratégias ambiciosas para identificar variantes genéticas e fatores ambientais que modulam o risco para DCCs. Existem inúmeras evidências de interações alélicas causando DCCs assim como da existência de modificadores da expressão da doença; no entanto, poucos dados estão disponíveis no que diz respeito aos fatores ambientais.

É importante, no entanto, entender as limitações do paradigma da haploinsuficiência em cardiopatias congênitas. Apesar de encontradas repetitivamente em famílias, estudos do gene NKX2.5 em pacientes com tetralogia de Fallot de origem não-sindrômica, mostraram que mutações neste gene estão representadas com baixíssima frequência, tipicamente inferior a 4%, em casos esporádicos (GOLDMUNTZ *et al.*, 2001). Ainda, haploinsuficiência de vários genes produzem fenótipos similares (heterogeneidade genética) e a mesma mutação em uma família é frequentemente não manifestada em outros membros (penetrância incompleta), ou são expressas como entidades funcional e anatomicamente diferentes (expressão variável) (BENSON *et al.*, 1998). Esses marcadores de DCCs indicam que pode haver uma base genética mais complexa do que puramente a explicação por haploinsuficiência.

Essas observações sugerem visões contrastantes sobre a etiopatogenia das doenças cardíacas. Em primeiro lugar é possível que existam mais genes, com características semelhantes aos já identificados em famílias causando DCC na população geral. Em outras palavras, seriam variantes genéticas raras conferindo um alto risco relativo. Contudo, uma porção substancial dos casos esporádicos também pode ser o resultado de interações entre variantes genéticas mais freqüentes, mas de menor risco relativo. Essa última visão representa um retorno parcial ao conceito de que as DCC seriam doenças poligênicas ou, mais propriamente, oligogênicas (RISCH, 2000; WEISS & TERWILLIGER, 2000; JAY *et al.*, 2005).

1.4 Doenças Complexas

São denominadas “complexas” as doenças que são, em última análise, determinadas por fatores que incluem tanto aspectos genéticos quanto ambientais. De fato, tanto a detecção quanto a caracterização de um fator contribuinte para o fenótipo de uma doença complexa são difíceis de serem conseguidos uma vez que a contribuição destes elementos pode ser pequena e, adicionalmente, ser camuflada por outros fatores. De fato, as diferentes vias metabólicas e os órgãos e sistemas fisiológicos atuam em conjunto em intrincados mecanismos bioquímicos e fisiológicos, que estão sob controle gênico e sujeitos à estímulos ambientais. Além disso, redundâncias e sistemas compensatórios são componentes que tornam o funcionamento do corpo humano um fenômeno biológico ainda mais complexo. Além destes aspectos, a estrutura genética de uma população em que se deseja estudar uma doença também pode ser bastante complexa uma vez que a população humana pode ser considerada bastante grande, além de ser relativamente antiga – suficiente para que haja um grande número de recombinações no genoma. Desta forma, essa população está sujeita à estratificações, divisões e redistribuições nos mais diversos ambientes. Diante destas características, diversos fenômenos podem ocorrer, tais como: herança poligênica, heterogeneidade genética, interações entre gene e ambiente e expressão de genes que são dependentes espacial e temporalmente. Esses fatores tornam a descoberta e a

caracterização de um gene bastante complicada mas, uma vez que genes possivelmente causadores de doença são identificados, estudos epidemiológicos podem ser utilizados de modo a se caracterizar a frequência de um gene deletério na população bem como os efeitos deste gene na presença de fatores ambientais (SCHORK, 1997).

Para a investigação de doenças de herança simples – Mendelianas – onde genes únicos conferem um alto risco relativo de causar doenças, os estudos de ligação e de clonagem posicional foram utilizados com grande sucesso. O mesmo não ocorreu na identificação de genes relacionados às doenças complexas, onde os estudos de associação têm sido amplamente utilizados. Estes estudos buscam uma correlação estatística entre variantes genéticas específicas e a doença. E é justamente o seu maior poder estatístico em detectar vários genes com efeitos pequenos que torna esse estudo mais robusto na investigação das doenças complexas (CARDON & BELL, 2001; RANSOM & SRIVASTAVA, 2007).

Um dos tipos de associação mais utilizados baseia-se na investigação de genes que já conhecidamente participam de vias importantes relacionadas à doença. Esse conhecimento é adquirido principalmente através de: (I) experimentos desenvolvidos em animais – modelo e que mostram os efeitos da manipulação do gene no organismo; (II) estudos de expressão gênica e (III) estudos em famílias afetadas. A partir da seleção destes genes candidatos é possível investigar variantes que podem causar mudanças na proteína ou em sua expressão ou estar co-segregando com mudanças funcionais ou ainda genotipar as variantes na população e usar ferramentas estatísticas para correlacionar estas variantes e o fenótipo. A partir desta perspectiva de associação, o estudo baseado em genes candidatos pode ser útil na exploração de vias potencialmente causais entre determinantes genéticos e doenças complexas (TABOR *et al.*, 2002).

Independentemente do modelo considerado é evidente que as respostas sobre a etiopatogenia em casos esporádicos de doenças cardíacas congênitas não poderão ser

respondidas com estudos em famílias afetadas, nas quais os estudos de ligação têm obtido considerável sucesso (TABOR *et al.*, 2002). Portanto, hoje é necessário desenvolver abordagens genéticas alternativas capazes de identificar genes de risco para o desenvolvimento de DCC esporádica.

1.5 Ácido Retinóico e Desenvolvimento

Genes pertencentes à via metabólica e de sinalização do ácido retinóico (AR) constituem alvos fascinantes à investigação como causadores diretos ou modificadores de DCCs. O AR é importante na padronização dos eixos do embrião e na organogênese em geral. O coração é particularmente afetado pelas variações na sinalização do ácido retinóico e a privação ou o excesso de vitamina A produz malformações cardíacas e vasculares, sugerindo uma ligação direta entre o estado nutricional e a teratogênese. Os efeitos cardíacos podem ser observados após a manipulação da sinalização do ácido retinóico nas fases de segmentação anteroposterior, no estabelecimento da lateralidade cardíaca e na proliferação e diferenciação dos cardiomiócitos ventriculares. Desde 1949, Wilson & Barch (WILSON & BARCH, 1949) já observavam que a privação de vitamina A durante a gestação causava malformações congênitas em vários órgãos / sistemas. Em 1950, Wilson & Warkany enfatizaram que a deficiência de ácido retinóico causava anomalias cardíacas congênitas em animais-modelo que mimetizavam aqueles que ocorriam espontaneamente no homem (WILSON & WARKANY, 1950). Na mesma época, Sawin demonstrava que fatores genéticos exerciam papel na determinação do arco aórtico em coelhos (SAWIN, 1949). Estes fatos já ilustravam que o desenvolvimento embrionário de um sistema poderia ser alterado tanto por fatores genéticos quanto por fatores ambientais. Em 1993, experimentos de privação de ácido retinóico em codorna propiciaram a observação dos efeitos em estágios precoces do desenvolvimento. Estes estudos puderam identificar os períodos no desenvolvimento embrionário no qual a vitamina A (ácido retinóico) é necessária para o desenvolvimento cardíaco; que o isômero all – *trans* do ácido retinóico é a forma biologicamente ativa necessária para o desenvolvimento cardíaco embrionário e que, além disso, podia-se recuperar o fenótipo

cardíaco normal em embriões deficientes em vitamina A com a administração de ácido retinóico (DERSCH & ZILE, 1993).

Dentre as enzimas sintetizadoras de ácido retinóico (RALDH1, RALDH2 e RALDH3), a RALDH2 foi identificada como a enzima mais eficiente na oxidação de retinaldeídos à ácido retinóico (ZHAO *et al.*, 1996). Além disso, é também a RALDH de atividade mais proeminente no desenvolvimento precoce de camundongos, sendo expressa em grande parte do tronco embrionário, e sendo que RALDH1 e RALDH3 têm expressão mais restrita à região ocular (NIEDERREITHER *et al.*, 1997). Também, a RALDH2 é a primeira RALDH a ter expressão detectada no embrião em desenvolvimento (NIEDERREITHER *et al.*, 1997; ULVEN *et al.*, 2000). Outro fato que mostra que a RALDH2 é a enzima mais importante no desenvolvimento embrionário se refere ao fato de que os embriões *knockout* que apresentam defeitos congênitos mais severos são os animais mutantes para o gene RALDH2 (NIEDERREITHER *et al.*, 1999; DUPE *et al.*, 2003; FAN *et al.*, 2003).

Em 1998, Moss e colaboradores demonstraram que: (I) os padrões de expressão do gene RALDH2 coincidia com os locais no coração onde o ácido retinóico era produzido, (II) que tanto a atividade da RALDH2 e do transgene controlado pelo elemento de resposta ao ácido retinóico (RARE) mostravam padrões distintos de expressão comparados ao padrão de expressão da enzima que degrada o ácido retinóico e que (III) a progressão da sinalização do AR durante o desenvolvimento cardíaco se dava de maneira dinâmica, de forma que a ação do AR progride caudorostralmente no tubo cardíaco linear do camundongo, estando completamente ausente das porções anteriores do coração (MOSS *et al.*, 1998). Desta forma, para estudar a especificação das estruturas de influxo do coração, foram gerados animais transgênicos por Xavier-Neto e colaboradores (XAVIER-NETO *et al.*, 1999) que possuíam o gene da fosfatase alcalina humana dirigido por um fragmento do promotor do gene de codorna SMyHC3 (SMyHC3-HAP), cuja expressão se sobrepunha aos locais de síntese e resposta ao ácido retinóico. Com estes estudos foi sugerido que a

expressão gênica atrial – específica seria controlada pela síntese localizada do ácido retinóico e que a exclusão do AR dos precursores ventriculares é essencial para a correta especificação dos ventrículos (ROSENTHAL & XAVIER-NETO, 2000; XAVIER-NETO *et al.*, 2001). Para entender como e quando o AR afeta os precursores cardíacos, foram examinados os padrões de expressão do gene RALDH2 em galinha e em camundongo na mesoderme lateral do embrião em desenvolvimento. A sinalização anteroposterior pelo ácido retinóico, analisada a partir da expressão de RALDH2, é caracterizada por um envolvimento progressivo dos precursores cardíacos por um campo de RALDH2 originado na forma de uma onda caudorostral dinâmica e evolucionariamente conservada, na mesoderme lateral (HOCHGREB *et al.*, 2003; SIMOES-COSTA *et al.*, 2005). Além disso, estudos em embriões de camundongos *knockout* para o gene RALDH2 mostraram que a produção do ácido retinóico por esta enzima é essencial para a sobrevivência do embrião e para a morfogênese precoce. Como resultado, estes animais morrem *in utero* e apresentam malformações cardíacas que consistem da formação de uma única cavidade, anterior, dilatada, ausência do dobramento cardíaco e diferenciação inapropriada dos cardiomiócitos ventriculares. Estes defeitos resultam do bloqueio da síntese de ácido retinóico, uma vez que a administração materna de ácido retinóico resgata a morfogênese cardíaca normal (NIEDERREITHER *et al.*, 1999; NIEDERREITHER *et al.*, 2001). Finalmente, utilizando-se animais que carregavam alelos hipomórficos do gene RALDH2 e que também morriam *in utero*, foi possível demonstrar defeitos cardíacos que exibiam as características da síndrome de Di George humana como defeitos de septação do trato de saída cardíaco e defeitos aórticos. Este estudo mostrou ainda que a diminuição da disponibilidade de ácido retinóico – seja por intervenção genética ou nutricional – poderia representar o principal modificador da expressão dos aspectos clínicos da síndrome de Di George (VERMOT *et al.*, 2003). Estudos em galinha mostraram também que não só a privação de ácido retinóico mas também o excesso de AR pode prejudicar a formação do embrião, levando a uma ampla gama de malformações, desde a completa ausência do coração até desenvolvimento incorreto do eixo anteroposterior (OSMOND *et al.*, 1991).

Os efeitos teratogênicos do excesso de ácido retinóico também foram observados em ratos, com grandes efeitos na ectoderme e na mesoderme (MORRISS, 1972).

1.6 A Via de Síntese do Ácido Retinóico

O AR é obtido da dieta, na forma de precursores tais como beta caroteno e vitamina A (retinol) que não podem ser sintetizados *de novo* pelos vertebrados ou da mobilização de estoques endógenos na forma de retinaldeídos ligados à proteínas do vitelo ou de ésteres retinil do fígado. Todos estes precursores de AR derivam do metabolismo de grandes compostos terpenóides, os carotenos. Estes precursores atuam como moléculas sinalizadoras em uma ampla faixa de processos biológicos, especialmente em desenvolvimento. Extensas evidências indicam que o desenvolvimento cardíaco é particularmente sensível à oferta de vitamina A. A vitamina A é processada à ácido retinóico através de duas etapas oxidativas consecutivas: de retinol à retinaldeído e de retinaldeído à ácido retinóico. Estas etapas são bem conhecidas e consideradas como a via canônica da biossíntese nos vertebrados amniotos. A etapa de formação (oxidação do retinol a retinaldeído) consiste de uma reação de oxidação reversível e requer a atividade de diversos membros da família das álcool desidrogenases (ADH) (PERLMANN, 2002). Trabalhos em camundongos mutantes homozigotos para ADH1, ADH3 e ADH4 indicam que as enzimas ADHs são as principais enzimas na conversão de retinol à retinaldeído em amniotos. As enzimas ADHs possuem uma distribuição tecidual bastante ampla, sendo a ADH3 ubíqua (MOLOTKOV *et al.*, 2002). Dessa forma pode-se considerar que todas as células embrionárias são capazes de oxidar retinol a retinaldeído e que a disponibilidade de AR é ditada pelas RALDHs (ALDH1As). A etapa posterior – oxidação do retinaldeído para ácido retinóico *all-trans* – é irreversível e requer a ação de membros da família das retinaldeído desidrogenases (RALDH) (PERLMANN, 2002) (figura 1), dentre as quais a RALDH2 é a principal enzima sintetizadora de ácido retinóico no embrião em desenvolvimento e durante o qual é expressa na mesoderme embrionária (ZHAO *et al.*, 1996; NIEDERREITHER *et al.*, 1997). Em contraste com a atividade das ADHs, que são

ubíquas devido à expressão onipresente de ADH3 e à ampla distribuição de ADH4 e ADH1, as RALDHs têm uma distribuição mais restrita (BLENTIC *et al.*, 2003) e perfis de expressão que quase completamente coincidem com a ativação da sinalização pelo AR. Dessa forma, pode-se concluir que são as enzimas RALDHs que delimitam, os locais onde o AR é sintetizado (revisado em (SIMOES-COSTA *et al.*, 2006).

Publicações bastante recentes têm revelado novos aspectos da via de síntese do ácido retinóico. A reação que oxida o retinol à retinaldeído é primariamente catalisada pelos membros da família das ADHs (álcool desidrogenases). No entanto, outra família de enzimas é capaz de catalisar esta etapa de formação. Um membro desta família – a enzima RDH10 – desempenha papel crítico na síntese de ácido retinóico no embrião. Este estudo mostra que proteínas mutantes para este gene perdem sua capacidade de oxidar retinol à retinaldeído, resultando em sinalização insuficiente pelo ácido retinóico e levando a malformações diversas, inclusive à morte embrionária *in utero*. Esses novos dados contestam a noção de que a oxidação de retinol à retinaldeído é uma etapa de caráter redundante e que ocorre uniformemente através do embrião (SANDELL *et al.*, 2007). Outro artigo, também de 2007, revela novos aspectos sobre o mecanismo molecular pelo qual a vitamina A é absorvida pelas células - mecanismo esse bem pouco conhecido até então. O principal carreador de vitamina A (retinol) do sangue para outros órgãos é a *retinol binding protein* (RBP). Neste estudo os autores identificaram um domínio protéico multitransmembrana – STRA6 – como sendo um receptor de membrana específico para RBP. Dessa forma, STRA6 parece ser o principal mediador da oferta de vitamina A celular (KAWAGUCHI *et al.*, 2007). Além disso, um estudo genético humano recente mostrou que mutações no gene humano STRA6 está associado com um amplo espectro de doenças congênitas – que incluem o coração (PASUTTO *et al.*, 2007). Por fim, em 2006, Chambers e colaboradores demonstraram um novo papel para uma enzima da família P450, a CYP1B1. As enzimas P450 fazem parte de uma grande família de proteínas evolutivamente conservadas. Dentre estas, as enzimas mais proeminentes durante o desenvolvimento são as CYP26s, que atenuam a extensão espaço-temporal da sinalização pelo ácido retinóico (ver mais adiante). No entanto, os

autores mostram a caracterização de uma enzima pertencente à família das CYPs – a CYP1B1 – como sendo uma enzima sintetizadora – e não degradadora – de ácido retinóico. Este trabalho foi o primeiro a mostrar que uma enzima da família P450 pode direcionar positivamente os processos de padronização mediados por ácido retinóico, mostrando, inclusive, a contribuição da enzima CYP1B1 na padronização anteroposterior do embrião de galinha (CHAMBERS *et al.*, 2007).

Apesar destas novas perspectivas quanto à via de síntese do ácido retinóico, evidências indicam ainda que a segunda etapa da síntese de AR prevalece como sendo a mais importante. Primeiro, o fenótipo do *knockout* de RALDH2 é claramente mais severo do que o observado para o mutante RDH10. Enquanto o embrião RALDH2^{-/-} morre por volta do estágio E9.5, os embriões mutantes RDH10 conseguem sobreviver até o estágio E.13 indicando que os embriões RALDH2^{-/-} sofrem uma maior perda de sinalização pelo AR. Além disso a enzima RDH10 faz parte de uma família de enzimas capazes de oxidar retinol à retinaldeído (a família das SDR – *shot-chain dehydrogenase / reductase family*) e, até agora, não foram observados fenótipos em embriões *knockout* para SDR ou ADH, sugerindo que estas famílias de enzimas não são necessárias individualmente para a síntese precoce de AR. Além disso, apesar de evidências sugerirem que a enzima CYP1B1 parece ser importante na síntese de AR, durante o desenvolvimento do embrião não se observam sobreposições entre os padrões de expressão desta enzima e da RALDH2, a não ser nos somitos. Isto indica que RALDH2 continua a ser a enzima mais importante na síntese do AR na região cardíaca. Quanto à via de sinalização pelo AR, algumas evidências também contestam a importância da via que transporta esta molécula aos locais-alvo. Uma delas é o fato de que camundongos *knockout* para a proteína que se liga ao ácido retinóico (RBP) são animais viáveis e férteis, além de não apresentarem fenótipos aparentes em estruturas do embrião (QUADRO *et al.*, 1999).

