FERNANDO JORGE PEGO CRISTO

Molecular and functional analysis of DAND5 in human Congenital Heart Disease (CHD)

UNIVERSIDADE DO ALGARVE

Departamento de Ciências Biomédicas e Medicina

2016

FERNANDO JORGE PEGO CRISTO

Molecular and functional analysis of DAND5 in human Congenital Heart Disease (CHD)

Doutoramento em Ciências Biomédicas

Trabalho efetuado sob a orientação de:

Professor Doutor José António Henriques de Conde Belo

UNIVERSIDADE DO ALGARVE

Departamento de Ciências Biomédicas e Medicina

2016

Molecular and functional analysis of DAND5 in human Congenital Heart Disease (CHD)

Declaração de autoria de trabalho

Declaro ser o autor deste trabalho, que é original e inédito. Autores e trabalhos consultados estão devidamente citados no texto e constam da listagem de referências incluída.

_____

Dedicated to all the people who never stop believing in me ……

“Your time is limited, so don‘t waste it living someone else‘s life. Don‘t be trapped by dogma – which is living with the results of other people‘s thinking. Don‘t let the noise of other‘s opinions drown out your own inner voice. And most important, have the courage to follow your heart and intuition. They somehow already know what you truly want to become. Everything else is secondary”

Steve Jobs

Acknowledgements

Durante estes últimos anos, muitos foram aqueles com quem, trabalhei, troquei opiniões, partilhei ideias, aprendi e muitas vezes me diverti. Tive alguns amigos que sempre me apoiaram, ouviram as minhas tristezas e as minhas alegrias. A todos aqueles que ao longo deste período me ofereceram uma imutável e inestimável ajuda, permitiram a realização deste trabalho e de alguma forma contribuíram para o meu enriquecimento pessoal quero expressar os meus sinceros agradecimentos de forma especial, pois especial foi também o que me transmitiram.

Dito isto, desejo expressar os meus sinceros agradecimentos:

Ao Prof. José A. Belo, meu orientador, pela confiança em mim depositada para levar a cabo este projeto no seu laboratório, por todo o suporte e pela disponibilidade e generosidade reveladas ao longo destes anos que temos trabalhado juntos.

Ao José Inácio por todas as perguntas “parvas” que teve de me responder e todos os dias que teve de me aturar, por me ter ensinado quase todas as técnicas que sei até hoje e principalmente pela pessoa que é. Obrigado Zé por toda a paciência, disponibilidade e companheirismo.

À Salomé Almeida, por todo o suporte, disponibilidade, competência e boa disposição. Muito obrigado Salomé por acreditares em mim, no meu trabalho e dizeres sempre aquelas palavrinhas para levantar o animo.

Ao pessoal do laboratório, nomeadamente, Elizabeth, Sara, Marta, Marisa, Rita, Margaret, João Facucho, João Furtado (el pollo), Carolina (Brasileira), Ana Rubina (Rubi), Maria Silva, Oriol Bover e Paulo Pereira pela forma como me receberam e por toda a amizade, companheirismo e ensinamento durante este longo percurso. Um agradecimento especial ao Tiago Justo pelo excelente ano em que decidimos aventurar-nos por Lisboa, pelo companheirismo e ensinamentos tantos profissional como pessoal.

Um muito muito obrigado a todos os meus amigos de S.Tiago e claro aos da faculdade que estão sempre presentes, João Baptista e Nídia Cunha. Um agradecimento especial à Raquel Santa Maria por todo o suporte e acreditar em mim até mais do eu. Muito obrigado és uma pessoa espetacular, um verdadeiro exemplo, tanto profissionalmente como pessoalmente.

Quero também agradecer ao “pessoal do CEDOC” pela forma como me receberam no vosso grupo de amigos e pelos dias e noite bem passadas. Em especial quero agradecer à Inês Santarino por tudo o que fez por mim e pela pessoa espetacular que é, a tua amizade foi e é muito importante para mim. Uma palavra de apreço também á “Tati”, é muito difícil encontrar pessoas como tu e com o teu caracter. Temos uma amizade para uma vida. Obrigado por todos os conselhos.

Quero agradecer também ao CBMR e CEDOC pelas excelentes condições de trabalho.

Não podia deixar de agradecer também à Carla Machado por todo o suporte, por todos as noites que me stressava e me aturavas, por toda a compreensão, carinho e amizade. És uma pessoa única.

Por último, aos meus avós, tia, irmão e aos meus pais: a vocês o meu muito obrigado por todo o apoio incondicional, carinho, compreensão e por estarem sempre presentes nos momentos mais difíceis. Obrigada por tudo, Amo-vos!

Resumo

Uma das questões fulcrais da biologia do desenvolvimento continua a ser como é que uma única célula, o oócito fecundado, se desenvolve num ser tão complexo e especializado, com os tecidos e órgãos diferenciados e organizados em torno de três eixos, o anterior-posterior (AP), dorso-ventral (DV) e esquerdo-direito (LR).

De um modo geral, o desenvolvimento embrionário humano pode ser dividido em três períodos: o primeiro, chamado de período zigótico, inicia-se com a fusão dos gametas masculino e feminino, formação do zigoto e termina com a implantação do blastocisto na parede do útero duas semanas após a fecundação. Durante o segundo período, ou período embrionário, que vai desde a terceira até à oitava semana de gestação, ocorre a gastrulação, com a formação e diferenciação dos folhetos embrionários (ectoderme, mesoderme e endoderme), e a formação dos órgãos e sistemas fisiológicos. O ultimo período, chamado fetal, vai desde o terceiro mês de gestação até ao nascimento e é caracterizado pelo crescimento do embrião, maturação e funcionalização das estruturas e órgãos.

Durante o desenvolvimento do embrião, o coração é o primeiro órgão interno a ser formado e a tornar-se funcional de modo a bombear sangue por volta das três semanas de gestação, suprindo as necessidades do embrião e garantindo a sua sobrevivência. Na maioria dos vertebrados o desenvolvimento cardíaco segue o mesmo padrão de formação. O coração forma-se a partir de duas regiões cardíacas bilaterais que advêm de uma região progenitora comum na gastrulação, a mesoderme cardíaca. A mesoderme cardíaca recebe informação posicional dos três eixos embrionários, porém o efeito do eixo Esquerda-Direita (ED) tem recebido particular atenção. Os dois "*pools*" bilaterais de células precursoras cardíacas mesodermais migram para a linha média onde se juntam e dão origem ao crescente cardíaco. Este crescente cardíaco vai formar posteriormente um tubo cardíaco linear primordial. De seguida, este tubo alonga e sofre um processo de "looping" para a direita levando à formação dos ventrículos e aurículas. Subsequentemente, ocorre a diferenciação, especialização e septação das quatro camaras cardíacas e válvulas

proporcionando as paredes necessárias para a separação física da circulação sistémica e pulmonar de sangue. Por último a formação do sistema de condução cardíaco e maturação dos cardiomiócitos, culminam na formação de um coração inteiramente funcional.