Depois de sintetizado, o AR atua de forma autócrina, parácrina e endócrina, cruzando a membrana celular e se ligando à seus receptores nucleares, *retinoid acid*

receptors (RARs), que são ativados pelo all-*trans* e 9-*cis* RA (os dois principais isômeros do ácido retinóico) e *retinoid X receptors* (RXRs), que são ativados especificamente pela 9-*cis* RA. Para desempenhar seu papel como receptores nucleares, RXRs formam heterodímeros com RARs (KASTNER *et al.*, 1997) e este complexo atua como fatores de transcrição interagindo especificamente com uma sequência do DNA (*retinoic acid response elements* ou RAREs) presentes em regiões regulatórias de genes susceptíveis ao controle da expressão gênica. Estudos em animais *knockout* para o gene RXRa demonstraram que o efeito diminuído da dosagem gênica em embriões heterozigotos gera um espectro de malformações cardíacas (defeitos no saco aórtico, no conotronco, na formação de válvulas e na formação e diferenciação do músculo ventricular) que em muito se assemelham às malformações cardíacas encontradas em humanos. Além disso, estes trabalhos mostram que animais duplos *knockout* para o gene RXRa morrem ainda *in utero* (GRUBER *et al.*, 1996; KASTNER *et al.*, 1997).

AR é degradado a metabólitos parcialmente ativos ou inativos pelas enzimas P450 tais como CYP26A, B e C (figura 1). A superfamília do citocromo P450 (CYP) catalisa um grande número de importantes reações biológicas. Mais especificamente, a CYP26 é essencial para a produção de 4-hidroxi-ácido retinóico e 4-oxo-ácido retinóico que são produtos de degradação inativos do ácido retinóico (PERLMANN, 2002). A função da CYP26, particularmente, parece ser a de proteger tecidos contra a exposição anormal ao ácido retinóico pois ela assume um padrão de expressão complementar ao da enzima produtora de ácido retinóico (SWINDELL & EICHELE, 1999; SWINDELL *et al.*, 1999), estabelecendo uma distribuição desigual do ácido retinóico ao longo do eixo anteroposterior do animal. Além disso, os animais mutantes homozigotos para o gene CYP26 também apresentam defeitos cardíacos (como tubo cardíaco dilatado, por exemplo) e morrem ainda durante os estágios embrionários ou mesmo logo após o nascimento (ABU-ABED *et al.*, 2001; SAKAI *et al.*, 2001). Esta enzima é encontrada em estágios precoces de embriões de camundongo e de galinha e experimentos *in vitro* e *in vivo* mostram, ainda, que sua expressão é regulada pelo ácido retinóico (RAY *et al.*, 1997; REIJNTJES *et al.*, 2004). Além disso, estudos utilizando-se animais duplo

mutantes nulos para o gene CYP26a1 e heterozigotos para o gene RALDH2 levaram à observação de que o fenótipo de CYP26a1^{-/-} podia ser resgatado sob condições nas quais os níveis de ácido retinóico estivessem diminuídos, ou seja, pela haploinsuficiência de RALDH2. Este estudo foi o primeiro a reportar uma mutação atuando como supressor dominante de uma mutação letal em mamíferos. Utilizando-se um transgene responsivo à ácido retinóico foi possível demonstrar que embriões CYP26^{-/-} têm atividade bastante aumentada do transgene e, desta forma, altos níveis de ácido retinóico do que os controles. Esses achados fornecem evidência de que o fenótipo de CYP26^{-/-} resulta dos efeitos deletérios do excesso de ácido retinóico devido à falta de catabolismo tecido-específica, mais do que pela sinalização dos metabólitos bioativos do ácido retinóico (NIEDERREITHER *et al.*, 2002).

Desta forma, RALDHs, RXRs, RARs e CYPs são todos alvos potenciais para alterações genéticas que poderiam produzir, ou pelo menos modificar, a expressão das DCCs. A recente descrição de fenótipos que se assemelham àqueles encontrados em pacientes portadores dos fenótipos encontrados na síndrome de DiGeorge (síndrome da deleção 22q11) em camundongos portadores de alelos *knockout* ou hipomórficos do gene *RALDH2*, assim como a expressão ectópica do gene *RALDH2* na mesoderme faríngea em camundongos duplos *knockout* Tbx-1 e Crkl sugere que os genes envolvidos na sinalização e no metabolismo do ácido retinóico são modificadores dos fenótipos de doenças cardíacas congênitas (GURIS *et al.*, 2006).

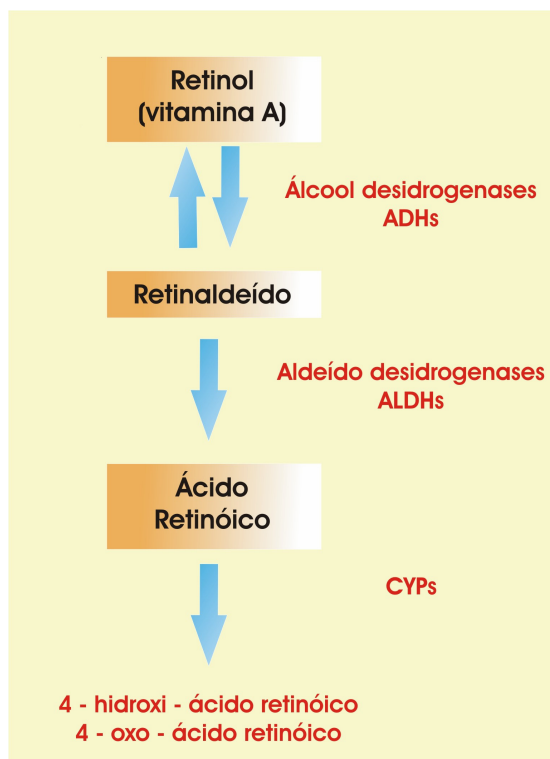


Figura 1 - A via de síntese e degradação do ácido retinóico. O retinol é oxidado à retinaldeído pelas ADHs (álcool desidrogenases) em uma etapa reversível e oxidado a ácido retinóico em uma etapa irreversível pelas ALDHs (aldeído desidrogenases), sendo a ALDH1a2 (RALDH2) a enzima mais importante neste processo. A etapa de degradação é mediada pelas enzimas da família P450 de citocromo – oxidases, as CYPs.

1.7 Variantes no Gene Raldh2 e Doenças Cardíacas Congênitas

Para estabelecer a relação entre as variantes gênicas do metabolismo do AR com as DCCs nós: (I) sequenciamos o gene da RALDH2 em pacientes portadores de doença cardíaca congênita em busca de variantes que pudessem conferir risco de modificar o fenótipo de DCCs e, em seguida, nós (II) investigamos, através de um estudo de associação genética baseado tanto em família como na população em geral, se a variação genética no gene *RALDH2* está associada com a modulação da susceptibilidade ou proteção às DCCs.

2. MATERIAIS E MÉTODOS

2.1 Desenho de Estudo

Combinando informações de marcadores polimórficos que foram encontrados por nós e cuja frequência na população foi maior do que 1%, com informações de SNPs selecionados de bases de dados públicas, de forma a cobrir todo o gene, nós determinamos a estrutura de desequilíbrio de ligação deste locus em nossa população. Uma alteração anteriormente descrita por Deak e colaboradores (DEAK *et al.*, 2005) e que foi associada à espinha bífida foi também utilizada. Finalmente, utilizando-se estas informações, nós testamos uma possível associação entre a variabilidade genética no gene *RALDH2* e as DCCs através de um teste de desequilíbrio de transmissão (TDT) que empregou trios compostos de indivíduos afetados e seus pais.

2.2 Estudo de Pacientes

Foram selecionados 101 trios (os pacientes - com cardiopatias congênitas diversas - e seus pais) que são atendidos no Ambulatório de Cardiologia Pediátrica, no Instituto do Coração (InCor – HCFMUSP), São Paulo, Brasil. Pacientes foram selecionados com base no diagnóstico de defeitos cardíacos congênitos e seus pais também foram convidados a participar do estudo. Sangue venoso foi obtido para extração de DNA genômico de cada paciente do estudo. O protocolo foi aprovado pelo Comitê de Revisão Institucional da Faculdade de Medicina da Universidade de São Paulo. Todos os participantes leram e assinaram um termo de consentimento aprovado. O DNA foi extraído de leucócitos de sangue periférico pelo procedimento *salting-out* padrão.

2.3 Ferramentas de Bioinformática e Amplificação por PCR

A localização (cromossomo 15) e a estrutura do gene RALDH2 (ALDH1A2) foram obtidas através da base de dados pública Ensembl (http://www.ensembl.org/Homo_sapiens/geneview?gene=ENSG00000128918). Os polimorfismos analisados no estudo baseado em família foram escolhidos na base de dados dbSNP (www.ncbi.nlm.nih.gov/SNP) com base em sua frequência e localização, de forma a cobrir todo o gene (quadro 1). Os pares de primers foram desenhados com o *Interactive Software Primer Design – GeneFisher* (www.genefisher.de). Para a análise de sítios de *splicing* foi utilizado o programa *ESEfinder* disponível na rede mundial de computadores (<http://www.genetics.med.ed.ac.uk/tutorials/msc/2005/frames.html>). Para avaliar as possíveis consequências estruturais no dobramento do RNA foi utilizado o programa Genebee (<http://www.genebee.msu.su/>). Para caracterizar as mutações encontradas no gene RALDH2 com relação à localização espacial na proteína foi utilizado o programa de bioinformática VMD (*Visual Molecular Dynamics*) (HUMPHREY *et al.*, 1996).

Quadro 1 – Marcadores e primers utilizados no estudo de TDT. Os marcadores foram escolhidos com base em sua frequência e localização.

SNP	Foward Primer	Reverse Primer	Localização	Tamanho do Fragmento	Alelos
1. rs2704188	tcccctaggctcctatctatgttatggatc	tcccttaagttgccacacag	região 5'	97 pb	C / T
2. rs1441815	attaatgccattataaaaggccttgagaat	ggtgaacgacttcct	intron 7	113 pb	G / T
3. rs3784259	gaaagtcattgcagca	acatgaaccaccacaca	intron 8	426 pb	A / G
4. rs1530293	cctccagtgtcccagagcccagacagtgt	gacgtggagagaggaagcag	região 3'	137 pb	A / G
5. rs1899430	gctctagaccctaaaaagccctagctttgaa	aaaatcatggaggaggagca	região 3'	179 pb	C / T

2.4 PCR e Sequenciamento de DNA

Na reação de PCR, tanto para o sequenciamento do gene RALDH2 quanto para o estudo de TDT (quadros 1 e 2), foi utilizada a enzima Easy Taq polimerase (Invitrogen) e seus reagentes apropriados. O programa utilizado para a amplificação por PCR consiste de uma etapa de denaturação inicial por 5 min. a 94°C, 35 ciclos e extensão

final por 10 min. a 72°C. Os ciclos principais consistem de 30s. à 94°C, temperatura de anelamento (quadro 3) por 30s. e 72°C por 3min. Os diferentes alelos dos marcadores utilizados no estudo de TDT foram identificados entre os indivíduos através de enzimas de restrição apropriadas. As seqüências de DNA foram determinadas por sequenciamento direto bidirecional. Para o sequenciamento, os produtos de PCR foram purificados com o kit E.Z.N.A. Cycle-Pure Kit (200), Omega Bio-Tek, USA, de acordo com as instruções.

Quadro 2 – Primers utilizados para o sequenciamento do gene RALDH2.

Exon / RALDH2	Foward Primer	Reverse Primer	Fragmento Amplificado
1	gggcgtcctaggtaggaa	cctcaccgctgaagaga	371 pb
2	ttcctaaattggcgttgagtc	tgctttgtctatatctgcctgtta	448 pb
3	tcagttgtgcctcttctc	tttcagcagaatggcaaat	364 pb
4	gatatgggttactggaagca	gatgtattaagtggcatgcc	589 pb
5	actttctgggagtgacca	aacacacatcgctgagga	186 pb
6	gagatttccaaggtcagtgctatt	aggaggaaacgtggctgatgaaag	433 pb
7	cttcgggagtgcttctga	cctttgtgggcttgggta	353 pb
8	tacccactaacttacatcttg	agtgtctttccccattatt	435 pb
9	ctagggccttggagagg	aaggacagagaagcataagaaat	462 pb
10	aaaaaccctaaaatgaaag	ttactgttgcaaaatgaata	381 pb
11	gactcttctgcataatctt	ggagtgtacccttttct	462 pb
12	tacactcctactgcttagat	gctccacgaaatgtttgtgaat	367 pb
13	cgcatacgtgaggaattttacat	ggtggccccttacagagtgc	418 pb

Quadro 3 – Temperatura dos primers utilizados no estudo de TDT e no sequenciamento do gene RALDH2.

Fragmento	Temperatura °C
RALDH2 exon 1	56.7
RALDH2 exon 2	60.7
RALDH2 exon 3	56.7
RALDH2 exon 4	56.7
RALDH2 exon 5	56.7
RALDH2 exon 6	61.8
RALDH2 exon 7	56.7
RALDH2 exon 8	50.5
RALDH2 exon 9	61.4
RALDH2 exon 10	48.2
RALDH2 exon 11	50.5
RALDH2 exon 12	53.4
RALDH2 exon 13	59.6
rs 1899430	56.7
rs 1530293	63.7
rs 1441815	45
rs 2704188	56.7
rs 3784259	56.7

2.5 Clonagem, Cultura de Células e Extração de RNA

Primeiramente, foi utilizada a técnica de PCR para amplificar o exon 4 do gene RALDH2 de um paciente e de um indivíduo controle. Estes fragmentos foram clonados no plasmídeo PCR-Script (PCR-Script Amp Cloning Kit da Stratagene, CAT# 21189.5). Para certificar que mutações não foram geradas durante a etapa de PCR, os plasmídeos foram sequenciados. Os fragmentos foram então subclonados no vetor pSPL3, específico para o estudo de *splicing* (CHURCH *et al.*, 1994). Para a clonagem no vetor pSPL3, o plasmídeo PCR-Script contendo o fragmento referente ao exon 4 (polimórfico e controle) foi digerido com as enzimas de restrição *EcoRV* and *SstI*. As mesmas enzimas foram utilizadas para digerir o vetor pSPL3. O vetor e o produto de digestão foram purificados em gel de agarose 1%.

Células COS - 7 e 293 EBNA foram cultivadas em meio DMEM com 10% de FBS, à 37°C, em 5% de CO₂. Foram utilizadas placas de 6 poços e as células foram cultivadas

até aproximadamente 75% de confluência. Para a transfecção nós utilizamos o carreador Lipofectamine™ Reagent (Invitrogen) de acordo com as instruções do fabricante. Cada poço foi transfectado com 1,6 ug de DNA. Tanto o experimento utilizando-se células COS-7 quanto o experimento utilizando-se células 293 EBNA foram repetidos duas vezes cada.

O RNA foi extraído 48h após o início da transfecção utilizando-se o reagente Trizol (GIBCO BRL, USA) de acordo com as instruções. As seguintes seqüências de primers foram utilizadas para amplificar os fragmentos de cDNA, sintetizados através de RT – PCR: SA2 - ATC TCA GTG GTA TTT GTG AGC; SD6 – TCT GAG TCA CCT GGA CAA CC (CHURCH *et al.*, 1994).

2.6 Análises Estatísticas

As frequências alélicas e genótípicas foram calculadas através de *gene counting*. O equilíbrio de Hardy-Weinberg foi testado com o teste de qui-quadrado. Nós utilizamos o software Haploview 3.2 para determinar o desequilíbrio de ligação entre os alelos, bem como para a determinação de haplótipos. Uma possível associação entre a presença de um alelo genotipado em particular e a doença foi testado através do teste TDT, avaliado com o software Haploview. Valores de p menores do que 0,05 foram considerados significantes.

3. RESULTADOS

3.1 Plano Inicial do Trabalho

1. O objetivo inicial do projeto era o de rastrear variantes no gene RALDH2 – sequenciando todos os exons (13 ao todo) – com capacidade de modular a susceptibilidade ou proteção às doenças cardíacas congênitas;
2. Nesta etapa nós encontramos duas variantes localizadas em um ESE (*exonic splicing enhancer*): A151G e G151T. Nós analisamos o dobramento do RNA em presença destas variantes e desenvolvemos ensaios funcionais de *splicing* para ambas;
3. Na segunda fase do projeto nós determinamos a estrutura haplotípica do gene RALDH2 com base em SNPs (*single nucleotide polymorphisms*) encontrados na rede mundial de computadores e, a partir daí, desenvolvemos um estudo de associação robusto do tipo TDT (*transmission disequilibrium test*) com estes marcadores;
4. Aplicamos também um estudo de associação simples, do tipo caso–controle, para o polimorfismo A151G (rs16939660) anteriormente associado à espinha bífida por Deak e colaboradores (DEAK *et al.*, 2005).

3.2 Sequenciamento do Gene RALDH2

Nós investigamos a hipótese de que mutações no gene RALDH2 poderiam estar relacionadas a doenças cardíacas congênitas. Desse modo, para rastrear variantes existentes no gene RALDH2 e que poderiam estar modificando a função do gene em pacientes portadores de cardiopatias congênitas, nós sequenciamos todos os exons do gene RALDH2 em aproximadamente 50 pacientes portadores de tetralogia de Fallot. A tetralogia de Fallot foi a doença cardíaca escolhida uma vez que é encontrada em maior quantidade em nosso ambulatório.

O gene da RALDH2 (ALDH1A2) é um gene relativamente grande, com 113.000pb e 13 exons (NM_003888) (figura 2), sendo a parte codificante composta de 1557pb (CCDS10163.1).

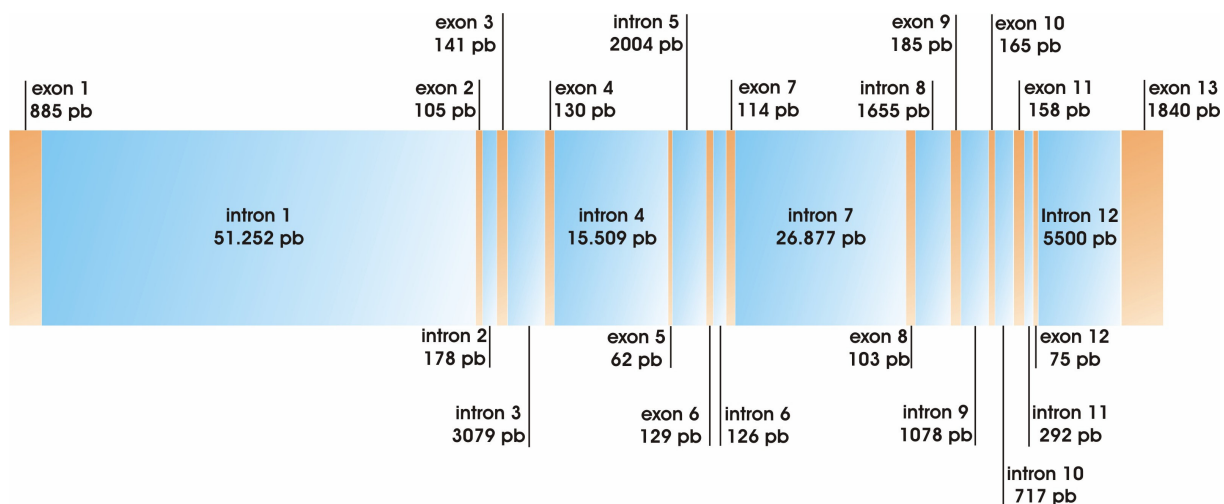


Figura 2 – Estrutura do gene RALDH2. O gene RALDH2 possui 13 exons e 113.000pb. A parte codificante é composta de 1557pb.

Nesta etapa nós encontramos duas mutações no exon 4, não descritas anteriormente. Uma mutação muda o nucleotídeo T para um nucleotídeo C (de ATT para ACT), responsável pela troca do aminoácido isoleucina pelo aminoácido treonina (figura 3). Esta mutação foi encontrada em um paciente. A troca do aminoácido isoleucina pelo aminoácido treonina representa uma troca de um aminoácido apolar para um aminoácido polar.

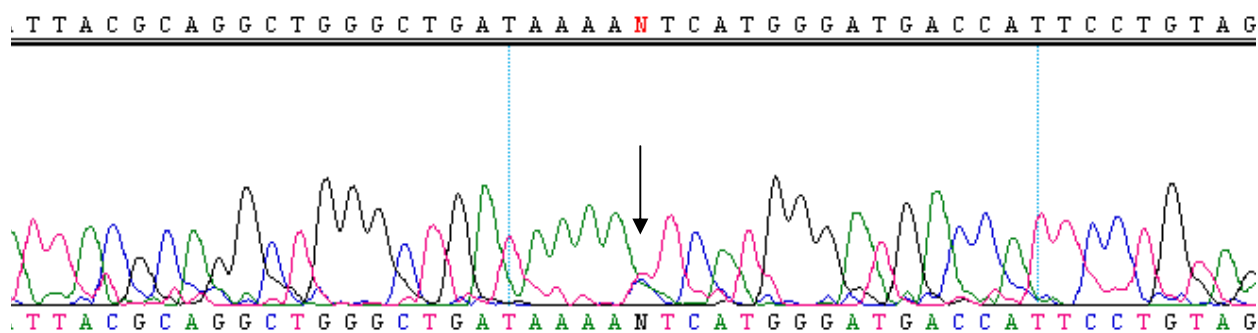


Figura 3 – Mutação de troca T por C, encontrada na posição 157, no exon 4, durante o sequenciamento do gene RALDH2 (T157C). Essa mutação troca o aminoácido isoleucina pelo aminoácido treonina.

A outra mutação é uma troca de G por T (de GCA para TCA), responsável pela troca do aminoácido alanina (apolar) pelo aminoácido serina (polar) (figura 4). Esta mutação também foi encontrada em um paciente.

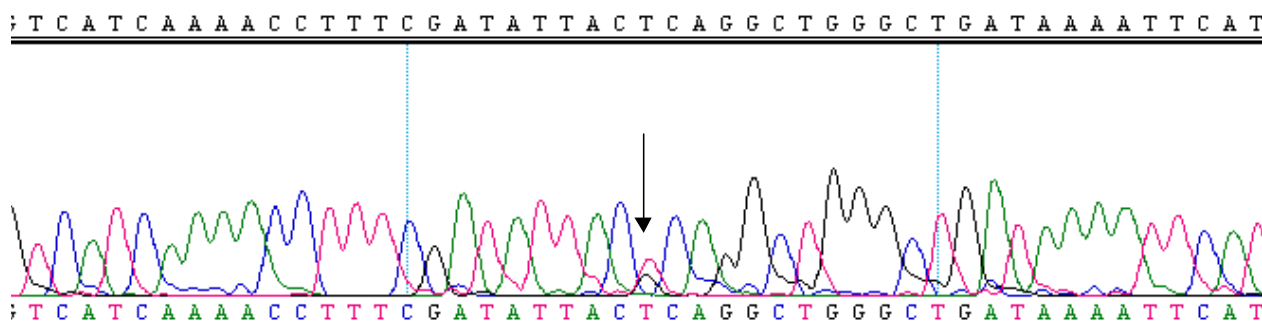


Figura 4 - Mutação de troca G por T, encontrada na posição 151, no exon 4, durante o sequenciamento do gene RALDH2 (G151T). Essa mutação troca o aminoácido alanina pelo aminoácido serina.

3.2.1 Estudo de Modelagem Molecular para as Mutações G151T e T157C

Para se ter uma idéia das alterações estruturais que estas mutações poderiam causar, um estudo de modelagem molecular foi realizado com a ajuda do aluno Tiago José Paschoal Sobreira, orientado pelo Dr. Paulo Sérgio Lopes Oliveira, ambos desenvolvendo suas atividades no Instituto do Coração – InCor. Com esta análise nós caracterizamos a estrutura protéica do gene RALDH2 e localizamos os aminoácidos mutados G151T e T157C (figuras 5 e 6, respectivamente). Isso foi possível através do programa de bioinformática VMD (*Visual Molecular Dynamics*) (HUMPHREY *et al.*, 1996).

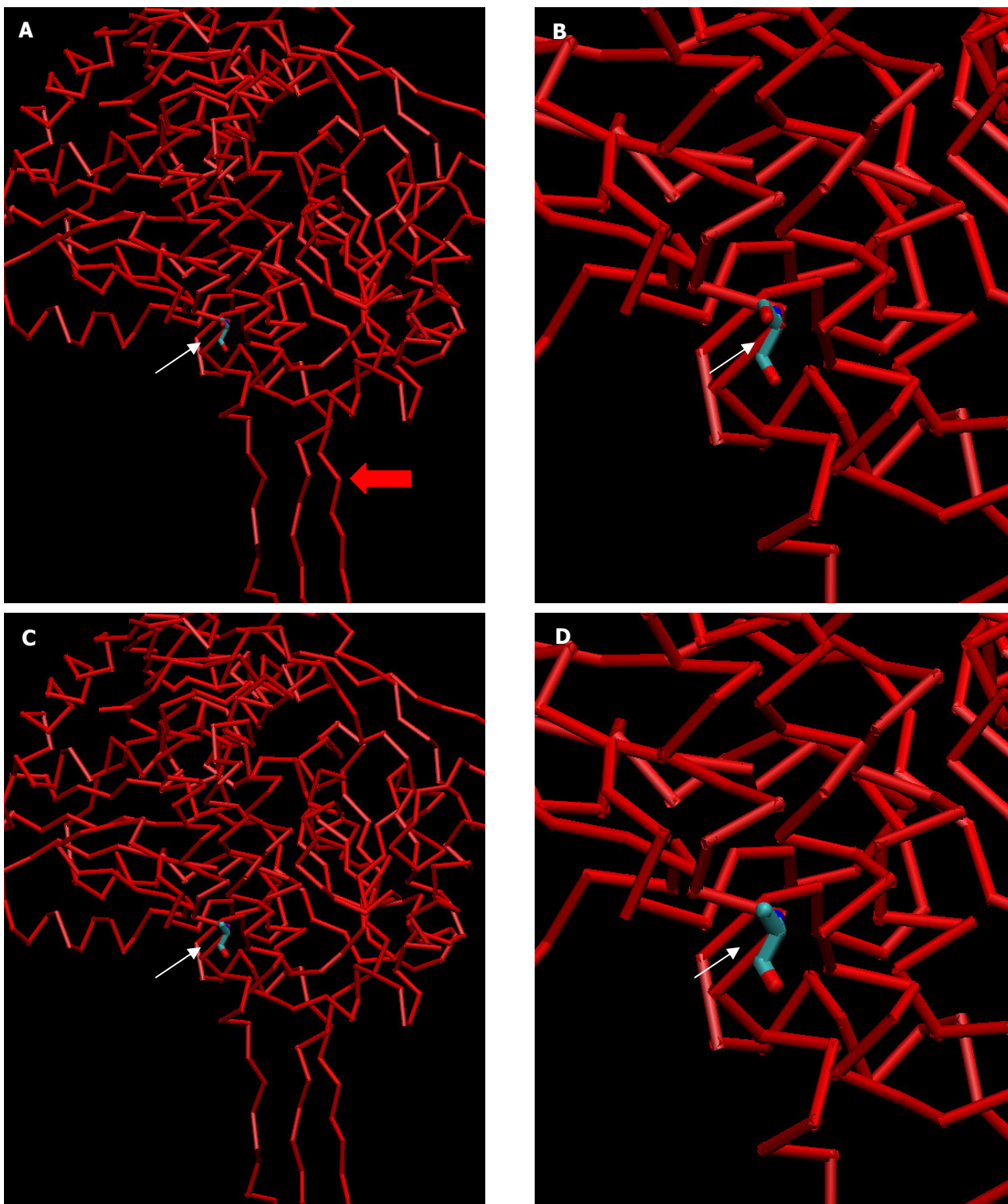


Figura 5 - Estrutura protéica da RALDH2 (vermelho). A e B: estrutura normal da proteína; aminóácido alanina (em destaque). C e D: proteína mutada; aminóácido serina (em destaque). A seta vermelha indica o segundo sítio de tetramerização da proteína (exon 13).

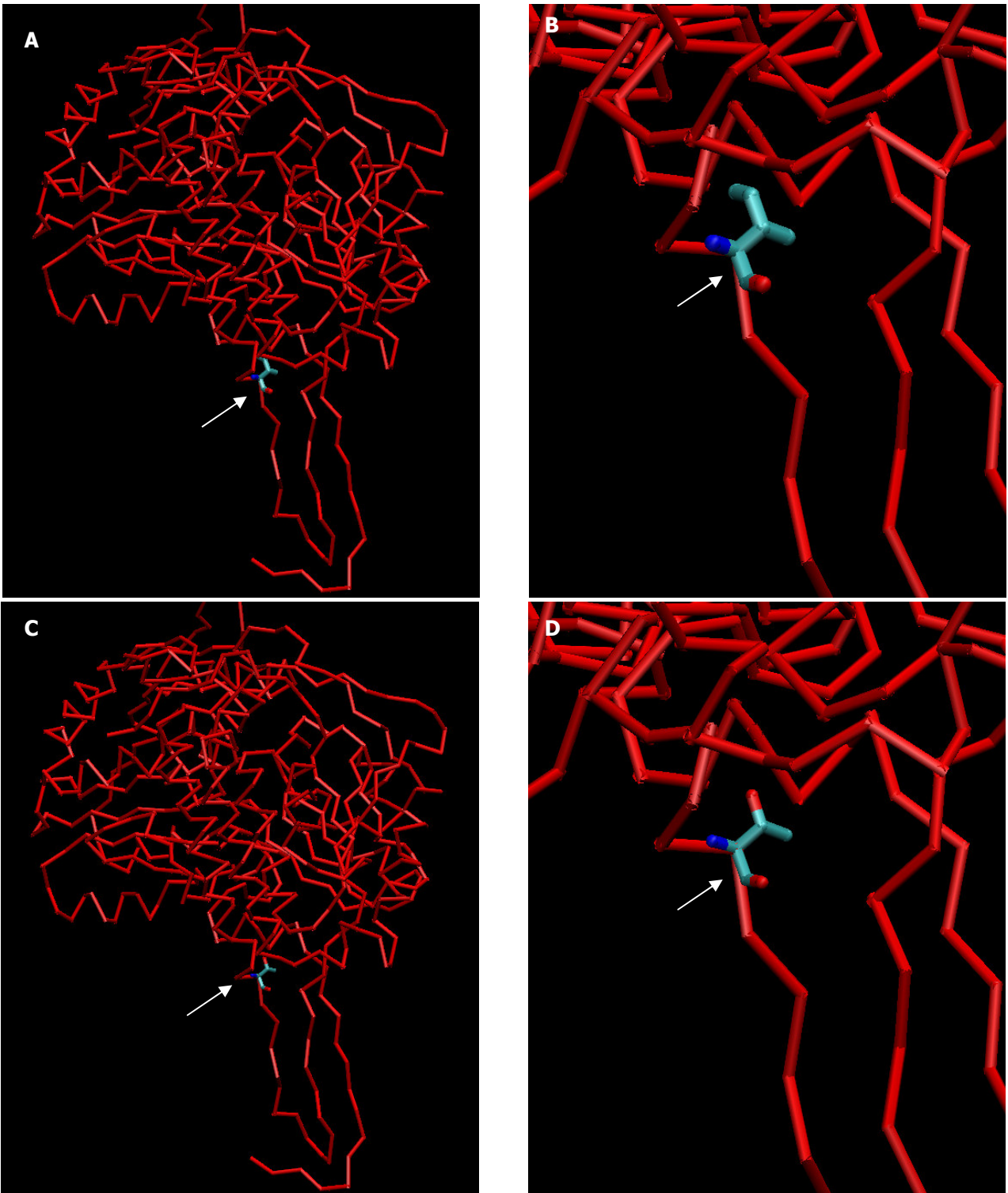


Figura 6 - Estrutura protéica da RALDH2 (vermelho). A e B: estrutura normal da proteína; aminoácido isoleucina (em destaque). C e D: proteína mutada; aminoácido treonina (em destaque).

A enzima RALDH2 é formada por 3 domínios: um domínio de ligação (1 – 136, 161 – 270), um domínio catalítico (271 – 484) e um domínio de tetramerização (137 – 160 [exon 4] e 485 – 500 [exon 13]) (LAMB & NEWCOMER, 1999), ambos os subdomínios de tetramerização estão muito próximos entre si na estrutura tridimensional da proteína (figura 5a, seta vermelha). Mutações nestes sítios podem diminuir ou até impedir a tetramerização da proteína. Além disso, trabalhos com o gene ALDH2 mostram que uma mutação de troca G por A, que substitui o aminoácido glutamato pelo aminoácido lisina na posição 487 do gene (exon 13 – sítio de tetramerização) estão relacionadas a deficiências no metabolismo de álcool na população asiática. Homozigotos para o alelo mutante ALDH2*2 não têm níveis detectáveis de atividade da enzima ALDH2 e, conseqüentemente, após a ingestão de álcool os indivíduos homozigotos sofrem com altos níveis de acetaldeído. Os indivíduos heterozigotos têm atividade enzimática reduzida, mas detectável (CRABB *et al.*, 2004).

3.2.2 Sequenciamento dos Fragmentos Referentes ao Exon 4 do Gene RALDH2 dos Pais dos Pacientes Afetados pelas Mutações G151T e T157C

De modo a verificar se as mutações G151T e T157C teriam sido herdadas pelos pacientes ou se seriam mutações *de novo* (não transmitidas pelos pais), nós sequenciamos o fragmento referente ao exon 4 do gene RALDH2 dos pais dos pacientes portadores destas mutações. Os resultados estão ilustrados na figura abaixo (figura 7):

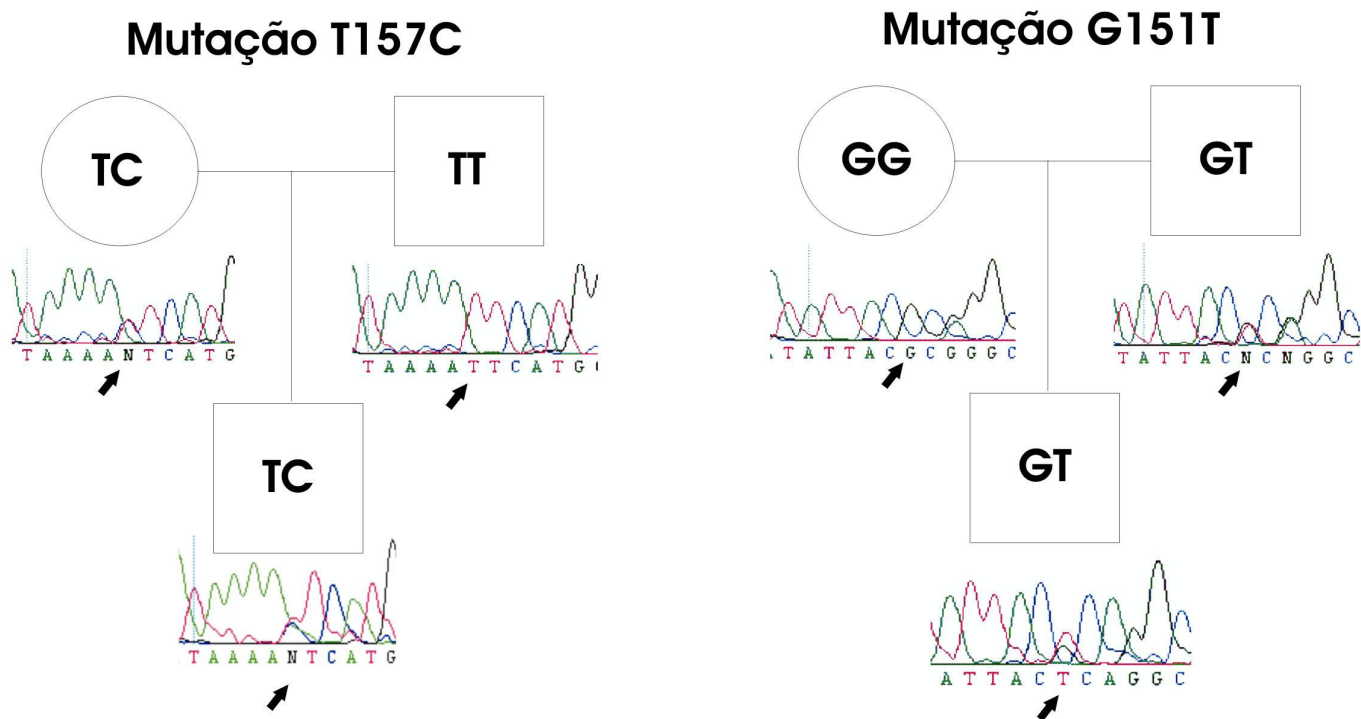


Figura 7 - Hereditogramas mostrando o padrão de herdabilidade das mutações G151T e T157C. À esquerda: o hereditograma mostra que a mãe do paciente portador da mutação T157C transmitiu-a para o filho. À direita: o pai do paciente da mutação G151T também é portador da mutação e transmitiu-a para o filho.