As doenças cardíacas congénitas (DCC) são a manifestação clínica de problemas durante o desenvolvimento embrionário do coração. Esta doença é a forma mais comum de defeito à nascença, ocorrendo em cerca de 8 em 1000 nados vivos, bem como a principal causa não-infeciosa de morte no primeiro ano de vida. Na maioria dos casos (80%), esta doença tem origem desconhecida e manifesta-se isoladamente em doentes não sindrómicos e não seguindo os padrões de hereditariedade definidos por Mendel, sendo considerada uma doença complexa com origem multifatorial. Esta complexidade advém do facto de a doença em muitos casos se apresentar geneticamente heterogenica, com penetrância reduzida e expressividade genética variável. Para além disso, fatores de risco ambientais foram também identificados como contribuído para o risco de desenvolver a doença. Nos dias de hoje, um modelo envolvendo tanto fatores genéticos como a interação genes-ambiente, é a explicação mais plausível para as altas frequências verificadas para a doença cardíaca congénita.

Nos últimos anos, vários esforços foram feitos de modo a se compreender os mecanismos moleculares que levam a uma má formação do coração. Estes estudos, usando principalmente animais modelo e genética molecular humana, da morfogénese do coração e dos primeiros passos da determinação do eixo de simetria esquerdo-direito durante a embriogénese, indicam que na maioria dos casos de distúrbios de lateralidade é observada uma malformação cardíaca complexa, sugerindo que as DCC podem dever-se a defeitos de lateralidade na morfogénese do coração.

Uma das redes regulatórias por detrás do estabelecimento do eixo esquerdo-direito é a via de sinalização Nodal, um fator de crescimento da família TGF-b, que é expresso assimetricamente do lado esquerdo do nó e placa lateral esquerda da mesoderme de ratinho. Os fenótipos cardíacos de ratinhos geneticamente modificados para desativar a sinalização de nodal ou a sua

regulação, demonstraram que o desenvolvimento cardíaco depende vivamente do estabelecimento do eixo E-D embrionário. Para além disso, foram identificadas variantes genéticas em genes envolvidos na via de sinalização Nodal em doentes com DCC.

O nosso laboratório identificou um membro da família Cerberus/DAN, o gene *Cerberus-like2* (*Cerl2*), que é assimetricamente expresso no lado direito do nó em ratinho. O gene *Cerl2* codifica para uma proteína secretada capaz de se liga diretamente a Nodal, inibindo a sua via de sinalização. Na presença de *Cerl2* a atividade da via de sinalização Nodal restringe-se apenas ao lado esquerdo do embrião, o que culmina com uma morfogénese normal dos órgãos do abdómem. Ratinhos knock-out para o gene *Cerl2*, em que o gene *Nodal* passa a poder ser expresso e induzir a sua cascata de sinalização em ambos os lados da placa lateral da mesoderme, revelaram uma ampla gama de defeitos de lateralidade - situs inversus, isomerismos e heterotaxia - e malformações cardíacas - formação incompleta das aurículas, defeitos no septo ventricular, falha na rotação das principais artérias (transposição das grandes artérias, ventrículo direito de dupla saída), looping cardíaco anormal (randomizado), posicionamento do ápex cardíaco e hipertrofia ventricular - o que nos levou a suspeitar que alterações neste gene poderiam estar associadas a doenças de lateralidade e/ou casos isolados de doenças cardíacas congénitas.

Dado que as vias genéticas que regulam o desenvolvimento cardíaco em ratinhos e humanos são conservadas, o principal objetivo desta tese foi o estudo de alguns dos genes homólogos humanos envolvidos na via de sinalização Nodal, com principal foco no gene *DAND5*, homólogo humano de *Cerl2,* mapeado no cromossoma 19, região 19p13.2, mas também verificando a existência de variantes nos genes *NODAL*, *PITX2C* e *CFC1* numa cohort de doentes com defeitos cardíacos congénitos com origem em perturbações do eixo esquerdo-direito.

A análise dos resultados da sequenciação de DNA do gene *DAND5* permitiu-nos identificar dois doentes com a mesma variação. Clinicamente, um dos doentes apresenta um fenótipo de defeito no septo ventricular com aorta

conectada aos dois ventrículos, atresia pulmonar e isomerismo esquerdo. O outro doente apresenta um caso extremo de tetralogia de Fallot (defeito no septo ventricular com aorta conectada aos dois ventrículos, hipertrofia do ventrículo esquerdo e atresia pulmonar). A variante, identificada nos dois doentes como heterozigótica, leva à substituição de um nucleótido guanina por uma adenina na posição 455 do exão 2, resultando na substituição de um aminoácido arginina por uma histidina (p.R152H) numa região altamente conservada e funcionalmente importante da proteína DAND5. De modo a entendermos o potencial efeito da alteração p.R152H na proteína DAND5, fizemos um estudo *in silico* recorrendo a vários programas bioinformáticos. Os resultados dessas predições foram inconclusivos uma vez que 3 programas apontam para um possível efeito nefasto enquanto que outro programa sugere que a variante não tem qualquer efeito na proteína. Para clarificarmos a função desta proteína variante, avaliámos o seu efeito na regulação da via de sinalização Nodal recorrendo a um ensaio funcional de luciferase. O resultado obtido mostra uma diminuição substancial da função da proteína mutante comparada com a proteína wild-type.

Embora os fenótipos dos doentes sejam muito semelhantes, não podemos fazer uma clara correlação do genótipo com o fenótipo pois a variante c.455G>A foi reportada em bases de dados publicas como variante de nucleótido único (rs45513495). Além disso, a mãe de um dos doentes é portadora da variante sem apresentar indícios de doença, sugerindo uma penetrância incompleta, expressividade genética variável ou o efeito de fatores ambientais. Estas observações estão de acordo com a complexidade das doenças cardíacas congénitas e/ou defeitos de lateralidade e podem refletir a acção de variantes modificadoras ou outras variantes genéticas na via de sinalização que podem exacerbar ou atenuar o efeito da variante p.R152H na proteína DAND5 e nos fenótipos dos doentes. Portanto, nós propomos que esta variação no gene *DAND5* pode ser um alelo de risco para o desenvolvimento de DCC e/ou defeitos de lateralidade.

Por esta razão, decidimos levar a cabo uma busca de possíveis alterações em genes que fazem parte desta cascata Nodal. Foram assim identificadas uma nova alteração na zona codificante do gene *CFC1*, sem efeito funcional

aparente, e quatro polimorfismos, um na zona codificante e outro no intrão do gene *NODAL* e dois na zona não codificante do gene *PITX2C* em doentes com uma vasta gama de defeitos. Tendo em conta a análise levada a cabo por um programa bioinformático que permite analisar se alterações genéticas podem afetar o processo de splicing, nós verificamos que de facto, as variantes genómicas encontradas fora das zonas codificantes dos genes *NODAL* e *PITX2C* podem resultar em moléculas de RNA mensageiro anormais, principalmente devido à criação ou disrupção de locais para a ligação da maquinaria de splicing. Quanto à variante na zona codificante do gene *NODAL*, esta encontra-se no exão 2, e leva à substituição do aminoácido Histidina por uma Arginina na posição 165 (p.H165R). Esta variante já tinha sido reportada em dois estudos de 2008/2009 nos quais os autores verificaram, recorrendo a um ensaio funcional de lucifarese, que a variante leva à redução da atividade da proteína NODAL e parece atuar como variante modificadora e como fator de risco quando associada a outras possíbeis alterações em genes da via Nodal, com os quais este normalmente interage.

Para além disto, e com o objetivo de modelar a doença e estudar os mecanismos moleculares por de trás de uma simples variação num nucleótido, geramos células estaminais induzidas de um dos doentes com a variante no gene *DAND5*. Estas células foram caracterizadas, apresentado uma morfologia, expressão de marcadores pluripotentes e cariotipo normais.

Em conclusão, embora não possamos fazer uma correlação dos genótipo-fenótipo, nem classificar as alterações identificadas neste estudo como alelos associadas a doença por si, estas variantes, podem, no entanto, aumentar a suscetibilidade para o desenvolvimento de DCC e/ou defeitos de lateralidade. Uma vez que a maioria dos doentes apresenta mais do que uma das variantes no seu genoma, o efeito cumulativo de cada variante na via parece aumentar ainda mais o risco para desenvolver doença. Portanto, o desequilíbrio dos níveis adequados da via de sinalização Nodal, em ambos os lados da placa lateral da mesoderme, devido a uma ou várias variantes nos seus componentes, é um denominador comum para defeitos de lateralidade e/ou doenças cardíacas congénitas.

Palavras-chave: DAND5; Doenças cardíacas congénitas Defeitos de lateralidade, Via de sinalização Nodal, Variantes alélicas; Modelação de doença, Células estaminais induzidas

Abstract

The majority of congenital heart disease (CHD) is sporadic, with a minority of cases associated with a known genetic abnormality. Combinations of genetic-environmental factors are implicated in the etiology of the disease. Recently, several studies, using mostly animal models, unraveled that perturbations in the molecular processes that precede the beginning of heart development might also be at the origin of CHD. In fact, some of the most complex CHDs are found associated with laterality defects, a disorder resulting from abnormal Left-Right axis formation. In our laboratory, the identified mouse Cerberus-like2 (Cerl2 – human DAND5), a protein that inhibit Nodal signaling, prompt us to study cardiac and laterality diseases, since the generated *Cerl2* KO mice display a wide range of laterality defects and CHD. Considering the high conservation of genetic pathways regulating cardiac development in mouse and human, the main objective of the present thesis was the study of human genes involved in the Nodal pathway, focusing mostly in *DAND5*, in a CHD and/or laterality defects patients cohort. The sequence analysis of *DAND5* revealed two patients displaying the same p.R152H variant, resulting in a substantial decreased in the function of the protein. We propose that p.R152H acts as a risk allele for CHD and/or laterality defects. In addition, we found two alterations in *NODAL*, two alterations in *PITX2C* and one alteration in *CFC1*. We hypothesized that the NODAL p.H165R variant can act as a common modifier and the intronic variants in *NODAL* and *PITX2C* might cause alterations in the splicing pattern of the mRNA molecules. Moreover, we generated patient-specific iPSCs to understand the molecular mechanisms of disease behind the *DAND5* nucleotide variant. Although we cannot make a clearly genotype-phenotype correlation, the variants here identified probably increase the disease susceptibility due to the resulting abnormal Nodal signaling. Because most of the patients presented more than one alteration, the cumulative effect of each variant within the pathway seems to enhance even more these risk. Therefore, the imbalance in dosage-sensitive Nodal signaling is a common denominator for laterality defects and associated CHDs.

Keywords: DAND5, Congenital Heart Diseases, Laterality defects, Nodal signaling, allelic variations, Disease modeling, iPSCs.

List of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Part I

Part II

Part III

List of abbreviations, acronyms and symbols

°C – Degrees Celsius
µg – Microgram
µl – Microliter
µM – Micro molar
3’ – 3 prime
5’ – 5 prime
β-GAL – β-galactosidase
µm – micrometer

A
A – Adenine
ActRIIA – Activin receptor type IIA
ActRIIB – Activin receptor type IIB
AKL7 – Activin receptor-like kinase 7
ALK4 – Activin receptor-like kinase 4
AP – Anterior-posterior
ARE – Activin responsive element

ASD – Atrial septal defects
ASEs – Asymmetric enhancers
AVC – Atrioventricular canal
AVSD – Atrioventricular septal defects

B
BLAST – Basic Local Alignment Search Tool
BMP – Bone morphogenetic protein
Bp – Base pair

C
C – Cytosine
Ca^{2+} – Calcium
Cerl2 – Cerberus like 2
CHD – Congenital heart disease/defects
CNC – Cardiac neural crest
CNVs – Copy number variants

D
Da – Dalton
DAND5 – DAN Domain Family Member 5

DILV – Double inlet left ventricle
DNA – Deoxyribonucleic acid dNTP – Deoxyribonucleotide triphosphate
DORV – Double outlet right ventricle
DV – Dorso-ventral

Dvl – Dishevelled

E
E – Embryonic day
ECM – Extracellular matrix
EGF-CFC – epidermal growth factor-Cripto-FRL1-Cryptic
EHF – Early head-fold
EMT – Epithelial-mesenchymal transition
EPE – Entidade Pública Empresarial
ESCs – Embryonic stem cells

F
FGF – Fibroblast growth factor
FHF – First heart field
FoxH1 – Forkhead box protein H1

G
G – Guanine
GAGs – Glycosaminoglycans
GATA4 – GATA Binding Protein 4
GDF1 – Growth differentiation factor 1

H
HAND2 – Heart- and neural crest derivatives-expressed protein 2
HEK – Human Embryonic Kidney
HTX – Heterotaxy

I
ICM – Inner cell mass
ICOs – Intraciliary calcium oscillations
Information
iPSCs - induced pluripotent stem cells
Isl1 – ISL LIM Homeobox 1
IV – *Inversus viscerum*
IVS – Interventricular septum
K
Kif3a – Kinesin family member 3A
Klf4 – Kruppel-like factor 4
KO – Knock-out

L
LA – Left atrium
Lefty - Left-right determination factor
LHF – Late head-fold stage
LPM – Lateral plate mesoderm
LR – Left-right
LRO – Left-right organizer
LV – Left ventricle

M

M – Molar
MAPCAs – Major Aorto-Pulmonary Collateral Arteries
Mef2c – Myocyte Enhancer Factor 2C
Mesp1 – Mesoderm posterior 1
Min – Minutes
ml – Milliliters
mRNA – Messenger ribonucleic acid

N
NCBI – National Center for Biotechnology
Nkx - Homeodomain factors
NVPs – Nodal vesicular parcels

O
Oct – Octamer-binding transcription factor
OFT – Outflow tract
OMIM – Online Mendelian Inheritance in Man

P
PA – Pulmonary atresia
PAM – Paraxial mesoderm
PBS – Phosphate buffer saline
PCP – Planar cell polarity
PCR – Polymerase chain reaction
PFP – Prospective floor plate
Pitx2 – Pituitary homeobox 2
PK – Proteinase K
Pkd1 – Polycystin-1
Pkd1l1 – Polycystin like 1
Pkd2 – Polycystin-2
PLSVC – Persistent left superior vena cava
PS – Primitive streak

Q
qPCR – Quantitative polymerase chain reaction

R
RA – Retinoic Acid
RA – Right atrium
RNA – Ribonucleic acid
rpm – Rotations per minute
RT – Room temperature
RT-PCR – Reverse transcription PCR
RV – Right ventricle

S
SELI – Self-enhancement lateral inhibition
SeV – Sendai virus
SHF – Second heart field
Shh – Sonic Hedgehog