Como observado no hereditograma à esquerda, a mãe do paciente portador da mutação T157C também possui a mutação e a transmitiu para o filho; o pai do paciente não possui a mutação. No hereditograma à direita nós podemos concluir que a mutação G151T foi transmitida do pai para o filho, sendo que a mãe possui genótipo normal.

3.2.3 Outras Mutações Encontradas no Sequenciamento do Gene RALDH2

Na região 5'UTR nós encontramos uma mutação já descrita na base de dados pública dbSNP (rs34645259), de troca G por A, próxima 6 pb ao sítio de iniciação (figura 8). Esta mutação foi encontrada em 4 pacientes, de 55 analisados.

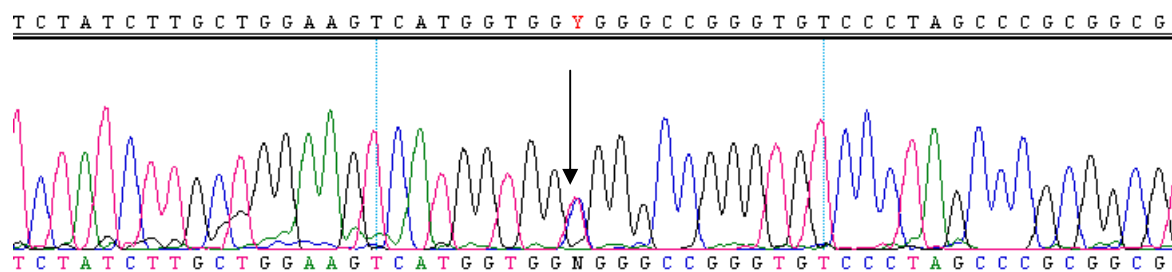


Figura 8 – Mutação de troca G por A (rs34645259) localizada na região 5'UTR do gene RALDH2. Essa mutação está localizada 6pb antes do sítio ATG e foi encontrada em aproximadamente 7% dos pacientes analisados.

No exon 9 (domínio catalítico da enzima) nós encontramos uma mutação já descrita (rs4646626), não - sinônima de troca G por A, e que é responsável pela troca do aminoácido valina pelo aminoácido isoleucina (figura 9). Bastante freqüente, de 45 pacientes analisados, 30 apresentaram essa mutação (aproximadamente 67% dos pacientes).

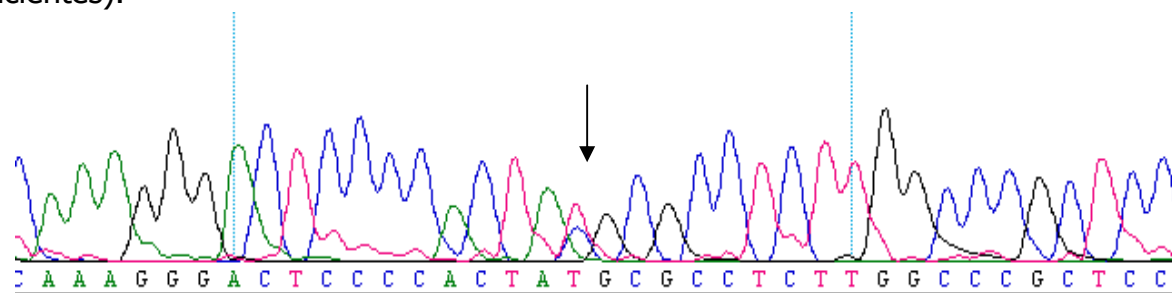


Figura 9: Mutação de troca G por A (rs4646626), não sinônima, encontrada no exon 9 do gene RALDH2 em aproximadamente 67% dos pacientes analisados. Essa mutação troca um aminoácido valina por um aminoácido isoleucina e está localizada no domínio catalítico da enzima.

No exon 11 nós encontramos uma mutação sinônima de troca C por A no primeiro codon do exon, já descrita (rs35251510), e que codifica o aminoácido isoleucina (figura 10). Esta mutação foi encontrada em 2 pacientes, de 55 analisados.

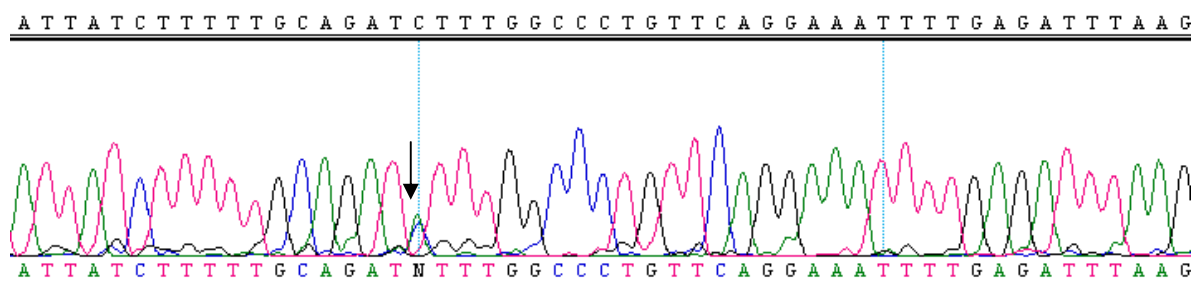


Figura 10 – Mutação de troca C por A, sinônima (aminoácido isoleucina), encontrada no exon 11 do gene RALDH2. Essa mutação foi encontrada em aproximadamente 3.5% dos pacientes analisados.

3.2.4 Análise do Dobramento do RNAm do Gene RALDH2

Nós utilizamos o programa ESEfinder (<http://exon.cshl.edu/ESE/>) que executa procuras por ESEs putativos nas seqüências requeridas utilizando matrizes de peso correspondentes às seqüências aos quais se ligam as quatro proteínas humanas do tipo SR (*serine-rich*), que são as proteínas que se ligam aos sítios de ESE. As matrizes são baseadas nos valores das freqüências derivadas do alinhamento das melhores seqüências obtidas por experimentos funcionais do tipo SELEX (*Systematic Evolution of Ligands by Exponential enrichment*) (CARTEGNI *et al.*, 2003). Nós utilizamos esse programa de bioinformática para prever se a localização do polimorfismo A151G e das mutações G151T e T157C coincidiam com sítios de ESEs. O resultado segue na figura abaixo (figura 11):

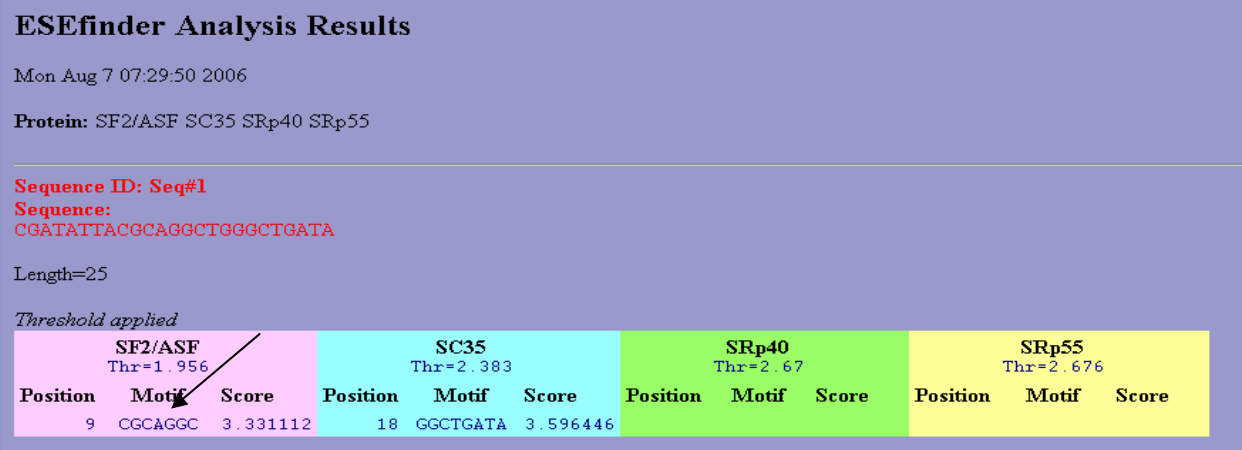


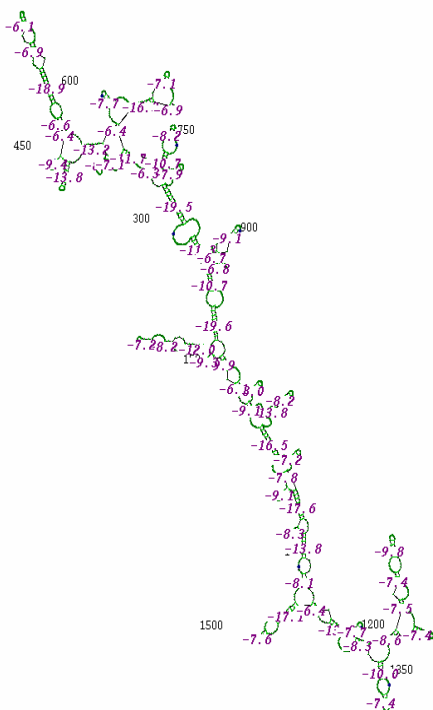
Figura 11 - Análise de *exonic splicing enhancers* através do programa ESEfinder para a sequência referente ao exon 4 do gene RALDH2. A seta indica que o polimorfismo A151G está inserido em um sítio de ESE em um local de ligação para proteínas SF2 / ASF, pertencentes à família de proteínas do tipo SR.

FONTE: <http://exon.cshl.edu/ESE/>

A análise da sequência do exon 4 do gene RALDH2 mostra que, exatamente onde se encontram o polimorfismo A151G e a mutação G151T, existe um possível sítio de ESE, onde se ligam as proteínas – do tipo SR (*serine-rich*) – SF2 / ASF. Já a mutação T157C, segundo a predição do programa, não está localizada em nenhum destes sítios.

Uma vez que mudanças sinônimas (mudanças que não trocam o aminoácido) podem afetar potencialmente a função por interferir na estrutura secundária do RNAm, nós investigamos se a transição A151G estaria relacionada a mudanças significativas neste parâmetro. A análise do dobramento do RNA indicou que a variante A151G foi associada com uma importante alteração estrutural refletida em grandes variações das energias livres calculadas (figura 12).

Free Energy of Structure = -294.1 kkal/mol



Free Energy of Structure = -333.2 kkal/mol

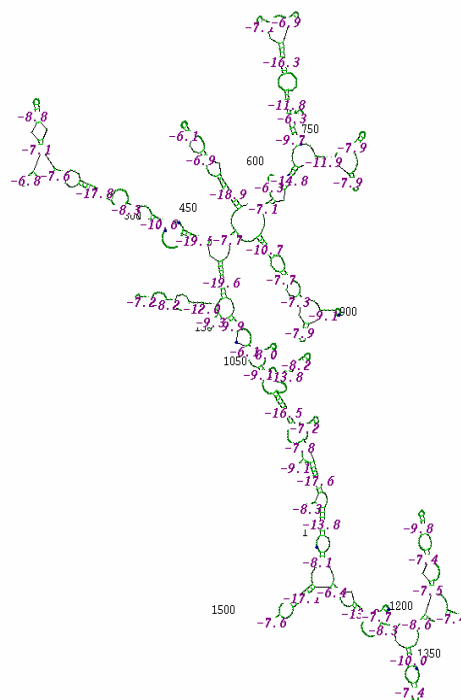


Figura 12: Análise da estrutura secundária do RNA do gene *RALDH2* sem o polimorfismo A151G (esquerda) e com o polimorfismo A151G (direita). Na presença do polimorfismo A151G observa-se grande variação na estrutura do RNAm, refletida na grande variação da energia livre da estrutura.

Para estabelecer o quão específicas eram as mudanças estruturais induzidas pelo polimorfismo A151G no RNAm, nós examinamos as conseqüências de transições simuladas no primeiro (G to A) e no segundo nucleotídeo (C to T) do cistron GCA que codifica para o aminoácido alanina, na posição 151. Subseqüentemente, nós avaliamos as mudanças induzidas pela replicação da mesma alteração A por G em todos os outros cístrons GCA, que codificam para alanina do gene *RALDH2*. Estes controles foram completados através da simulação das alterações estruturais induzidas por outros polimorfismos sinônimos descritos no gene *RALDH2*. Esta análise indicou que nenhuma das transições na primeira ou na segunda posição do cistron GCA, codificando alanina, na posição 151, nem as transições A por G no terceiro nucleotídeo de todos os outros

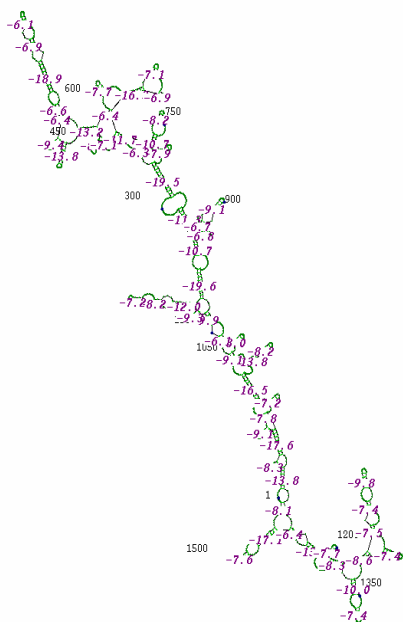
cístrons GCA foram tão importantes quanto os efeitos na estrutura secundária do RNAm causados pela transição A151G (quadro 4). Da mesma forma, as outras substituições sinônimas descritas para o gene *RALDH2* falharam em induzir mudanças importantes na estrutura do RNAm.

Quadro 4 - Simulações de transições no primeiro e no segundo nucleotídeo do cistron GCA na posição 151 e da alteração A por G em todos os outros cistrons GCA, que codificam o aminoácido alanina.

Descrição	Localização	Energia
RNA RALDH2 selvagem		- 294,1 kkal / mol
RNA mutado troca A por G	A151G	- 333,2 kkal / mol
Transição de G por A no primeiro nucleotídeo do aminoácido	A151G	- 292,8 kkal / mol
Transição de C por T no segundo nucleotídeo do aminoácido	A151G	- 288,8 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A72G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A75G	- 296,5 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A80G	- 309,6 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A83G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A110G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A118G	- 290,8 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A121G	- 295,0 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A211G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A218G	- 302,2 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A248G	- 330,8 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A249G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A260G	- 295,5 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A274G	- 299,2 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A322G	- 293,4 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A447G	- 297,2 kkal / mol
Troca A por G no terceiro nucleotídeo do códon GCA	A463G	- 293,4 kkal / mol

A análise do dobramento do RNAm, quando a mutação G151T está presente, também mostrou – embora de forma mais discreta – alteração na estrutura deste RNA (figura 13):

Free Energy of Structure = -293.4 kkal/mol



Free Energy of Structure = -292.2 kkal/mol

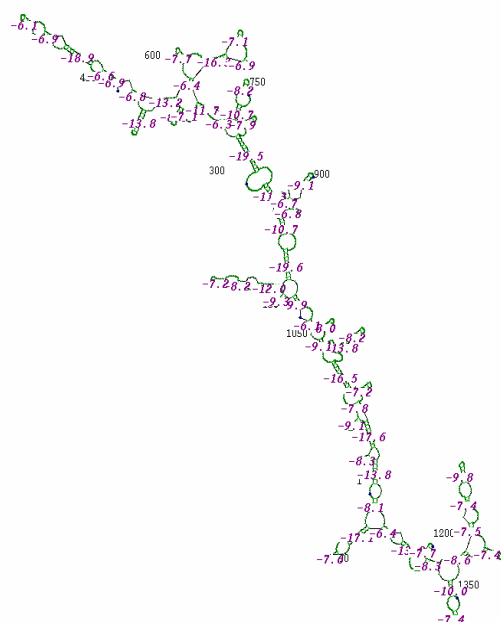


Figura 13: Estrutura do RNAm do gene RALDH2 normal (à esquerda) e em presença da mutação G151T (à direita). Observa-se, quando a mutação G151T está presente, mudança na estrutura do RNAm e na energia livre da estrutura.

3.2.5 Estudos de *Splicing* para o Polimorfismo A151G

Uma vez que o polimorfismo A151G está localizado em um provável ESE (*exonic splicing enhancer*), nós desenvolvemos um ensaio para determinar os efeitos funcionais do polimorfismo A151G no *splicing*.