SI – *Situs inversus*
SMAD – Mothers against decapentaplegic homolog
SNPs – Single nucleotide polymorphisms
Sox2 – Sex determining region Y-box 2

T
T – Thymine
TAPVC – Total anomalous pulmonary venous connection
Tbx – T-box transcription factor
TF – Transcription Factor
TGA – Transposition of great arteries
TGF-β – Transforming growth factor beta
Tm – Temperature of melting

U
U – Unit
UTR – Untranslated region

V
Vangl1 – Van Gogh like 1
VSD – Ventricular septal defects

W
w/v – Weight/volume
Wnt – Wingless-Type MMTV Integration Site Family
WT – Wild-type

Chapter 1

General Introduction

1. The primordium of life

Human prenatal development starts at the time of fertilization and can be divided into three different periods, blastocyst or zygotic period; embryonic period and fetal period (Cochard & Netter, 2002). The zygotic period, which comprises the first two weeks of gestation, is a period of cell proliferation and leads to implantation of the blastocyst into the uterine wall. Through the embryonic period (third to the eighth week of gestation), the gastrulation occurs and most of the important organs and physiological systems develop. At this time, the embryo assumes a human appearance. The fetal period extends from the third month until birth. It is a period of growth, maturation and functionalization of the structures and organs. During these time course periods, the initial fertilized egg undergoes a systematic series of sequential changes to become increasingly complex, differentiated and ultimately giving rise to a fully developed human organism.

1.1. Early embryogenesis

Conception or fertilization is the fusion of a male spermatozoid with a female oocyte. This process, which normally occurs in the upper part of the uterine tube, marks the beginning of the human development and leads to the formation of a genetically unique and specialized cell, the zygote (Moore, Persaud, & Torchia). Immediately after fertilization, this newly formed totipotent cell makes its way to the uterus, a journey that takes five to seven days in humans. As it travels, the zygote undergoes a series of repeated mitotic divisions, called cleavages. Cleavage divisions result in a rapid increase in the number of cells that also become smaller with each division and are at this time called blastomeres. The blastomeres continue to divide and as the developing human enters in the uterus, by approximately three days after fertilization, it undergoes a compaction phenomenon leading to the formation of a 16-cell morula (Carlson, 2014; Moore et al.).

Posteriorly, by day four, a fluid starts to penetrate into the intercellular spaces of the morula and progressively forms a single blastocyst cavity. As the fluid

continues to increase into the blastocyst cavity, it separates the blastomeres and we can differentiate two groups of cells (Moore et al.). At this time, the embryo is a hollow ball of hundreds of cells, called blastocyst, with an internal mass of cells, the inner cell mass (ICM) or embryoblast, at one pole of the blastocyst, and an outer cell layer that forms the epithelial wall of the blastocyst, the trophoblast. The trophoblast will form the extraembryonic structures and the fetal component of the placenta whereas the cells of the ICM are pluripotent stem cells that can give rise to all cell types of the three embryonic germ layers, i.e., Ectoderm, Mesoderm, and Endoderm, and the Germ Cell Lineage, as well as to the non-trophoblast tissues (Yolk sac, Allantois, and Amnion) that support the developing embryo (Carlson, 2014; Moore et al.; Schoenwolf et al.).

Around day six of development, the 150 micrometers sized embryo starts to invade the endometrial epithelium and by the end of the second week of human gestation the blastocyst is completely implanted in the endometrium (Cochard & Netter, 2002; Moore et al.). As implantation proceeds, the trophoblast proliferates quickly and differentiates into the cytotrophoblast, an inner actively proliferating layer, and into the syncytiotrophoblast, which erodes the endometrial tissues, whereas the blastocyst changes morphologically resulting in a flat bilaminar plate of epithelial cells, the embryonic disc. This bilaminar embryonic disc consists of an upper layer of columnar cells, the epiblast, and a lower layer of cuboidal cell, the hypoblast or primitive endoderm. The epiblast contains the pluripotent cells that will form the three germ layers and the hypoblast will give rise to the yolk sac (Carlson, 2014; Moore et al.; Schoenwolf et al.). The formation of the bilaminar embryonic disc marks the appearance of the first axis in the embryo, the primitive dorsal-ventral axis with the epiblast defining the dorsal side and the hypoblast defining the ventral side of the embryo (Carlson, 2014; Moore et al.; Schoenwolf et al.).

At the start of the third week of human development, the cells of the embryonic epiblast begin to differentiate in a process known as gastrulation – formation of the three germ layers. This process, which marks the beginning of morphogenesis (creation of body shape), initiates between days fourteen and sixteen after fertilization and at this stage the embryo is known as gastrula (Schoenwolf et al.). The first morphological sign of gastrulation is the formation

of a longitudinal midline structure, the primitive streak (PS), on the surface of the epiblast, near the caudal end of the bilaminar embryonic disc (Moore et al.). This structure results from the proliferation and movement of the epiblast cells to the median plane of the embryonic disc. As the primitive streak elongates by the addition of cells to its caudal end, its cranial end proliferates to form the primitive node or Hensen's node (Moore et al.). The primitive node contains a circular depression called the primitive pit, which is continuous caudally down the midline of the primitive streak with a trough like depression called the primitive groove (Carlson, 2014). This groove results from the invagination movement (inward movement) of cells from the epiblast through the primitive streak, a process called ingression (Schoenwolf et al.).

Shortly after the primitive streak appears, cells from the epiblast move toward the primitive streak, ingress in this structure and then migrates, in a temporal and spatial well-defined sequence, into the interior of the embryonic disc, to form two new germ layers. During this process, the cells change their structure and organization, undergoing an epithelial-mesenchymal transition (EMT). While in the epiblast they present the typical properties of epithelial cells, i.e. closely connected, well-defined apical and basal surfaces, expression of E-cadherin, when they become free of the epiblastic layer in the primitive groove, they assume the morphology and characteristics of mesenchymal cells (spindle-shaped morphology, expression of N-cadherin instead of E-cadherin), which are able to migrate as individual cells if provided with the proper extracellular environment (Carlson, 2014).

The first epiblast cells to ingress in the primitive streak invade the hypoblast, displacing its cells and replacing them with a layer of definitive endoderm. Others epiblast ingressing cells remain there and organize into the space between the epiblast and definitive endoderm to form a third germ layer – the intraembryonic mesoderm. As soon as the definitive endoderm and intraembryonic mesoderm are formed, epiblast cells no longer migrate toward the PS nor ingress through it. This remaining epiblast cells now constitutes the last germ layer, the ectoderm (Carlson, 2014; Moore et al.; Schoenwolf et al.). The embryonic ectoderm gives rise to peripheral and central nervous systems, the epidermis, hair, nails, sensory epithelia of eyes, ears and nose, many

connective tissues of the head, some glands and neural crest cells, which contribute to the formation of many organic systems (Moore et al.; Schoenwolf et al.). Embryonic endoderm is the source of the epithelial linings of the digestive and respiratory tracts, auditory tube, urinary bladder and most of the urethra. It also contributes to the parenchyma of the liver, pancreas thymus, thyroid and parathyroid glands (Carlson, 2014; Gilbert, 2010; Moore et al.; Schoenwolf et al.). Embryonic mesoderm gives rise to striated and smooth muscles, blood cells, blood vessels lineage, serosal linings of all body cavities, ducts and organs of the reproductive and excretory systems, and most of the cardiovascular system. It is the source of all connective tissues, including cartilage, bones, tendons, ligaments, dermis, and stroma (connective tissue) of internal organs, excluding the head and limbs (Carlson, 2014; Gilbert, 2010; Moore et al.; Schoenwolf et al.). By the end of the third week, the process of gastrulation is completed and the bilaminar embryonic disc was converted into a flat three-layered disc, composed of ectoderm, endoderm and mesoderm that will give rise to all the tissues and organs of the embryo. Furthermore, gastrulation is responsible for the definition of all major body axes through the formation of a temporary structure called the primitive streak. The primitive streak establishes a visible longitudinal axis of bilateral symmetry around which all embryonic structures will organize and align.