A estratégia que nós desenvolvemos para a identificação de seqüências transcritas de DNA genômico clonados foi uma estratégia baseada em amplificação exônica. Esta metodologia se utiliza do plasmídeo pSPL3 que contém as seqüências codificantes do gene da β -globina de coelho e estas seqüências codificantes são separadas por uma porção do gene HIV-*tat* (BURN *et al.*, 1995). O DNA genômico é

então clonado no intron do gene *tat* e os subclones resultantes são transfectados preferencialmente em células COS-7, uma linhagem de células de macaco que, pela manipulação artificial, levou à integração de um segmento do genoma de SV40 que contém uma origem de replicação SV40 defectiva. A sequência de SV40 do vetor, integrado nas células COS-7, permite ao vetor contendo uma origem de replicação do SV40 funcional, replicar independentemente do DNA celular (STRACHAN, 2002). Os exons que estão subclonados e que sofrem *splicing* junto ao transcrito de globina – *tat* são recuperados por RT-PCR. Como a expressão do DNA genômico que está subclonado é direcionado por um promotor viral, as características espaciais e temporais de expressão não são um problema (BURN *et al.*, 1995).

Este ensaio utiliza-se de um vetor específico para o estudo de *splicing*, o vetor pSPL3 (CHURCH *et al.*, 1994). Como dito anteriormente, este vetor é formado por dois exons curtos separados por um longo intron. O sítio múltiplo de clonagem do vetor está localizado entre o exons, no intron (figura 14). Quando o vetor pSPL3 é transfectado na célula, a maquinaria de *splicing* remove o intron original, restando somente os dois exons originais (1 e 2). Se um fragmento experimental – contendo um exon flanqueado por sequências intrônicas – é inserido neste vetor e essa construção é transfectada, dois cenários são possíveis: quando os sinais de *splicing* são fortes todas as sequências intrônicas são adequadamente removidas restando, além dos exons 1 e 2 originais do vetor, o exon experimental; quando os sinais de *splicing* são muito fracos, além da remoção dos introns originais ocorre também a exclusão do fragmento inserido.

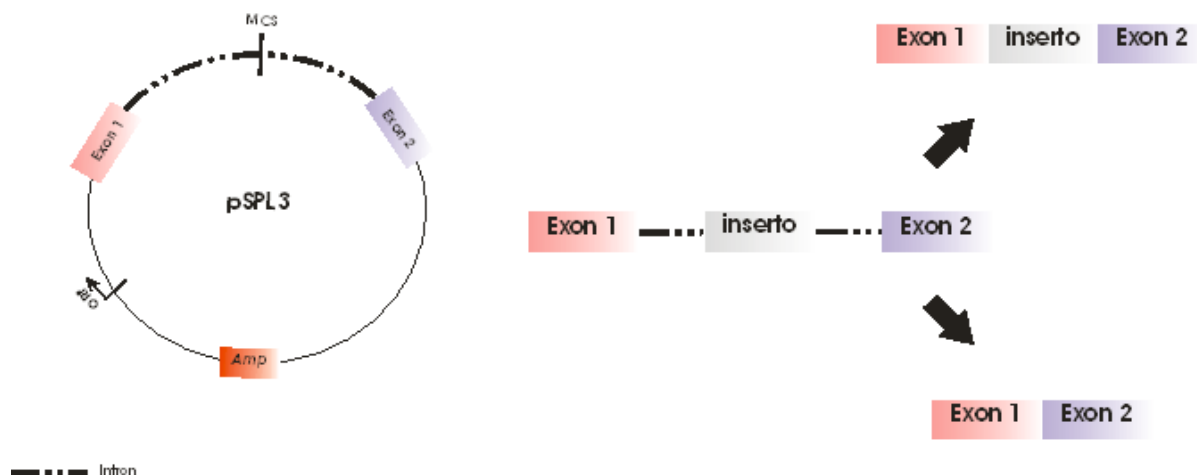


Figura 14: Estudo de *splicing* empregando o vetor pSPL3. Este vetor contém dois exons separados por um longo intron onde está inserido o sítio múltiplo de clonagem. O inserto experimental (o exon flanqueado por curtos fragmentos de suas seqüências intrônicas) é subclonado neste vetor e transfectado em células COS-7. A maquinaria de *splicing* da célula pode: (I) remover os introns e reter os exons, inclusive o experimental, ou (II) se os sinais de *splicing* forem muito fracos, remover inclusive o exon experimental.

Após a transfecção com as construções (pSPL3 + exon 4 com o polimorfismo A151G; pSPL3 + exon 4 sem o polimorfismo; apenas o vetor pSPL3; apenas lipofectamina - controle) nós amplificamos os fragmentos de cDNA, sintetizados a partir do RNA extraído das células por RT-PCR, com diferentes ciclos de PCR (figura 15). Nós comparamos, utilizando o programa de análise de imagem *image pro plus*, as taxas de exclusão do exon 4 entre os fragmentos que contêm e aqueles que não contêm o polimorfismo (quadro 5). Uma diferença de aproximadamente 9% foi verificada, isto é, foi observada 9% a mais de exclusão do exon 4 quando a variante A151G está presente (tabela 1).

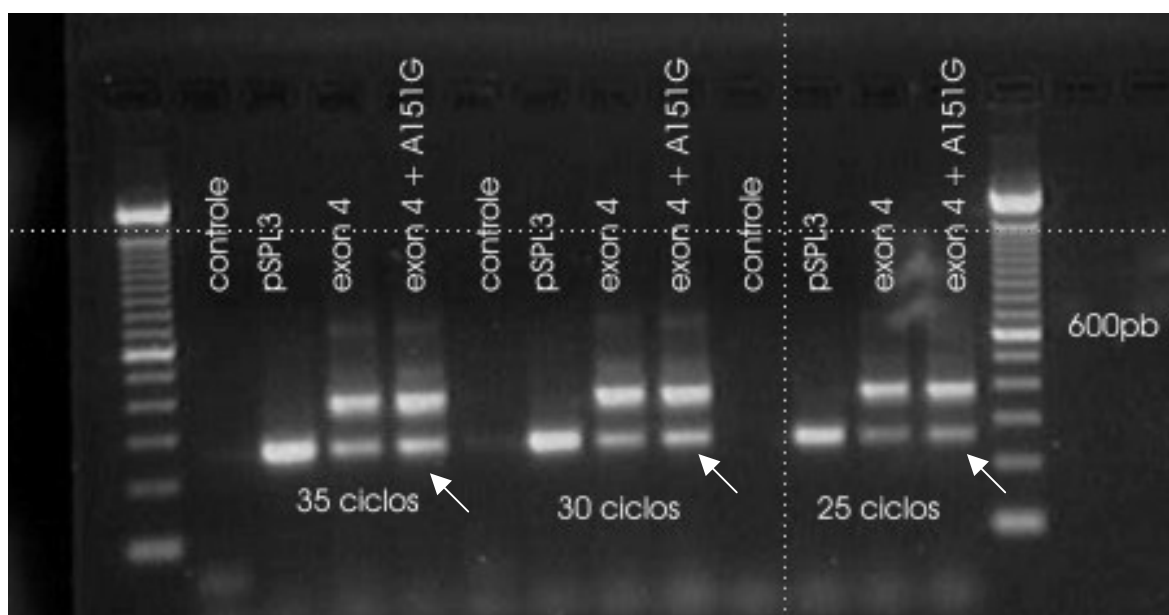


Figura 15: resultado dos diferentes ciclos de amplificação por PCR do cDNA sintetizado a partir do RNA extraído de células COS – 7. Nota-se uma maior exclusão do exon 4 polimórfico (da ordem de aproximadamente 9%) em relação ao exon 4 selvagem.

Quadro 5 – Ensaio de *splicing* para o polimorfismo A151G: medidas de densidade das bandas através do programa Image Pro Plus.

	35 ciclos	30 ciclos	25 ciclos
Exon 4 (total)	58.891	56.879	39.3237
fragmento com a inclusão do exon	32.914	34.112	31.718
fragmento com a exclusão do exon	25.977	22.677	7.6057
Exon 4 + A151G (total)	46.804	50.645	28.5703
fragmento com a inclusão do exon	30.332	34.935	23.768
fragmento com a exclusão do exon	16.472	15.710	4.8023

Tabela 1 – Razão (em porcentagem) das taxas de inclusão e exclusão dos exons experimentais.

	35 ciclos	30 ciclos	25 ciclos
Exon 4			
fragmento com a inclusão do exon	64,8%	68,9%	83,1%
fragmento com a exclusão do exon	35,2%	31,1%	16,9%
Exon 4 + A151G			
fragmento com a inclusão do exon	55,9%	60%	80,6%
fragmento com a exclusão do exon	44,1%	40%	19,4%

3.2.6 Estudos de *Splicing* para a Mutação G151T

Uma vez que a mutação G151T também está localizada no ESE, nós empregamos a mesma metodologia para o estudo desta mutação. Como resultado nós observamos um visível aumento do sinal de *splicing*, uma vez que não foi observada exclusão do exon experimental (figura 16).

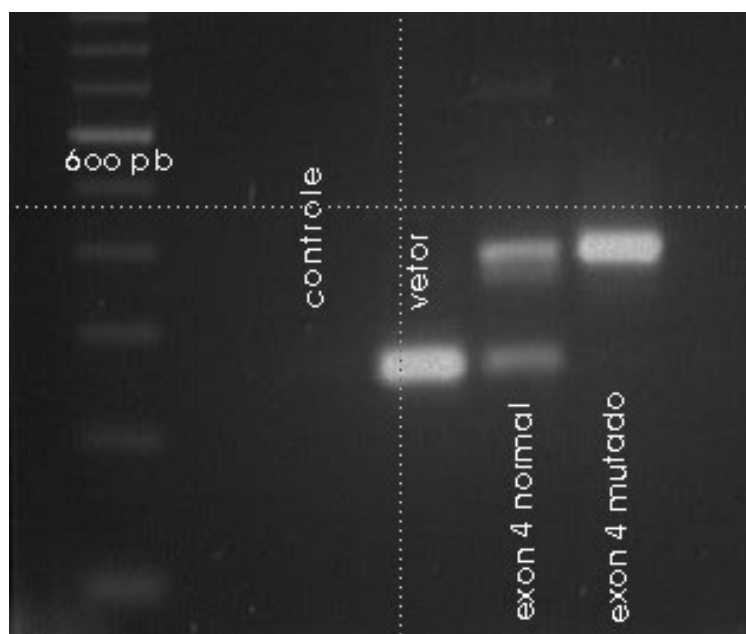


Figura 16: Ensaio de *splicing* para a mutação G151T. Observa-se, quando a construção contém o exon 4 selvagem do gene RALDH2, uma pequena taxa de exclusão já observada anteriormente, no experimento com o polimorfismo A151G. Esta exclusão, porém não é observada quando utiliza-se a construção contendo a mutação G151T.

Observa-se, quando a construção contém o exon 4 selvagem do gene RALDH2, uma pequena taxa de exclusão já observada anteriormente, no experimento com o polimorfismo A151G. Esta exclusão, porém não é observada quando utiliza-se a construção contendo a mutação G151T.

Nosso trabalho mostra, neste experimento, que a mutação pode estar aumentando a inclusão do exon 4 no transcrito total. Esse aumento pode, possivelmente: (I) estar relacionado à uma melhoria de um ESE pré existente; (II) estar relacionado à criação de um elemento de ESE (uma vez que o programa que nós utilizamos – ESEfinder – é baseado em predição, é possível que este sítio não exista *in vivo*); ou ainda (III) essa mutação pode estar inativando um sítio de ESS (*exonic splicing silencer*) e, desta forma, aumentando a taxa de inclusão do exon 4 no *splicing*. Como nossas análises de bioinformática anteriores já mostraram que essa mutação

possivelmente se encontra em um ESE (figura 11), a primeira explicação parece ser a mais plausível.

3.2.7 Conseqüências para a Proteína RALDH2 com a Exclusão do Exon 4

Caso haja a exclusão, de forma aberrante, do exon 4 devida à qualquer uma destas duas variantes, as conseqüências para a proteína seriam bastante importantes. Como se pode observar na figura a seguir (figura 17), o último aminoácido do exon 4 (asparagina) é compartilhado com o exon 5. Havendo a exclusão do exon 4, ocorre uma mudança no quadro de leitura da proteína, mudando o primeiro aminoácido do exon 5 de asparagina para metionina. Além da troca de todos os aminoácidos subsequentes, no exon 6 ocorre ainda a inserção de um códon de parada prematuro. O exon 3 não foi esquematizado pois, uma vez que não compartilha aminoácido com o exon 4, não ocorre qualquer mudança em seu quadro de leitura.

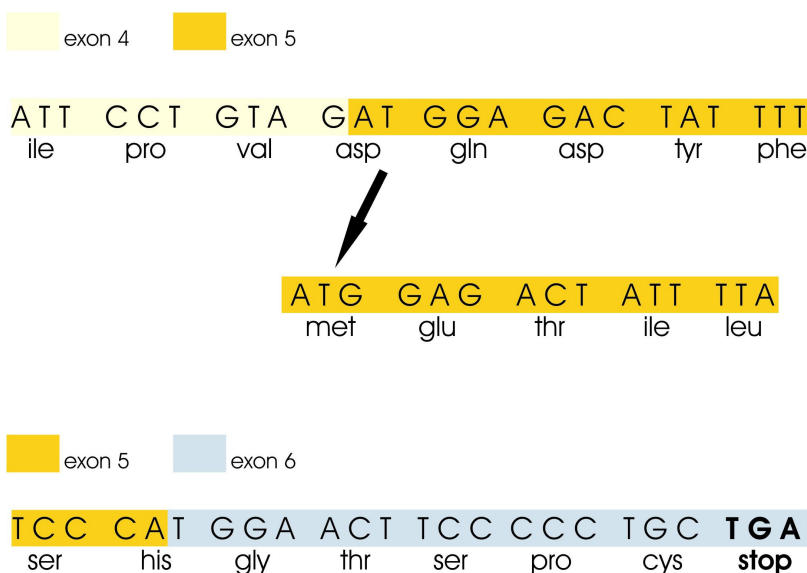


Figura 17 – Simulação do erro no quadro de leitura da proteína RALDH2 com a possível exclusão do exon 4. Se houver a exclusão do exon 4, muda-se o quadro de leitura a partir do início do exon 5 e, como conseqüência, gera-se um códon de parada no exon 6.

3.3 Estudo de Associação

Na etapa seguinte do nosso trabalho, nós investigamos a hipótese de que polimorfismos no gene *RALDH2* estariam associados a doenças cardíacas congênitas. Para isso nós utilizamos SNPs (*single nucleotide polymorphisms*) disponíveis em base de dados públicas (rs2704188; rs1441815; rs1530293; rs1899430; rs3784259) (<http://www.ncbi.nlm.nih.gov/projects/SNP/>) e um polimorfismo anteriormente associado à espinha bífida por Deak e colaboradores (A151G) (DEAK *et al.*, 2005).

3.3.1 Análise da Estrutura de Desequilíbrio de Ligação

Estes marcadores polimórficos foram primeiramente utilizados para determinar a estrutura de desequilíbrio de ligação (LD) entre eles.

A figura 18 representa a estrutura de desequilíbrio de ligação para os marcadores selecionados no gene *RALDH2*, testados nos 101 trios de pacientes / pais. O polimorfismo A151G não foi incluído em virtude de sua baixa frequência na população.

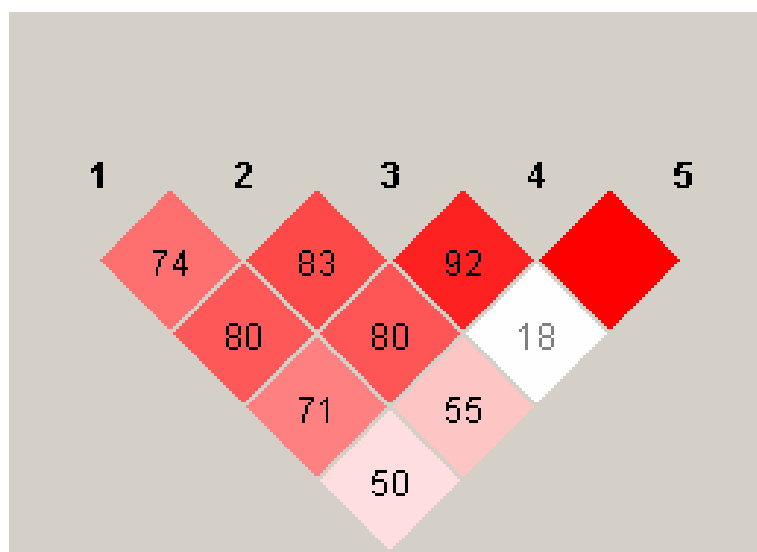


Figura 18 - Desequilíbrio de ligação entre os marcadores: 1. rs2704188; 2. rs1441815; 3. rs3784259; 4. rs1530293; 5. rs1899430.

Como mostrado, uma forte relação entre os marcadores analisados foi demonstrada através da alta pontuação obtida na comparação par a par (sendo o mais alto grau conseguido entre 3 e 4, com 92% de desequilíbrio de ligação), e dois blocos de desequilíbrio de ligação na região abrangida puderam ser definidos, um entre os marcadores de 1 a 4 e outro entre os marcadores 4 e 5.

3.3.2 Teste de Desequilíbrio de Transmissão (TDT)

Finalmente, os marcadores analisados, baseados no desequilíbrio de ligação, frequência e distribuição ao longo do gene RALDH2, foram incluídos em um estudo de associação realizado com 101 trios (pacientes com DCCs diversas e seus pais). Este estudo é chamado TDT (*transmission disequilibrium test*) e é um teste de associação baseado em família que avalia se a frequência de transmissão dos alelos dos pais heterozigotos da criança afetada desvia dos 50% esperados para a frequência Mendeliana (SCHAID, 1998).

Para a análise do teste de desequilíbrio de transmissão (TDT) nós utilizamos o algoritmo Haploview. Controles de qualidade baseados na repetição da genotipagem de alguns indivíduos foram utilizados para excluir possíveis erros de genotipagem.

Como mostrado na tabela 2, nenhum dos marcadores analisados mostrou associação significativa com DCCs quando testado com o teste TDT. Esta conclusão pode ser observada na coluna T:U (*transmitted : untransmitted*). Os alelos polimórficos foram distribuídos em taxas bastante equivalentes para todos os marcadores. Essa ausência de associação mostrou-se mesmo quando os marcadores foram testados sob um estudo de associação padrão utilizando-se os genótipos dos pais como controle (dado não mostrado).