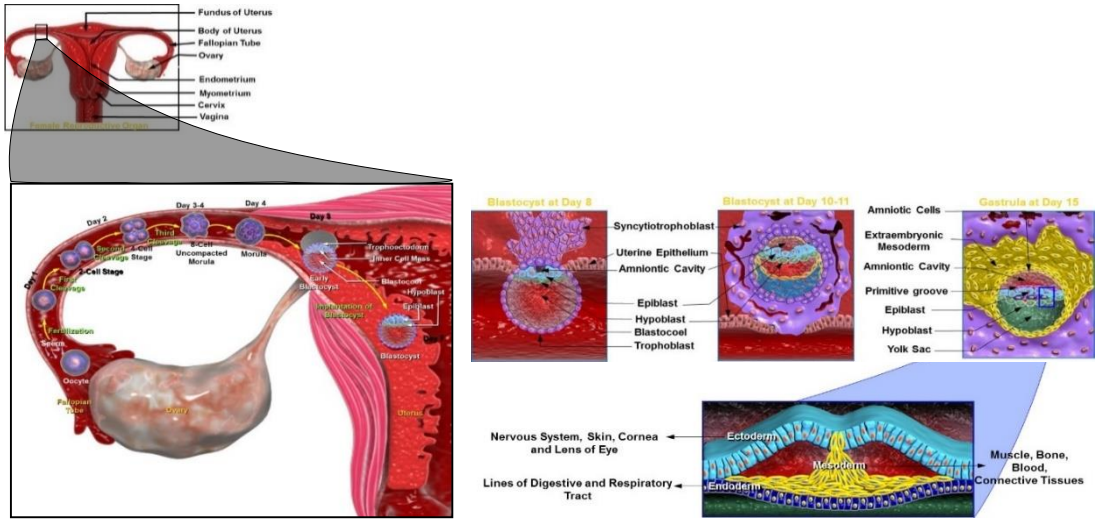


Figure 1.1 – Summary of the first three week of human development.

During the first two weeks of gestation, the initial zygote embraces a period of cell proliferation that culminates with the implantation in the uterine wall and appearance of a bilaminar embryonic disc consisting of the epiblast and hypoblast. During the third week, gastrulation occurs and the bilaminar embryonic disc was converted in a three-layered disc, composed of ectoderm, endoderm and mesoderm.

2. Patterning the embryo – Left-Right Axis Specification

2.1. Anterior-posterior and dorsal-ventral axis

As three-dimensional objects, all vertebrate and invertebrate organisms present three different body axes, the anterior-posterior (AP) axis, the dorso-ventral (DV) axis and the left-right (LR) axis. The formation of these axes however varies from species to species. Whereas in fish and frog the dorso-ventral and anterior-posterior axes are determined at fertilization by the distribution of the yolk and the entry position of the sperm (Gilbert, 2010), in mouse and human, when the embryo implants into the wall of the uterus it is still symmetric without any axe. In these embryos, as embryogenesis proceeds, the DV axis is specified as the proximal-distal axis from the implantation site. Subsequently, the AP axis is arbitrarily determined in the plane perpendicular to the DV axis (Alarcon & Marikawa, 2003; Beddington & Robertson, 1999). So, after implantation, the future AP axis is marked anteriorly by the polarization of the head primordium and posteriorly (tail) by the appearance of the primitive streak, while the prospective DV axis is defined by the formation of the three germ layers, the endoderm marking its ventral side and the ectoderm being dorsal (Hamada & Tam, 2014; Hirokawa et al., 2009).

2.2. Left – right asymmetry

Theoretically, in a three-dimensional structure, once the orientation of two orthogonal axes is acquired, the orientation of a third axis would be spatially delineated by default (Hamada & Tam, 2014). This axiom of geometry seems to fit to the formation of the three body axes.

During development, once the AP and DV axes are established, it is argued that the LR axis is consistently oriented relative to the dorsal–ventral and anterior–posterior axes (Hashimoto & Hamada, 2010). Although many insights had been made regarding the generation of the anterior-posterior and dorsal-ventral axes, much remains to be understood about how the LR axis is consistently oriented with respect to the other two and what are the molecular mechanisms

underlying the process of LR asymmetry generation (Vandenberg & Levin, 2012). In the past 25 years, this process has been extensively studied and many genes and pathways were uncovered. Nonetheless, the precise mechanisms have not been determined making difficult to conceptualize therapeutic or preventative approaches (Amack, 2014).

In the next pages, I will provide a review about the current knowledge of how left-right symmetry is established in animal models. Resuming the agreements and disagreements in the field and the proposed models to explain how by breaking the bilateral symmetry and creating LR differences in the body plan, the initial symmetric embryo is transformed into a complex organism, externally bilateral symmetric and internal asymmetric.

2.2.1. The three main steps

The establishment of left-right (LR) asymmetry during early embryogenesis is critical for the organization of the embryo body plan and for the correct positioning and morphogenesis of the internal organs (Shiratori & Hamada, 2014). There is a general agreement in the field that the process by which this LR asymmetry is generated requires and can be divided in three different main phases/steps (Figure 1.2): In the first step, there is a breaking of the initial bilateral symmetry, allowing that the LR axis is consistently oriented relative to the dorsal-ventral and anterior-posterior axes. The second step consists in the transference of an LR-biased signal (or signals) to the lateral plate mesoderm (LPM), which leads to differential gene expression of signaling molecules in cell fields on either side of the midline. In the third step, this asymmetric gene expression of signaling molecules drives changes in cell behavior, such as proliferation and migration rate, which lead to the LR asymmetric position and morphogenesis of the internal organs (Shiratori & Hamada, 2006, 2014; Vandenberg et al., 2013; Vandenberg & Levin, 2013). In the last years, this process has been extensively studied to decipher the molecular mechanisms underlying these three steps of left-right asymmetry establishment. These studies revealed a general agreement between researchers regarding the main molecular pathways involved in steps two and three, decoding that these

cascades are highly conserved from sea urchin to mouse (Nakamura & Hamada, 2012). In contrast, the origin of the symmetry-breaking event (first step) is controversial because no unifying mechanism between model organisms has been identified to date and it seems that this process has diverged during evolution (Nakamura & Hamada, 2012; Sauer & Klar, 2012).