Tabela 2 – Resumo dos resultados das freqüências de transmissão dos alelos selecionados.

SNP	Localização	Alelos	T:U	Chi Square	Valor de p
1. rs2704188	região 5´	C/T	42:38	0.2	0.6547
2. rs1441815	intron 7	G/T	43:36	0.62	0.431
3. rs3784259	intron 8	A/G	44:43	0.011	0.9146
4. rs1530293	região 3´	A/G	36:33	0.13	0.718
5. rs1899430	região 3´	C/T	29:28	0.018	0.8946

3.3.3 Análise da Estrutura Haplotípica

A estrutura haplotípica foi então determinada e os cinco blocos de haplótipos mais freqüentes estão resumidos na tabela 3. Entretanto, nenhuma freqüência haplotípica mostrou-se significativa uma vez que as freqüências de distribuição entre os blocos mostraram-se bastante equivalentes, não podendo, desta forma, serem consideradas associadas à doença.

Tabela 3 – Resultado do estudo de haplótipos dos cinco blocos mais freqüentes no gene RALDH2.

Bloco	Haplótipo	Freqüências	T:U	Chi Square
CTAGC	0.307	37.3 : 39.6	0.069	0.7925
TGGAC	0.251	35.8 : 29.1	0.694	0.4048
TGAGC	0.109	13.4 : 19.4	1.105	0.2931
TGGGT	0.075	12.1 : 15.3	0.375	0.5402
TTAGC	0.052	10.5 : 10.4	0.0010	0.9804

3.3.4 Estudo de Associação Caso – Controle para a Variante A151G

Para a variante A151G nós desenvolvemos um estudo de associação caso – controle, uma vez que esta variante não foi incluída no teste TDT por estar em baixa freqüência na população, inferior a 1%, não sendo, desta forma, considerada um SNP. Esta mesma alteração foi descrita por Deak e colaboradores (DEAK *et al.*, 2005) como um polimorfismo associado a espinha bífida. Para determinar a freqüência estatística do

polimorfismo A151G, nós sequenciamos 201 pacientes com tetralogia de Fallot e um grupo controle composto de 44 indivíduos normais. A tetralogia de Fallot foi a cardiopatia congênita utilizada neste teste por se apresentar em maior quantidade em nosso ambulatório. O polimorfismo A151G esteve presente em 7 pacientes, representando assim 3.5% do total. No grupo controle nós encontramos a variante em 3 pessoas. Isto representa 6,8% dos indivíduos normais. Essa diferença percentual entre os grupos não pode ser considerada estatisticamente significativa.

4. DISCUSSÃO

Neste estudo nós avaliamos o potencial do gene RALDH2 como candidato em doenças cardíacas congênitas. Para isto nós utilizamos duas estratégias. Em primeiro lugar nós avaliamos a possibilidade de que mutações no gene RALDH2, *per se*, pudessem produzir defeitos cardíacos congênitos. A metodologia utilizada foi o sequenciamento de todos os exons do gene RALDH2 em 50 pacientes portadores da tetralogia de Fallot. Em segundo lugar nós investigamos a hipótese de que o gene RALDH2, embora não pudesse produzir defeitos cardíacos *per se*, pudesse proteger ou suscetibilizar indivíduos a defeitos cardíacos congênitos.

4.1 Mutações Não-Conservativas Encontradas no Sequenciamento do Gene RALDH2

Como resultado do sequenciamento nós encontramos duas mutações não-conservativas no exon 4. Uma dessas variantes é uma mutação de troca G por T, na posição 151 (G151T) e que é responsável pela troca do aminoácido alanina (apolar) pelo aminoácido serina (polar). O aminoácido alanina é um aminoácido hidrofóbico (não polar) e isto faz com que este tipo de aminoácido esteja localizado interiormente nas proteínas, longe do meio aquoso. Já o aminoácido serina possui característica polar. Por serem hidrofílicos, aminoácidos polares são encontrados geralmente na superfície de proteínas hidrossolúveis, onde não somente contribuem para a solubilidade da proteína como também formam sítios de ligação para moléculas com carga. O fato de essas duas mutações trocarem a polaridade dos aminoácidos indica que pode haver, como consequência, impacto significativo na função protéica uma vez que um fator importante que governa o dobramento de uma proteína é a distribuição de seus aminoácidos polares e não polares. Uma vez que a polaridade das cadeias laterais dos aminoácidos é uma das forças responsáveis pela forma final da estrutura tridimensional das proteínas, esses tipos de mutações podem ter grande impacto na função protéica (HARVEY, 2000).

A outra mutação é uma troca de T por C, na posição 157 (T157C) e é responsável pela troca do aminoácido isoleucina (apolar) pelo aminoácido treonina (polar). A treonina é um aminoácido desprovido de carga mas que possui grupos hidroxila polares que participam de pontes de hidrogênio com outras moléculas polares. Já a isoleucina é um aminoácido apolar cujas cadeias laterais consistem inteiramente de hidrocarbonetos.

Ambas as mutações foram encontradas em apenas um paciente cada e não haviam sido descritas anteriormente. Por causa das características mencionadas, o achado destas mutações são consistentes com a idéia de que o gene RALDH2, *per se*, pode promover doença cardíaca congênita.

4.1.1 Estudo de Modelagem Molecular para as Mutações G151T e T157C

Para avaliar possíveis mudanças estruturais que as mutações G151T e T157C poderiam causar na proteína, nós realizamos um estudo de modelagem molecular. Assim, foi possível caracterizar a estrutura protéica do gene RALDH2 e localizar os aminoácidos mutados G151T e T157C. Nós observamos, em ambos os casos, que ocorre mudança estrutural no local onde as mutações estão inseridas. O exon 4 é um dos domínios de tetramerização da proteína – juntamente com o exon 13 – e ambos os subdomínios estão bastante próximos entre si na estrutura da proteína. Trabalhos recentes mostram que mutações nestes sítios podem, aparentemente, diminuir ou até impedir a tetramerização da proteína (PARK *et al.*, 2002; SUMANDEA & L, 2005). Outros trabalhos têm associado mutações no exon 13 da enzima ALDH2 com deficiência no metabolismo de etanol na população asiática (CRABB *et al.*, 2004). De posse destas informações é possível que variantes em um dos sub-domínios de tetramerização possam eventualmente afetar a tetramerização como um todo, levando a distúrbios associados a qualquer um desses dois sub-domínios protéicos (exons 4 e 13).

4.1.2 Sequenciamento dos Fragmentos Referentes ao Exon 4 do Gene RALDH2 dos Pais dos Pacientes Afetados pelas Mutações G151T e T157C

Nós sequenciamos também os fragmentos relativos ao exon 4 dos pais dos pacientes afetados pelas mutações G151T e T157C. Mutações de novo (não herdadas) são um evento relativamente raro e, como era de se esperar, o que nós observamos é que essas mutações foram herdadas pelos pacientes. O paciente que possui a mutação G151T herdou-a do pai e o paciente que possui a mutação T157C herdou-a da mãe. Mais estudos envolvendo outros membros destas duas famílias podem ser conduzidos para avaliar se esta mutação está definitivamente associada à cardiopatia congênita familiar.

4.2 Variante Conservativa Encontrada no Sequenciamento do Gene RALDH2

Outra mutação encontrada no gene RALDH2 foi uma mutação já descrita (rs4646626) no domínio catalítico da enzima (exon 9). Esta mutação de troca G por A troca um aminoácido valina (apolar) por um aminoácido isoleucina (apolar). Esta variante se mostrou bastante freqüente uma vez que nós a encontramos em 67% dos pacientes analisados. Esta variante pode ser classificada como um polimorfismo, uma vez que é também encontrada em indivíduos normais. Mutações em domínios catalíticos podem levar à diminuição ou completa abolição da função enzimática e mutações nestes sítios já têm sido relacionadas ao câncer (MILED *et al.*, 2007) e à erros no processo de metilação do DNA (GOWHER *et al.*, 2006).

4.3 Variantes Sinônimas Encontradas no Sequenciamento do Gene RALDH2

No exon 11 nós encontramos um polimorfismo já descrito (rs35251510) de troca C por A no primeiro códon do exon e que codifica o aminoácido isoleucina. No nosso grupo de pacientes essa alteração se mostrou menos freqüente, estando presente em 3.5% dos indivíduos afetados.

A terceira variante encontrada no exon 4 foi um polimorfismo sinônimo de troca A por G, na posição 151 (A151G) que codifica o aminoácido alanina. Este polimorfismo se mostrou bastante interessante pois em um trabalho publicado em 2005 por Deak e colaboradores (DEAK *et al.*, 2005) esta variante foi associada à espinha bífida. Nós observamos que a variante A151G produz mudanças substanciais na conformação do RNA do gene RALDH2 (variações na estrutura secundária do RNAm podem interferir na ligação de proteína importantes para o processo de *splicing*) (figura 12) e, além disso, está localizada em um ESE (*exonic splicing enhancer*), em um sítio de ligação para a proteína ASF / SF2, pertencentes à família de proteínas do tipo SR (figura 11). Por todas essas evidências nós consideramos uma investigação mais profunda à respeito do impacto desta variante para a função da enzima RALDH2.

4.4 Conseqüências para a Proteína RALDH2 com a Exclusão do Exon 4 pela Maquinaria de *Splicing*

Como o polimorfismo A151G está localizado em um ESE (bem como a mutação G151T), nós analisamos quais seriam as conseqüências para a estrutura da proteína se o gene RALDH2 tivesse seu exon 4 excisado em conseqüência de um erro de leitura da maquinaria de *splicing*. Essa exclusão levaria a um completo erro de leitura a partir do primeiro aminoácido do exon 5, com conseqüente geração de um códon de terminação no exon 6 (figura 17). Em doenças cardíacas congênitas são vários os exemplos de mutações que dão origem à erros no quadro de leitura da proteína e que geram códons de terminação prematuros. Mutações que dão origem à codons de terminação prematuros, por mudança no quadro de leitura da proteína, já foram bem descritas para as doenças cardíacas congênitas. Uma mutação no gene JAG1 (c.1880_1881insA), que provoca erro no quadro de leitura, revelou ser responsável por grande variabilidade no fenótipo de pacientes que são portadores da síndrome de Allagile. A síndrome de Allagile é uma desordem autossômica dominante que, entre outros órgãos, atinge também o coração (HARENDZA *et al.*, 2005). Em outro trabalho também envolvendo o gene JAG1, Jurkiewicz e colaboradores (JURKIEWICZ *et al.*, 2005) anotaram 16 novas

mutações neste gene sendo que 7 destas mutações provocam mudança no quadro de leitura. Além disso, foi observado que 75% destas mutações inserem um codon de parada prematuro na proteína. Também já foi descrita uma mutação no gene NOTCH1 que resulta em erro de leitura seguido pela geração de um codon de parada prematuro. As geração dessa proteína truncada leva à um espectro de anomalias no desenvolvimento da valva aórtica. Ainda, a geração desses transcritos truncados proporciona evidência que haploinsuficiência de NOTCH1 resulta em doença cardíaca (GARG *et al.*, 2005). Estudos em famílias portadoras de defeitos de septação cardíaca isolados identificaram uma mutação transcrionalmente inativa que causa erro de leitura no gene GATA4 (E359del) e segrega com defeitos de septação cardíaca (GARG *et al.*, 2003)

4.5 Ensaio de *splicing* para o polimorfismo A151G

Para avaliar a possibilidade de que o polimorfismo A151G pudesse alterar o padrão normal de *splicing*, nós desenvolvemos um estudo funcional. A análise dos resultados mostrou que uma diferença de aproximadamente 9% foi verificada quando comparamos a taxa de exclusão do exon 4 entre as seqüências polimórficas e não-polimórficas. Entretanto, as conseqüências *in vivo* dessas diferenças são ainda desconhecidas.

4.5.1 *Splicing* Alternativo e Doenças Humanas

Um gene típico, em vertebrados superiores, é composto de vários exons curtos separados por longas seqüências intrônicas. Para gerar moléculas de RNAm funcionais, estes introns precisam ser excisados e os exons, reunidos, da forma mais precisa e eficiente possível. A produção de várias isoformas destes transcritos pelos vários tipos de *splicing* alternativos é um evento relativamente comum nas células de mamíferos. Estes eventos são responsáveis por muita da complexidade do proteoma e explica a descoberta inesperada de que o genoma humano consiste de apenas 31.000 a 39.000

genes. Ainda, análises genômicas de *splicing* alternativos indicam que de 40% a 60% dos genes têm formas alternativas geradas por *splicing*. A necessidade de se regular um evento tão complexo e que gera tanta variabilidade introduz um requerimento extra de sinais que possam modular o *splicing* de forma específica tanto do ponto de vista do desenvolvimento quanto da perspectiva celular. Além destes sinais, elementos *cis* regulatórios estão presentes em seqüências codificantes de genes e que permitem à maquinaria de *splicing* modular a seleção de sítios de *splicing* alternativos (CARTEGNI *et al.*, 2002; MODREK & LEE, 2002). Essas seqüências exônicas podem desempenhar função importante tanto no *splicing* constitutivo quanto no *splicing* alternativo. Estes elementos podem agir de forma a estimular (como *enhancers*) ou reprimir (*silencers*). *Exonic splicing enhancers* (ESEs) são seqüências discretas junto aos exons e que servem como sítios de ligação para proteínas específicas do tipo *serine-rich* (SR) (BLENCOWE, 2000; CARTEGNI *et al.*, 2002). A descoberta de um número cada vez maior de ESEs começa a chamar a atenção para a possibilidade de que estes elementos possam ser alvos de mutações que possam estar envolvidas em doenças humanas de base genética. De fato, Xu *et al.*, em 2005, demonstraram que camundongos *knockout* coração-específicos de uma proteína do tipo SR – ASF / SF2 – produzia cardiomiopatia por perda da regulação de eventos específicos de *splicing* alternativos durante o desenvolvimento pós natal precoce (XU *et al.*, 2005). Além disso, estudos em linhagens celulares de galinha mostram ser este um gene essencial, uma vez que a depleção de ASF / SF2 faz com que o pré-RNAm processado de forma incompleta se acumule, causando morte celular (WANG *et al.*, 1996). Outro dado importante relacionado às doenças genéticas humanas é que aproximadamente 15% das mutações pontuais que estão ligadas à doenças humanas de base genética, causam defeitos no *splicing* (KRAWCZAK *et al.*, 2007). Tanto é assim que para diversas doenças já existem estudos mostrando o impacto de polimorfismos que alteram a transcrição e o *splicing* de genes importantes (AZNAREZ *et al.*, 2003; KWOK *et al.*, 2005; CARTEGNI *et al.*, 2006). Um dos exemplos mais representativos é o dos estudos que envolvem a atrofia espinhal muscular (SMA). Essa desordem é devida à perda de ambas as cópias do gene SMN1. SMN2 é um parálogo praticamente idêntico ao gene SMN1 e está presente tanto em

indivíduos normais quanto em indivíduos afetados pela doença. No entanto, SMN2 é parcialmente funcional. Uma mutação silenciosa de base única, de troca C por T resulta na inclusão precária do exon 7 no RNAm de SMN2. Dessa forma, sem o exon 7 o resultado é uma proteína aberrante, aparentemente inativa, que não consegue complementar o defeito de SMN1, dando origem ao fenótipo de SMA. Resumindo, o exon 7 do gene SMN2 conhecidamente contém um sítio ESE que é reconhecido por proteínas do tipo SR. Este sítio é abolido por uma mudança de base única no gene SMN2, que inibe a inclusão do exon 7 (CARTEGNI & KRAINER, 2002).

Outros estudos têm demonstrado que alterações sinônimas podem causar fenótipos diretamente, agindo em diferentes níveis do processo de *splicing* do RNAm: inativando sítios de *splicing* (estima-se que esse evento represente 5.6% de todas as mutações pontuais que ocorrem nos exons do gene humano P53; inativando *exonic splicing enhancers (ESE)*, inativando / criando *exonic splicing silencers (ESS)* ou induzindo mudanças estruturais no RNA que modificam a interação entre os transcritos imaturos e os fatores de *splicing*, levando à inclusão ou excisão de exons. O processo de *splicing* é um passo crítico na expressão gênica. A seleção ou exclusão de um exon em uma molécula de RNAm permite à um único gene gerar múltiplas proteínas com diferentes estruturas primárias. Nestes casos, o processo é chamado *splicing* alternativo. O processo de *splicing*, seja ele constitutivo ou alternativo, é catalisado por uma maquinaria extremamente complexa chamada spliceossomo que, estima-se, possui aproximadamente 145 proteínas diferentes em seu complexo. Entretanto, para que a ação de diversas proteínas diferentes ocorra de forma correta é necessário o posicionamento do spliceossomo nos sítios de *splicing*. Entre os fatores que podem influenciar na identificação exon – intron pelo spliceossomo estão a presença de *silencers (exonic or intronic splicing silencers)* e *enhancers (exonic or intronic splicing enhancers)*.

SNPs (*single nucleotide polymorphisms*) podem dar origem a duas ou mais formas alélicas de RNAm. Essas variantes podem possuir funções biológicas distintas como

resultado de diferenças em sua estrutura que interage com outros componentes celulares. Alguns trabalhos sugerem que estes SNPs podem dar origem à diferentes dobramentos do RNAm. Em 1999, Shen et al. (SHEN *et al.*, 1999) mostraram que substituições aparentemente benignas podiam induzir dobramentos estruturais diversos na estrutura de pré – RNAm tanto *in vivo* quanto *in vitro*. De fato, esses estudos mostram a contribuição de variações genéticas comuns à diversidade estrutural do RNAm e sugerem um papel mais amplo do que previamente se pensava a respeito dos efeitos de SNPs na estrutura do RNAm e, conseqüentemente, na função biológica. Buratti e colaboradores (BURATTI *et al.*, 2004) mostraram, através de padrões de imunoprecipitação, a presença de interações bastante complexas entre proteínas do tipo SR e o RNA que são perdidas com variações na estrutura secundária, que ocorrem depois da introdução de variações em sítios ESE e ESS. Para as regiões de *exonic splicing enhancers* (ESEs) já existem programas computacionais que podem prever essas regiões.

4.6 Ensaio de *splicing* para a mutação G151T

Como a mutação G151T também está localizada em um sítio putativo de ESE, nós aplicamos o mesmo tipo de metodologia para essa mutação. O que nós observamos foi um aumento no sinal de *splicing*, uma vez que não foi observada exclusão do exon experimental portador da mutação G151T. Esse aumento pode estar relacionado à uma melhoria no sítio ESE ou mesmo à criação de um sítio ESE, uma vez que o método que nós utilizamos para encontrar possíveis sítios de ESE é um método baseado em predição.

4.6.1 Mutações em Sítios de *Splicing* que Podem Aumentar a Especificidade pelo Sítio ESE

De maneira bastante parecida, em um trabalho de 1999 sobre mutações em genes relacionados à demência que afetam o *splicing* alternativo do RNAm, D'Souza et

al. mostraram em um ensaio de *splicing* que quando a sequência referente ao exon 10 do gene tau era testada, tanto transcritos contendo o exon 10 eram produzidos quanto transcritos que não possuíam esse exon. Eles observaram também que quando uma mutação – a N279K – ou mesmo um polimorfismo silencioso – L284L – estavam presentes, observava-se um aumento na inclusão deste exon 10, não sendo praticamente mais detectados níveis de transcritos que não possuíam este exon (D'SOUZA *et al.*, 1999). Outro trabalho, realizado por Moseley et al. mostrou que uma transição de A por G no exon 3 do gene GH1, encontrada em famílias afetadas pela desordem IGH II, interrompe um ESE. Essa mutação também causa perda parcial ou total do exon 3 (MOSELEY *et al.*, 2002). Neste trabalhos os autores chegam à conclusão de que essa transição leva à IGH II por perturbar um sítio de ESE, necessário para o *splicing* do gene GH1. O fato de mutações localizadas em ESEs causarem uma quantidade limitada e variável de exclusão exônica sugere que essas mutações podem não ser funcionalmente nulas o tempo todo e sim podem estar representando alelos com penetrância limitada.

4.7 Mutação Localizada na Região 5'UTR do Gene RALDH2

Durante o sequenciamento do gene RALDH2 nós também encontramos uma mutação que está localizada na região 5'UTR do gene, 6 pares de base antes do sítio ATG. Essa mutação já havia sido descrita anteriormente (rs34645259) e foi encontrada em aproximadamente 7% das pessoas afetadas. Por estar localizada próximo ao sítio de iniciação, seriam precisos estudos funcionais para determinar o impacto desta mutação.

4.8 Estudos de Associação

Como já descrito anteriormente, nesta etapa do nosso projeto nós trabalhamos com a hipótese de que variantes no gene RALDH2 pudessem conferir proteção ou susceptibilidade a doenças cardíacas congênitas.

4.8.1 Estudo de Associação do Tipo Caso-Controle para a Variante A151G

Devido à recente associação da variante A151G com espinha bífida por Deak e colaboradores (DEAK *et al.*, 2005), nós desenvolvemos um estudo baseado em casos e controles. Para avaliar a hipótese de associação com doenças cardíacas congênitas nós sequenciamos o exon 4 de 201 pacientes com tetralogia de Fallot e verificamos que o polimorfismo A151G esteve presente em 3.5% do total. No grupo controle – formado por 44 indivíduos saudáveis – a variante se apresenta em 6,8% dos indivíduos. Isto quer dizer que a prevalência desta variante entre os grupos não apresenta diferença significativa, não sendo portanto possível ser associada com a doença cardíaca congênita.

4.8.3 Análise da Estrutura de Desequilíbrio de Ligação

Para testar a hipótese de associação do gene RALDH2 com o risco de predisposição ou proteção às doenças cardíacas congênitas nós utilizamos mais marcadores do gene RALDH2, encontrados na base de dados pública.

Dentre estes marcadores nós escolhemos cinco *single nucleotide polymorphisms* (SNPs) baseados em suas frequências e localização no gene, para determinar a estrutura de desequilíbrio de ligação entre eles. O desequilíbrio de ligação (associação de uma alelo marcador particular, com uma doença, em nível populacional) é bastante útil pois proporciona a melhor expectativa de reduzir a região candidata no mapeamento doença–marcador e pode levar a uma alta resolução no mapeamento genético, principalmente porque muito do genoma está contido em segmentos de alto desequilíbrio de ligação, onde variantes estão altamente correlacionadas entre si, de forma que muitos cromossomos carregam apenas um de somente uns poucos haplótipos comuns. De fato, umas poucas centenas de milhares de SNPs bem escolhidos são suficientes para providenciar informação sobre a maioria da variação comum no genoma (HIRSCHHORN & DALY, 2005). Além disso, para serem úteis, os marcadores

testados no estudo de associação precisam ser o próprio alelo causal ou estar altamente correlacionado (em desequilíbrio de ligação) com o alelo causal.

Nós encontramos uma forte relação entre os marcadores, isto é, nós conseguimos estabelecer forte desequilíbrio de ligação entre eles (somente os dados referentes a variante A151G não estão presentes entre os resultados pois esta variante aparece em uma frequência muito baixa). A análise de desequilíbrio de ligação em um estudo de associação é importante pois se um locus possui um alelo que predispõe à doença e este alelo está em desequilíbrio de ligação com alelos próximos ao local do marcador, este fenômeno pode ser explorado para mapear um gene que possivelmente cause a doença. Entretanto, nós não encontramos evidência de associação entre estes marcadores e as doenças cardíacas congênitas. O desequilíbrio de ligação entre os marcadores nos permitiu desenvolver um estudo de haplótipo envolvendo estes marcadores. Isto é importante pois os haplótipos definem as unidades funcionais dos genes, a variação em uma população está estruturada em haplótipos, além de serem úteis para reduzir os problemas inerentes aos testes de associação. Entretanto, as frequências encontradas por nós, também não nos permitiu associar os haplótipos com as DCCs.

4.8.4 Teste de Desequilíbrio de Transmissão

Um problema frequentemente encontrado em estudos de associação diz respeito à presença de estruturas populacionais não detectadas. Estratificação populacional é a presença de múltiplos subgrupos junto à uma população e que podem diferir na prevalência de doenças. Esta estrutura populacional pode mimetizar o sinal de associação, de forma que um marcador genético apresente frequências diferentes em diferentes subgrupos, e levar à falso positivos, assim como pode-se perder os efeitos reais deste tipo de estudo (MARCHINI *et al.*, 2004). O TDT, proposto por Spielman (SPIELMAN *et al.*, 1993), é um teste robusto uma vez que não é sensível à população estratificada já que utiliza membros de uma mesma família, evitando-se assim

associações espúrias. Os marcadores polimórficos – freqüentemente mudanças sinônimas – têm sido tradicionalmente usados neste modelo de estudo pois podem carregar informações úteis em estudos populacionais em virtude de desequilíbrios de ligação putativos com variantes realmente importantes como: alterações de perda de sentido, sem sentido, ou alterações no *splicing* que contêm alto risco de causar fenótipos.

Para analisar uma possível associação entre os marcadores do gene RALDH2 e as doenças cardíacas congênitas nós utilizamos o teste de TDT (teste de desequilíbrio de transmissão). Os trios familiares foram a base deste teste que comparou o número de alelos do tipo A observados, que são transmitidos para a amostra de afetados com aqueles esperados para a transmissão Mendeliana. Um excesso de alelos A (ou B) entre os membros afetados indica que um locus de susceptibilidade à doença está em ligação e em desequilíbrio de ligação com o locus marcador. Este estudo tem também uma grande vantagem uma vez que evita-se falsos positivos que surgem quando estratificação populacional está presente.

Em nossa análise nós observamos que a freqüência de distribuição entre os alelos (selvagens e polimórficos) foi bastante equivalente (tabela 2). Dessa forma nós não podemos associar estas variantes ao risco de doenças cardíacas congênitas.

5. CONCLUSÃO

As mutações não-conservativas (G151T e T157C) encontradas no sequenciamento do gene RALDH2 são consistentes com a idéia de que o gene RALDH2, *per se*, pode promover doenças cardíacas congênitas. Demonstramos em nosso estudo que a mutação G151T está relacionada com um aumento da especificidade do *splicing*. Além disso, nós desmonstramos que essas mutações podem estar sendo transmitidas de maneira familiar, já que foram encontradas nos pais dos pacientes afetados.

Apesar de estar localizado em um ESE e de provocar grandes mudanças na estrutura do RNAm, o ensaio funcional de *splicing* para o polimorfismo A151G mostrou uma discreta diferença entre os transcritos produzidos (da ordem de 9% de diferença). Além disso, o estudo de associação caso-controle onde este polimorfismo foi utilizado não mostrou diferença estatística significativa entre os grupos.

No estudo de desequilíbrio de ligação nós encontramos uma forte relação entre os marcadores. No entanto, tanto no estudo de haplótipos quanto no estudo de TDT a frequência de transmissão entre os blocos / alelos se mostrou bastante parecida. Não podemos, assim, associar estas variantes às doenças cardíacas congênitas.

REFERÊNCIAS

- ABU-ABED, S.; DOLLE, P.; METZGER, D.; BECKETT, B.; CHAMBON, P.; PETKOVICH, M. The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. **Genes Dev**, v. 15, n. 2, p. 226-240. 2001.
- AZNAREZ, I.; CHAN, E.M.; ZIELENSKI, J.; BLENCOWE, B.J.; TSUI, L.C. Characterization of disease-associated mutations affecting an exonic splicing enhancer and two cryptic splice sites in exon 13 of the cystic fibrosis transmembrane conductance regulator gene. **Hum Mol Genet**, v. 12, n. 16, p. 2031-2040. 2003.
- BENSON, D.W.; SHARKEY, A.; FATKIN, D.; LANG, P.; BASSON, C.T.; MCDONOUGH, B.; STRAUSS, A.W.; SEIDMAN, J.G.; SEIDMAN, C.E. Reduced penetrance, variable expressivity, and genetic heterogeneity of familial atrial septal defects. **Circulation**, v. 97, n. 20, p. 2043-2048. 1998.
- BLENCOWE, B.J. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. **Trends Biochem Sci**, v. 25, n. 3, p. 106-110. 2000.
- BLENTIC, A.; GALE, E.; MADEN, M. Retinoic acid signalling centres in the avian embryo identified by sites of expression of synthesising and catabolising enzymes. **Dev Dyn**, v. 227, n. 1, p. 114-127. 2003.
- BRUNEAU, B.G. The developing heart and congenital heart defects: a make or break situation. **Clin Genet**, v. 63, n. 4, p. 252-261. 2003.
- BRUNEAU, B.G.; NEMER, G.; SCHMITT, J.P.; CHARRON, F.; ROBITAILLE, L.; CARON, S.; CONNER, D.A.; GESSLER, M.; NEMER, M.; SEIDMAN, C.E.; SEIDMAN, J.G. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. **Cell**, v. 106, n. 6, p. 709-721. 2001.
- BURATTI, E.; MURO, A.F.; GIOMBI, M.; GHERBASSI, D.; IACONCIG, A.; BARALLE, F.E. RNA folding affects the recruitment of SR proteins by mouse and human polypurinic enhancer elements in the fibronectin EDA exon. **Mol Cell Biol**, v. 24, n. 3, p. 1387-1400. 2004.
- BURN, T.C.; CONNORS, T.D.; KLINGER, K.W.; LANDES, G.M. Increased exon-trapping efficiency through modifications to the pSPL3 splicing vector. **Gene**, v. 161, n. 2, p. 183-187. 1995.
- CARDON, L.R.; BELL, J.I. Association study designs for complex diseases. **Nat Rev Genet**, v. 2, n. 2, p. 91-99. 2001.
- CARTEGNI, L.; CHEW, S.L.; KRAINER, A.R. Listening to silence and understanding nonsense: exonic mutations that affect splicing. **Nat Rev Genet**, v. 3, n. 4, p. 285-298. 2002.
- CARTEGNI, L.; HASTINGS, M.L.; CALARCO, J.A.; DE STANCHINA, E.; KRAINER, A.R. Determinants of exon 7 splicing in the spinal muscular atrophy genes, SMN1 and SMN2. **Am J Hum Genet**, v. 78, n. 1, p. 63-77. 2006.
- CARTEGNI, L.; KRAINER, A.R. Disruption of an SF2/ASF-dependent exonic splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1. **Nat Genet**, v. 30, n. 4, p. 377-384. 2002.
- CARTEGNI, L.; WANG, J.; ZHU, Z.; ZHANG, M.Q.; KRAINER, A.R. ESEfinder: A web resource to identify exonic splicing enhancers. **Nucleic Acids Res**, v. 31, n. 13, p. 3568-3571. 2003.

CHAMBERS, D.; WILSON, L.; MADEN, M.; LUMSDEN, A. RALDH-independent generation of retinoic acid during vertebrate embryogenesis by CYP1B1. **Development**, v. 134, n. 7, p. 1369-1383. 2007.

CHURCH, D.M.; STOTLER, C.J.; RUTTER, J.L.; MURRELL, J.R.; TROFATTER, J.A.; BUCKLER, A.J. Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. **Nat Genet**, v. 6, n. 1, p. 98-105. 1994.

CRABB, D.W.; MATSUMOTO, M.; CHANG, D.; YOU, M. Overview of the role of alcohol dehydrogenase and aldehyde dehydrogenase and their variants in the genesis of alcohol-related pathology. **Proc Nutr Soc**, v. 63, n. 1, p. 49-63. 2004.

DEAK, K.L.; DICKERSON, M.E.; LINNEY, E.; ENTERLINE, D.S.; GEORGE, T.M.; MELVIN, E.C.; GRAHAM, F.L.; SIEGEL, D.G.; HAMMOCK, P.; MEHLTRETTER, L.; BASSUK, A.G.; KESSLER, J.A.; GILBERT, J.R.; SPEER, M.C. Analysis of ALDH1A2, CYP26A1, CYP26B1, CRABP1, and CRABP2 in human neural tube defects suggests a possible association with alleles in ALDH1A2. **Birth Defects Res A Clin Mol Teratol**, v. 73, n. 11, p. 868-875. 2005.

DERSCH, H.; ZILE, M.H. Induction of normal cardiovascular development in the vitamin A-deprived quail embryo by natural retinoids. **Dev Biol**, v. 160, n. 2, p. 424-433. 1993.

D'SOUZA, I.; POORKAJ, P.; HONG, M.; NOCHLIN, D.; LEE, V.M.; BIRD, T.D.; SCHELLENBERG, G.D. Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements. **Proc Natl Acad Sci U S A**, v. 96, n. 10, p. 5598-5603. 1999.

DUPE, V.; MATT, N.; GARNIER, J.M.; CHAMBON, P.; MARK, M.; GHYSELINCK, N.B. A newborn lethal defect due to inactivation of retinaldehyde dehydrogenase type 3 is prevented by maternal retinoic acid treatment. **Proc Natl Acad Sci U S A**, v. 100, n. 24, p. 14036-14041. 2003.

FAN, X.; MOLOTKOV, A.; MANABE, S.; DONMOYER, C.M.; DELTOUR, L.; FOGLIO, M.H.; CUENCA, A.E.; BLANER, W.S.; LIPTON, S.A.; DUESTER, G. Targeted disruption of *Aldh1a1* (*Raldh1*) provides evidence for a complex mechanism of retinoic acid synthesis in the developing retina. **Mol Cell Biol**, v. 23, n. 13, p. 4637-4648. 2003.

GARG, V.; KATHIRIYA, I.S.; BARNES, R.; SCHLUTERMAN, M.K.; KING, I.N.; BUTLER, C.A.; ROTHROCK, C.R.; EAPEN, R.S.; HIRAYAMA-YAMADA, K.; JOO, K.; MATSUOKA, R.; COHEN, J.C.; SRIVASTAVA, D. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. **Nature**, v. 424, n. 6947, p. 443-447. 2003.

GARG, V.; MUTH, A.N.; RANSOM, J.F.; SCHLUTERMAN, M.K.; BARNES, R.; KING, I.N.; GROSSFELD, P.D.; SRIVASTAVA, D. Mutations in NOTCH1 cause aortic valve disease. **Nature**, v. 437, n. 7056, p. 270-274. 2005.

GOLDMUNTZ, E.; GEIGER, E.; BENSON, D.W. NKX2.5 mutations in patients with tetralogy of fallot. **Circulation**, v. 104, n. 21, p. 2565-2568. 2001.

GOWHER, H.; LOUTCHANWOOT, P.; VOROBJEVA, O.; HANDA, V.; JURKOWSKA, R.Z.; JURKOWSKI, T.P.; JELTSCH, A. Mutational analysis of the catalytic domain of the murine Dnmt3a DNA-(cytosine C5)-methyltransferase. **J Mol Biol**, v. 357, n. 3, p. 928-941. 2006.

GRUBER, P.J.; EPSTEIN, J.A. Development gone awry: congenital heart disease. **Circ Res**, v. 94, n. 3, p. 273-283. 2004.

GRUBER, P.J.; KUBALAK, S.W.; PEXIEDER, T.; SUCHOV, H.M.; EVANS, R.M.; CHIEN, K.R. RXR alpha deficiency confers genetic susceptibility for aortic sac, conotruncal, atrioventricular cushion, and ventricular muscle defects in mice. **J Clin Invest**, v. 98, n. 6, p. 1332-1343. 1996.

GURIS, D.L.; DUESTER, G.; PAPAIOANNOU, V.E.; IMAMOTO, A. Dose-dependent interaction of Tbx1 and Crkl and locally aberrant RA signaling in a model of del22q11 syndrome. **Dev Cell**, v. 10, n. 1, p. 81-92. 2006.

HARENDZA, S.; HUBNER, C.A.; GLASER, C.; BURDELSKI, M.; THAISS, F.; HANSMANN, I.; GAL, A.; STAHL, R.A. Renal failure and hypertension in Alagille syndrome with a novel JAG1 mutation. **J Nephrol**, v. 18, n. 3, p. 312-317. 2005.