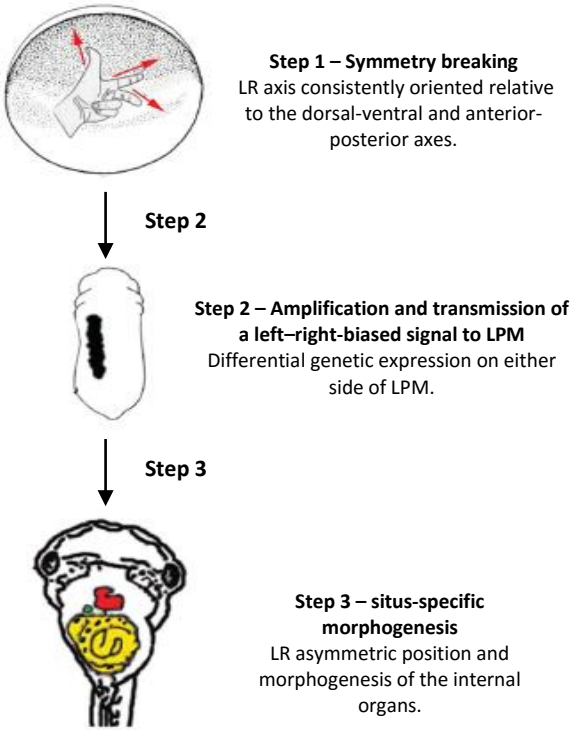


Figure 1.2 – General process of LR asymmetry generation.

Generation of LR asymmetry comprises 3 steps. In the first one, LR axis is oriented relative to the DV and AP axes due to the action of an asymmetric signal. The second steps involve the transference of this signal to the LPM and differential genetic expression on either side of LPM. The last step relay on the LR asymmetric position and morphogenesis of the internal organs.
(Adapted from: (Vandenberg, Lemire, & Levin, 2013))

2.2.1.1. Models for symmetry breaking (first step)

The mechanism beyond the process of symmetry breaking was put in debate with the conceptual F-molecule model (Figure 1.3) proposed by Brown and Wolpert in the nineties (Brown & Wolpert, 1990). In their analysis, they postulated that a symmetric embryo must distinguish is left from its right side after the establishment of the anterior-posterior and dorsal-ventral axis, by consistent orientation of an intracellular component that is chiral by virtue of its biochemical structure (Vandenberg & Levin, 2013). This “F-molecule” would

have three arms and when two arms are aligned with the preexisting axis, the third arm would be automatically aligned along the future LR axis (Brown & Wolpert, 1990; Hashimoto & Hamada, 2010; Vandenberg & Levin, 2013). Therefore, in the field, there is an agreement about the need of an "F" molecule for the origin of asymmetry prior asymmetric transcriptional events. However, there is considerable debate on some major issues: what is the chiral element that first breaks symmetry? When during embryogenesis is asymmetry initiated? How conserved across diverse species are these mechanisms, and which model systems best represent the "general case"?(Vandenberg et al., 2013; Vandenberg & Levin, 2013). In recent years, two models, which are cytoskeleton-dependent and whose mechanism of action fulfill the chiral "F" molecule hypothesis proposed by Brown and Wolpert, were identified and debated. One model predicts that unidirectional cilia-driven extracellular fluid flow (nodal flow), produced on the LR organizer, during gastrulation, is the origin of LR asymmetry, whereas the second model – intracellular model – postulates that LR asymmetry originates during very early developmental stages due to the chiral nature of the cytoskeleton (Levin, 2003, 2004; Vandenberg et al., 2013; Vandenberg & Levin, 2013). Although both mechanisms implicate the cytoskeleton, as the cilium is effectively a microtubular extension of the cytoskeleton, and lead to asymmetric *Nodal* activity (May-Simera & Kelley, 2012), they act at two very different developmental stages and seem to have less in common. Here, I will introduce only the first model.

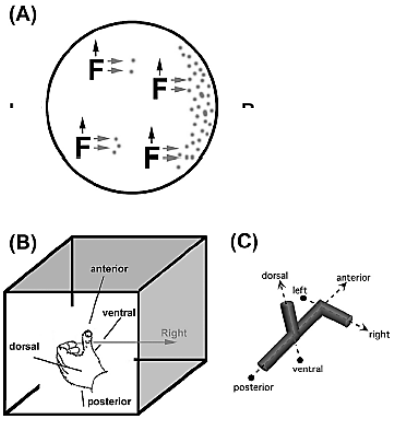


Figure 1.3 – "F" molecule model.

(A) The chirality of a molecule (schematized by an F) tethered with respect to the other axes can orient the direction of the LR axis. (B, C) Three-dimension representation of how an "F" molecule allows each cell to know which direction is left and which is right.
(Adapted from: (Levin, 2004))

2.2.1.1.1. The leftward flow model

In some mammals (mouse, rabbit and humans), fish and amphibians the LR symmetry seems to be broken by the action of rotating cilia in a pit like structure called the “embryonic node” in mouse, “Kupffer’s vesicle’’ in zebrafish, and ‘‘gastrocoel roof plate’’ in frogs (Shiratori & Hamada, 2006; Yoshiba & Hamada, 2014; S. Yuan, Zhao, Brueckner, & Sun, 2015). The mouse node (Figure 1.4A), formed during gastrulation, is a transient small triangular structure with 50–100 μm in width and 10–20 μm in depth, located between the anterior notochord and the primitive streak in the ventral midline of the embryo (Hamada & Tam, 2014; Hirokawa et al., 2009). There are two types of cells in this cavity, the node pit cells, which are columnar epithelial cells located in the central region of the ventral node and the crown cells, squamous epithelial cells located on the edge of the node (Yoshiba & Hamada, 2014). Although both cell types present monocilia that project into the extraembryonic space, the crown cells have in most of the cases (90%) immotile cilia that are believed to sense the nodal flow, whereas, most of pit cells possess motile rotating cilia which are responsible for the generation of the asymmetric left-directed flow of extracellular fluid (nodal flow) in the node cavity (Figure 1.4B) (Hirokawa, Tanaka, & Okada, 2012).

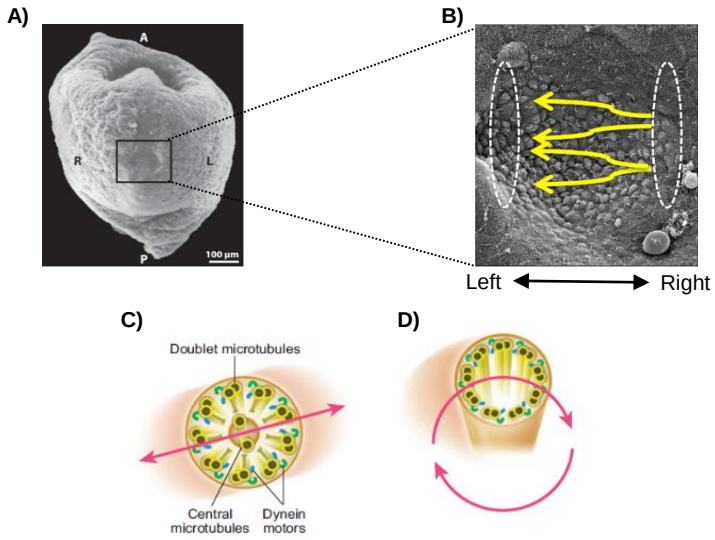


Figure 1.4 – Node structure and cilia microtubules organization.