HARVEY, L.B., A.; ZIPURSKY, L.S.; MATSUDAIRA, P.; BALTIMORE, D.; DARNELL, J. **Molecular Cell Biology**. 2000

HARVEY, R.P. NK-2 homeobox genes and heart development. **Dev Biol**, v. 178, n. 2, p. 203-216. 1996.

HARVEY, R.P. Patterning the vertebrate heart. **Nat Rev Genet**, v. 3, n. 7, p. 544-556. 2002.

HIRSCHHORN, J.N.; DALY, M.J. Genome-wide association studies for common diseases and complex traits. **Nat Rev Genet**, v. 6, n. 2, p. 95-108. 2005.

HOCHGREB, T.; LINHARES, V.L.; MENEZES, D.C.; SAMPAIO, A.C.; YAN, C.Y.; CARDOSO, W.V.; ROSENTHAL, N.; XAVIER-NETO, J. A caudorostral wave of RALDH2 conveys anteroposterior information to the cardiac field. **Development**, v. 130, n. 22, p. 5363-5374. 2003.

HOFFMAN, J.I. Incidence of congenital heart disease: II. Prenatal incidence. **Pediatr Cardiol**, v. 16, n. 4, p. 155-165. 1995.

HUMPHREY, W.; DALKE, A.; SCHULTEN, K. VMD: visual molecular dynamics. **J Mol Graph**, v. 14, n. 1, p. 33-38, 27-38. 1996.

JAY, P.Y.; ROZHITSKAYA, O.; TARNAVSKI, O.; SHERWOOD, M.C.; DORFMAN, A.L.; LU, Y.; UYEYAMA, T.; IZUMO, S. Haploinsufficiency of the cardiac transcription factor Nkx2-5 variably affects the expression of putative target genes. **Faseb J**, v. 19, n. 11, p. 1495-1497. 2005.

JURKIEWICZ, D.; POPOWSKA, E.; GLASER, C.; HANSMANN, I.; KRAJEWSKA-WALASEK, M. Twelve novel JAG1 gene mutations in Polish Alagille syndrome patients. **Hum Mutat**, v. 25, n. 3, p. 321. 2005.

KASTNER, P.; MESSADDEQ, N.; MARK, M.; WENDLING, O.; GRONDONA, J.M.; WARD, S.; GHYSELINCK, N.; CHAMBON, P. Vitamin A deficiency and mutations of RXRalpha, RXRbeta and RARalpha lead to early differentiation of embryonic ventricular cardiomyocytes. **Development**, v. 124, n. 23, p. 4749-4758. 1997.

KAWAGUCHI, R.; YU, J.; HONDA, J.; HU, J.; WHITELEGGE, J.; PING, P.; WIITA, P.; BOK, D.; SUN, H. A membrane receptor for retinol binding protein mediates cellular uptake of vitamin A. **Science**, v. 315, n. 5813, p. 820-825. 2007.

KRAWCZAK, M.; THOMAS, N.S.; HUNDRIESER, B.; MORT, M.; WITTIG, M.; HAMPE, J.; COOPER, D.N. Single base-pair substitutions in exon-intron junctions of human genes: nature, distribution, and consequences for mRNA splicing. **Hum Mutat**, v. 28, n. 2, p. 150-158. 2007.

KWOK, J.B.; HALLUPP, M.; LOY, C.T.; CHAN, D.K.; WOO, J.; MELLICK, G.D.; BUCHANAN, D.D.; SILBURN, P.A.; HALLIDAY, G.M.; SCHOFIELD, P.R. GSK3B polymorphisms alter transcription and splicing in Parkinson's disease. **Ann Neurol**, v. 58, n. 6, p. 829-839. 2005.

LAMB, A.L.; NEWCOMER, M.E. The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity. **Biochemistry**, v. 38, n. 19, p. 6003-6011. 1999.

LYONS, I.; PARSONS, L.M.; HARTLEY, L.; LI, R.; ANDREWS, J.E.; ROBB, L.; HARVEY, R.P. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. **Genes Dev**, v. 9, n. 13, p. 1654-1666. 1995.

MARCHINI, J.; CARDON, L.R.; PHILLIPS, M.S.; DONNELLY, P. The effects of human population structure on large genetic association studies. **Nat Genet**, v. 36, n. 5, p. 512-517. 2004.

MILED, N.; YAN, Y.; HON, W.C.; PERISIC, O.; ZVELEBIL, M.; INBAR, Y.; SCHNEIDMAN-DUHOVNY, D.; WOLFSON, H.J.; BACKER, J.M.; WILLIAMS, R.L. Mechanism of two classes of cancer mutations in the phosphoinositide 3-kinase catalytic subunit. **Science**, v. 317, n. 5835, p. 239-242. 2007.

MODREK, B.; LEE, C. A genomic view of alternative splicing. **Nat Genet**, v. 30, n. 1, p. 13-19. 2002.

MOLOTKOV, A.; FAN, X.; DELTOUR, L.; FOGLIO, M.H.; MARTRAS, S.; FARRES, J.; PARES, X.; DUESTER, G. Stimulation of retinoic acid production and growth by ubiquitously expressed alcohol dehydrogenase Adh3. **Proc Natl Acad Sci U S A**, v. 99, n. 8, p. 5337-5342. 2002.

MORRIS, G.M. Morphogenesis of the malformations induced in rat embryos by maternal hypervitaminosis A. **J Anat**, v. 113, n. Pt 2, p. 241-250. 1972.

MOSELEY, C.T.; MULLIS, P.E.; PRINCE, M.A.; PHILLIPS, J.A., 3RD. An exon splice enhancer mutation causes autosomal dominant GH deficiency. **J Clin Endocrinol Metab**, v. 87, n. 2, p. 847-852. 2002.

MOSS, J.B.; XAVIER-NETO, J.; SHAPIRO, M.D.; NAYEEM, S.M.; MCCAFFERY, P.; DRAGER, U.C.; ROSENTHAL, N. Dynamic patterns of retinoic acid synthesis and response in the developing mammalian heart. **Dev Biol**, v. 199, n. 1, p. 55-71. 1998.

NIEDERREITHER, K.; ABU-ABED, S.; SCHUHBAUR, B.; PETKOVICH, M.; CHAMBON, P.; DOLLE, P. Genetic evidence that oxidative derivatives of retinoic acid are not involved in retinoid signaling during mouse development. **Nat Genet**, v. 31, n. 1, p. 84-88. 2002.

NIEDERREITHER, K.; MCCAFFERY, P.; DRAGER, U.C.; CHAMBON, P.; DOLLE, P. Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development. **Mech Dev**, v. 62, n. 1, p. 67-78. 1997.

NIEDERREITHER, K.; SUBBARAYAN, V.; DOLLE, P.; CHAMBON, P. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. **Nat Genet**, v. 21, n. 4, p. 444-448. 1999.

NIEDERREITHER, K.; VERMOT, J.; MESSADDEQ, N.; SCHUHBAUR, B.; CHAMBON, P.; DOLLE, P. Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. **Development**, v. 128, n. 7, p. 1019-1031. 2001.

OLSON, E.N. A decade of discoveries in cardiac biology. **Nat Med**, v. 10, n. 5, p. 467-474. 2004.

OSMOND, M.K.; BUTLER, A.J.; VOON, F.C.; BELLAIRS, R. The effects of retinoic acid on heart formation in the early chick embryo. **Development**, v. 113, n. 4, p. 1405-1417. 1991.

PARK, S.; JOHNSON, M.E.; FUNG, L.W. Nuclear magnetic resonance studies of mutations at the tetramerization region of human alpha spectrin. **Blood**, v. 100, n. 1, p. 283-288. 2002.

PASUTTO, F.; STICHT, H.; HAMMERSEN, G.; GILLESSEN-KAESBACH, G.; FITZPATRICK, D.R.; NURNBERG, G.; BRASCH, F.; SCHIRMER-ZIMMERMANN, H.; TOLMIE, J.L.; CHITAYAT, D.; HOUGE, G.; FERNANDEZ-MARTINEZ, L.; KEATING, S.; MORTIER, G.; HENNEKAM, R.C.; VON DER WENSE, A.; SLAVOTINEK, A.; MEINECKE, P.; BITOUN, P.; BECKER, C.; NURNBERG, P.; REIS, A.; RAUCH, A. Mutations in STRA6 cause a broad spectrum of malformations including anophthalmia, congenital heart defects, diaphragmatic hernia, alveolar capillary dysplasia, lung hypoplasia, and mental retardation. **Am J Hum Genet**, v. 80, n. 3, p. 550-560. 2007.

PERLMANN, T. Retinoid metabolism: a balancing act. **Nat Genet**, v. 31, n. 1, p. 7-8. 2002.

PERLOFF, J.K. Congenital heart disease and pregnancy. **Clin Cardiol**, v. 17, n. 11, p. 579-587. 1994.

QUADRO, L.; BLANER, W.S.; SALCHOW, D.J.; VOGEL, S.; PIANTEDOSI, R.; GOURAS, P.; FREEMAN, S.; COSMA, M.P.; COLANTUONI, V.; GOTTESMAN, M.E. Impaired retinal function and vitamin A availability in mice lacking retinol-binding protein. **Embo J**, v. 18, n. 17, p. 4633-4644. 1999.

RANSOM, J.; SRIVASTAVA, D. The genetics of cardiac birth defects. **Semin Cell Dev Biol**, v. 18, n. 1, p. 132-139. 2007.

RAY, W.J.; BAIN, G.; YAO, M.; GOTTLIEB, D.I. CYP26, a novel mammalian cytochrome P450, is induced by retinoic acid and defines a new family. **J Biol Chem**, v. 272, n. 30, p. 18702-18708. 1997.

REAMON-BUETTNER, S.M.; BORLAK, J. Somatic NKX2-5 mutations as a novel mechanism of disease in complex congenital heart disease. **J Med Genet**, v. 41, n. 9, p. 684-690. 2004.

REIJNTJES, S.; GALE, E.; MADEN, M. Generating gradients of retinoic acid in the chick embryo: Cyp26C1 expression and a comparative analysis of the Cyp26 enzymes. **Dev Dyn**, v. 230, n. 3, p. 509-517. 2004.

RISCH, N.J. Searching for genetic determinants in the new millennium. **Nature**, v. 405, n. 6788, p. 847-856. 2000.

ROSENTHAL, N.; HARVEY, R.P. Single allele mutations at the heart of congenital disease. **J Clin Invest**, v. 104, n. 11, p. 1483-1484. 1999.

ROSENTHAL, N.; XAVIER-NETO, J. From the bottom of the heart: anteroposterior decisions in cardiac muscle differentiation. **Curr Opin Cell Biol**, v. 12, n. 6, p. 742-746. 2000.

SAKAI, Y.; MENO, C.; FUJII, H.; NISHINO, J.; SHIRATORI, H.; SAIJOH, Y.; ROSSANT, J.; HAMADA, H. The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the antero-posterior axis within the mouse embryo. **Genes Dev**, v. 15, n. 2, p. 213-225. 2001.

SANDELL, L.L.; SANDERSON, B.W.; MOISEYEV, G.; JOHNSON, T.; MUSHEGIAN, A.; YOUNG, K.; REY, J.P.; MA, J.X.; STAEBLING-HAMPTON, K.; TRAINOR, P.A. RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development. **Genes Dev**, v. 21, n. 9, p. 1113-1124. 2007.

SAWIN, P. Genetic background of transformation of aortic sacs in rabbit. **Anat. Rec.**, v. 100, n., p. 708. 1949.

SCHAID, D.J. Transmission disequilibrium, family controls, and great expectations. **Am J Hum Genet**, v. 63, n. 4, p. 935-941. 1998.

SCHORK, N.J. Genetics of complex disease: approaches, problems, and solutions. **Am J Respir Crit Care Med**, v. 156, n. 4 Pt 2, p. S103-109. 1997.

SCHOTT, J.J.; BENSON, D.W.; BASSON, C.T.; PEASE, W.; SILBERBACH, G.M.; MOAK, J.P.; MARON, B.J.; SEIDMAN, C.E.; SEIDMAN, J.G. Congenital heart disease caused by mutations in the transcription factor NKX2-5. **Science**, v. 281, n. 5373, p. 108-111. 1998.

SEIDMAN, J.G.; SEIDMAN, C. Transcription factor haploinsufficiency: when half a loaf is not enough. **J Clin Invest**, v. 109, n. 4, p. 451-455. 2002.

SHEN, L.X.; BASILION, J.P.; STANTON, V.P., JR. Single-nucleotide polymorphisms can cause different structural folds of mRNA. **Proc Natl Acad Sci U S A**, v. 96, n. 14, p. 7871-7876. 1999.

SIMOES-COSTA, M.S.; AZAMBUJA, A.P.; XAVIER-NETO, J. The search for non-chordate retinoic acid signaling: lessons from chordates. **J Exp Zool B Mol Dev Evol**, v., n., p. 2006.

SIMOES-COSTA, M.S.; VASCONCELOS, M.; SAMPAIO, A.C.; CRAVO, R.M.; LINHARES, V.L.; HOCHGREB, T.; YAN, C.Y.; DAVIDSON, B.; XAVIER-NETO, J. The evolutionary origin of cardiac chambers. **Dev Biol**, v. 277, n. 1, p. 1-15. 2005.

SPIELMAN, R.S.; MCGINNIS, R.E.; EWENS, W.J. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). **Am J Hum Genet**, v. 52, n. 3, p. 506-516. 1993.

SRIVASTAVA, D.; OLSON, E.N. A genetic blueprint for cardiac development. **Nature**, v. 407, n. 6801, p. 221-226. 2000.

STRACHAN, T.R., A. P. **Genética Molecular Humana**. 2002. 262-264 p.

SUMANDEA, C.A.; L, W.M.F. Mutational effects at the tetramerization site of nonerythroid alpha spectrin. **Brain Res Mol Brain Res**, v. 136, n. 1-2, p. 81-90. 2005.

SWINDELL, E.C.; EICHELE, G. Retinoid metabolizing enzymes in development. **Biofactors**, v. 10, n. 2-3, p. 85-89. 1999.

SWINDELL, E.C.; THALLER, C.; SOCKANATHAN, S.; PETKOVICH, M.; JESSELL, T.M.; EICHELE, G. Complementary domains of retinoic acid production and degradation in the early chick embryo. **Dev Biol**, v. 216, n. 1, p. 282-296. 1999.

TABOR, H.K.; RISCH, N.J.; MYERS, R.M. Candidate-gene approaches for studying complex genetic traits: practical considerations. **Nat Rev Genet**, v. 3, n. 5, p. 391-397. 2002.

ULVEN, S.M.; GUNDERSEN, T.E.; WEEDON, M.S.; LANDAAS, V.O.; SAKHI, A.K.; FROMM, S.H.; GERONIMO, B.A.; MOSKAUG, J.O.; BLOMHOFF, R. Identification of endogenous retinoids, enzymes, binding proteins, and receptors during early postimplantation development in mouse: important role of retinal dehydrogenase type 2 in synthesis of all-trans-retinoic acid. **Dev Biol**, v. 220, n. 2, p. 379-391. 2000.

VERMOT, J.; NIEDERREITHER, K.; GARNIER, J.M.; CHAMBON, P.; DOLLE, P. Decreased embryonic retinoic acid synthesis results in a DiGeorge syndrome phenotype in newborn mice. **Proc Natl Acad Sci U S A**, v. 100, n. 4, p. 1763-1768. 2003.

WANG, J.; TAKAGAKI, Y.; MANLEY, J.L. Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. **Genes Dev**, v. 10, n. 20, p. 2588-2599. 1996.

WEISS, K.M.; TERWILLIGER, J.D. How many diseases does it take to map a gene with SNPs? **Nat Genet**, v. 26, n. 2, p. 151-157. 2000.

WILSON, J.G.; BARCH, S. Fetal death and maldevelopment resulting from maternal vitamin A deficiency in the rat. **Proc Soc Exp Biol Med**, v. 72, n. 3, p. 687-693, illust. 1949.

WILSON, J.G.; WARKANY, J. Cardiac and aortic arch anomalies in the offspring of vitamin A deficient rats correlated with similar human anomalies. **Pediatrics**, v. 5, n. 4, p. 708-725. 1950.

XAVIER-NETO, J.; NEVILLE, C.M.; SHAPIRO, M.D.; HOUGHTON, L.; WANG, G.F.; NIKOVITS, W., JR.; STOCKDALE, F.E.; ROSENTHAL, N. A retinoic acid-inducible transgenic marker of sino-atrial development in the mouse heart. **Development**, v. 126, n. 12, p. 2677-2687. 1999.

XAVIER-NETO, J.; ROSENTHAL, N.; SILVA, F.A.; MATOS, T.G.; HOCHGREB, T.; LINHARES, V.L. Retinoid signaling and cardiac anteroposterior segmentation. **Genesis**, v. 31, n. 3, p. 97-104. 2001.

XAVIER-NETO, J.P., A.C.; AZAMBUJA, A.P. **Bases Embriológicas e Genéticas das Doenças Cardíacas Congênitas**: Manole, Barueri. 2005. 1761 - 1781 p. (In: Tratado de Cardiologia SOCESP)

XU, X.; YANG, D.; DING, J.H.; WANG, W.; CHU, P.H.; DALTON, N.D.; WANG, H.Y.; BERMINGHAM, J.R., JR.; YE, Z.; LIU, F.; ROSENFELD, M.G.; MANLEY, J.L.; ROSS, J., JR.; CHEN, J.; XIAO, R.P.; CHENG, H.; FU, X.D. ASF/SF2-regulated CaMKII δ alternative splicing temporally reprograms excitation-contraction coupling in cardiac muscle. **Cell**, v. 120, n. 1, p. 59-72. 2005.

YUAN, Z.R.; KOHSAKA, T.; IKEGAYA, T.; SUZUKI, T.; OKANO, S.; ABE, J.; KOBAYASHI, N.; YAMADA, M. Mutational analysis of the Jagged 1 gene in Alagille syndrome families. **Hum Mol Genet**, v. 7, n. 9, p. 1363-1369. 1998.

ZHAO, D.; MCCAFFERY, P.; IVINS, K.J.; NEVE, R.L.; HOGAN, P.; CHIN, W.W.; DRAGER, U.C. Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. **Eur J Biochem**, v. 240, n. 1, p. 15-22. 1996.