A) Scanning electron microscope of a mouse node at 7.5 days postcoitum. **B)** Direction of the leftward fluid flow across the ventral node. Motile cilia are located at the central region of the node, while immotile cilia are present at the periphery of the node cavity (marked by dashed lines). **C)** Normal cilia with nine pairs of 9+2 arrangement microtubules connected with dynein motors. **D**) The central pair of microtubules is missing in immotile primary cilia and nodal cilia. In nodal cilia, the dynein motors remain in a chiral arrangement and produce a rotation-like movement (arrows). (Adapted from: (Hamada & Tam, 2014; Hirokawa, Tanaka, & Okada, 2009))

2.2.1.1.1.1. Cilia and the generation of the unidirectional leftward flow

Normal motile ciliated cells present an axoneme composed of nine pairs of double microtubules arranged longitudinally along the axes and a pair of microtubules, referred as central pair, in the center of the cilia (9+2 arrangement) (Figure 1.4C) (Hirokawa et al., 2009). This central pair of microtubules, which defines the direction of the beating plane, is absent in the axoneme of immotile cilia cells and nodal cells. Thus, based on the microtubule arrangement (9+0) (Figure 1.4D), the node monociliated cells were thought to lack motility (Hirokawa et al., 2009). However, in 1998, Nonoka and co-workers examined the behavior of these monocilia using video-light microscopy. They discovered, for the first time, that the monocilia presented in the mouse node are motile but instead of move back and forth, like most motile cilia or flagella, they rotate vigorously at approximately 600rpm (Nonaka et al., 1998). Furthermore, when they added fluorescent beads in the extraembryonic fluid at the node region of wild-type embryos, they verify that the cilia of these cells rotate in a clockwise direction, generating a leftward (right-to-left) flow of fluid in the node cavity, which they designated nodal flow (Takeda et al., 1999) (Figure 1.5A). Okada also supported this hypothesis in 1999, by the analysis of node monocilia motility of *inversus viscerum* (*iv*) mutant mice, which harbors a mutation in the axonemal dynein protein *Dnah11 (Lrd)* gene. These results confirmed that the monocilia of heterozygous *iv* mutant embryos also rotates as fast as the wild-type embryo (~600rpm) and this movement can also generate a leftward flow of extraembryonic fluid in the node cavity. In contrast, in the *iv* mutant homozygous mice, the cilia are almost immotile and the flow absent (Okada et al., 1999), suggesting that in fact cilia are the responsible for the generation of the nodal flow.

But how and why is the flow generated in a leftward direction by the rotational movement of the motile monocilia? The answers to these questions rely in the posterior position of the basal body of each cilium, which lead to a posterior tilting of the cilia in the node cells. High-speed video microscopy observation revealed that the node cilia do not project vertically from the node cells, but they

are tilted posteriorly at an average angle of 30°- 40º (Nonaka et al., 2005), which allows the cilia to orient along the anterior-posterior and dorsal-ventral axes, as the F-molecule model proposed by Brown and Wolpert. These observations also revealed that a rightward rotational movement of the cilium occurs close to the cell surface, whereas a leftward motion is away from the surface (Hirokawa et al., 2009). Based on this, hydrodynamic principles predict that the shear resistance of the cell surface slows the rightward swing, but leaves the leftward stroke unimpeded to generate a leftward fluid flow, leading to a more efficient sweep of the cilia towards the left as opposed to the right (Okada, Takeda, Tanaka, Izpisua Belmonte, & Hirokawa, 2005). However, these observations do not explain the posterior tilting of the cilia. This tilting could be explained by the convex curvature of the apical membrane of the node cells and the shifting in the position of the basal body (Figure 1.5B), which seems to be related with planar cell polarity (PCP). Indeed, some core PCP components, like Dishevelled (Dvl), Van Gogh like 1 (Vangl1) and Prickle2 have polarized subcellular localizations (Figure 1.5B) in the node cells and knockout mice for these components result in disrupted posterior position of nodal cilia, abnormal nodal flow and LR defects (Hirokawa et al., 2012; Yoshiba & Hamada, 2014). Thus, the posterior positioning of the basal body, ensued from a mechanism of PCP, results in the posterior tilting of the cilia, which in turn is responsible for the generation of a unidirectional leftward flow that disrupts the initial LR symmetry in the mouse embryo.

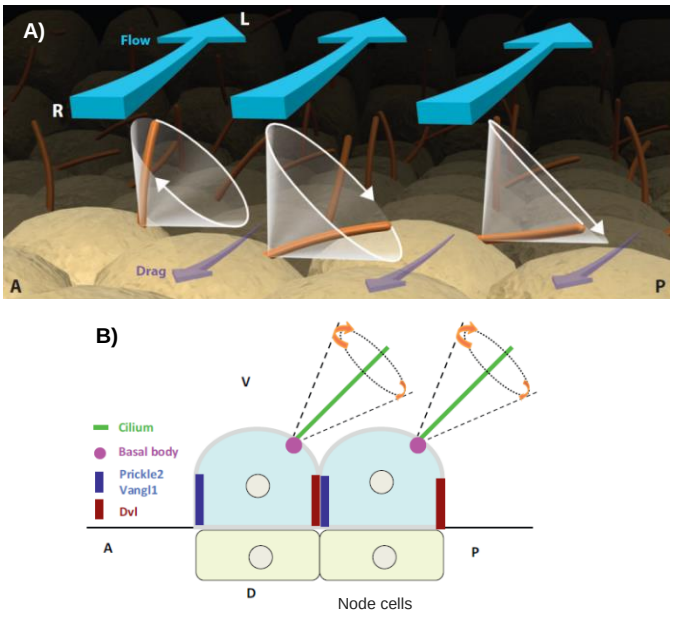


Figure 1.5 – Leftward nodal flow in the ventral node and posterior tilting of cilia.

A) Unidirectional leftward flow (blue) is generated by the clockwise rotation of nodal cilia due to the posteriorly tilting of these cilia. **B)** Posterior tilting of cilia result from the posterior position of the basal body, which seems to be related with the subcellular localization of some core planar cell polarity molecules, Dishevelled, Van Gogh like 1 and Prickle2. Adapted from: (Yoshiba & Hamada, 2014).

2.2.1.1.1.2. Sensing the flow – morphogen/ chemosensor model

It is widely accepted that while central cilia of pit cells are mostly motile and generate the nodal flow, peripheral cilia of crown cells are mostly immotile and act as sensors of the fluid flow. However, how these cells at the edge of the node sense the flow is still debated. Two main model hypotheses have evolved to answer to this question. One hypothesis, termed “morphogen gradient” or chemosensor model, was firstly proposed by Nonaka and colleagues in 1988 (Figure 1.6A) (Nonaka et al., 1998). This model assumes that signaling morphogens, with sizes between 20-40 kDa, like Sonic Hedgehog (Shh), Fibroblast growth factor (FGF), Retinoic Acid (Ra) or Nodal are secreted into the node cavity and transported, due to the action of the unidirectional nodal flow, towards the left side of the embryo, where they bind to cell surface receptors and initiate a signaling cascade (Hirokawa, Tanaka, Okada, & Takeda, 2006). This stationary asymmetric accumulation of a determinant morphogen triggers signaling events that cement the asymmetry in the developing embryo. Although simple and straightforward, this hypothesis is questionable whether nodal flow can generate a chemical gradient in a closed cavity of node, where the leftward flow is balanced by the rightward counterflow (Hirokawa et al., 2006). Furthermore, what is the identity of the morphogen itself?

In a more sophisticated version of the chemical model (Figure 1.6B), the authors observed flowing materials, named nodal vesicular parcels (NVPs) (Hirokawa et al., 2012) that are released from slowly growing filopodia-like processes in all regions of the ventral node into the extracellular fluid flow, fragmented by contacting rotational cilia and finally absorbed by left-sided node cells (Hirokawa et al., 2012). Furthermore, pharmacological studies suggest that FGF, Shh and RA regulate the release of NVPs and that NVPs seem to carry

the signal to elevate the calcium (Ca^{2+}) concentration on the left side of the node, which in turn is thought to influence the asymmetric gene expression in adjacent cells (Hirokawa et al., 2006). Indeed, Shh and RA are unsheathed and transported in the NVPs but they are unlikely to be the molecular determinant of LR polarity because asymmetric expression of downstream Shh or retinoic acid response genes has not been detected in or near the mouse node.

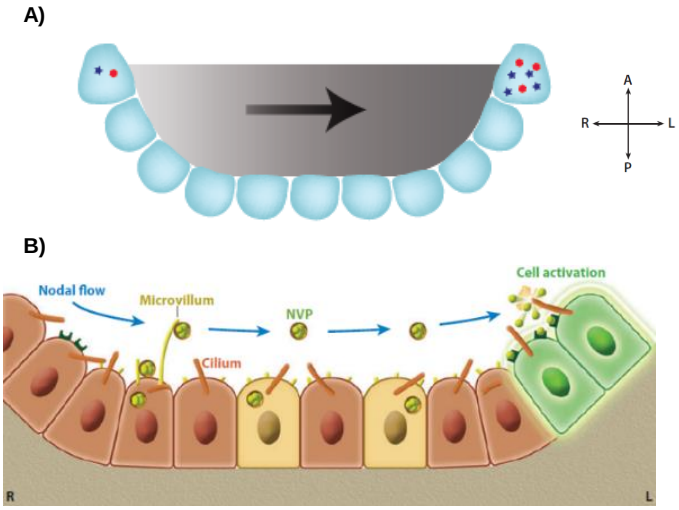


Figure 1.6 – The two versions of the "morphogen gradient"/chemosensor model.

A) First "morphogen gradient" model proposed by Okada and colleagues. In this model, a morphogen secreted into the node becomes asymmetrically accumulated on the left side due to the action of the nodal flow. Resulting thereby in symmetry breaking. **B)** Recent version of the "morphogen gradient" model propose that nodal vesicular parcels (NVPs) containing a morphogen are carried to the left side of the node by nodal the flow. These NVPs are then fragmented and absolved by left-sided node cells, resulting therefore in the breaking of symmetry. Adapted from: (Norris, 2012).

2.2.1.1.1.3. Sensing the flow – Two cilia/mechanosensor model

An alternative to the chemosensor hypothesis is that the hydrodynamic force generated by the flow is mechanosensed in the responding cells, most likely the crown cells, in the peripheral region of the ventral node (McGrath, Somlo, Makova, Tian, & Brueckner, 2003). This model, named "two-cilia" model (Figure 1.7), postulates that motile cilia at the center of the node produce an asymmetric left-directed flow of extracellular fluids detected by mechanical bending of a second non-motile group of cilia, which contain the Ca^{2+} permeable cation channel Polycystin-2 (Pkd2) exclusively on the left edge of the

node. This results in the elevation of intracellular Ca^{2+} and leads to subsequent asymmetric patterns of gene expression (Hamada & Tam, 2014; McGrath et al., 2003; Yoshiba & Hamada, 2014).

Some evidences are in agreement with this idea, the first comes from kidney cells in which immotile cilia respond to flow-induced bending mechanical stress by initiating an influx of extracellular Ca^{2+} through mechanosensory complex channels (Pkd1/Pkd2 (Praetorius & Spring, 2001). Surprisingly, nodal cilia cells (motile and immotile) do not express *Pkd1* and in its absence LR asymmetry is not disrupted (Karcher et al., 2005), suggesting that in nodal cells, Pkd2, which is thought not to be a sensory protein per se, will form a sensory complex with another partner. In fact, in mouse and fish cilia, Pkd2 physically interact and form a complex with a paralogue of Pkd1, Pkd1l1. Moreover, *PKd1l1* is expressed in a pattern that corresponds spatially and temporally to the establishment of LR asymmetry and both proteins localize in motile and immotile cilia of the ventral node (Field et al., 2011; Kamura et al., 2011). So, in nodal cells, the same process of sensing the flow may occur via the flow-sensing complex Pkd1l1/Pkd2, being Pkd1l1 the sensor and Pkd2 the effector. Second, asymmetric elevation of intracellular Ca^{2+} became randomized in *iv* mice (flow absent) and absent when Pkd2 was deleted (Babu & Roy, 2013). Moreover, rescues assays of the *Pkd2* gene in specific regions of the *Pkd2* deficient mice revealed that *Pkd2* expression is not necessary in pit cells, composed mostly of motile cilia, for correct LR development. However, correct LR development requires *Pkd2* expression specifically in crown cells, where most immotile cilia are located, suggesting that Pkd2 function is needed in the immotile cilia to sense and respond to the flow (Norris & Grimes, 2012). Lastly, an elegant work by Hamada's lab in Japan has provided proof of the sensory function of cilia in Nodal cascade induction. First, the authors verified that only two rotating cilia were required to establish normal sidedness. Second, they verified that artificially applied fluid flow was able to rescue the Nodal cascade in a mouse model of Kartagener syndrome, i.e. in embryos without flow due to immotile cilia, in a mechanism dependent on the calcium channel Pkd2.

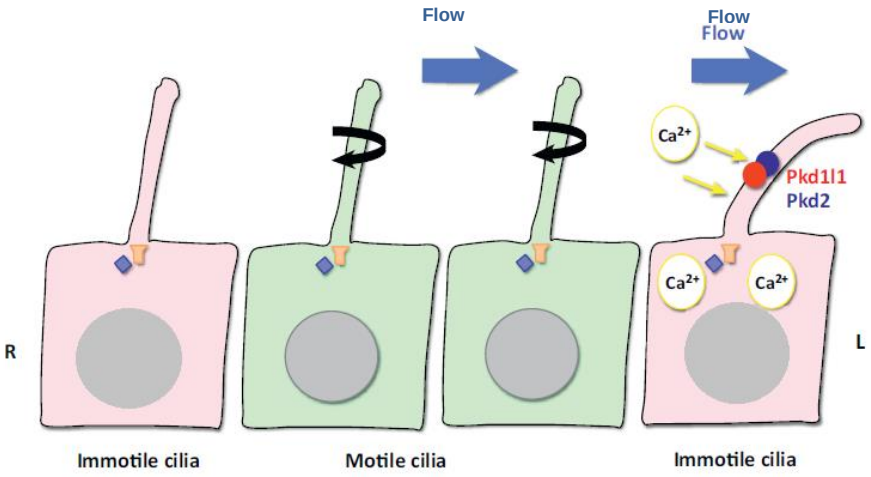


Figure 1.7 – "Two cilia" model.

Motile cilia at the center of the node generate a leftward nodal flow that is detected by mechanical bending of immotile cilia cells in the left edge of the node. These cells contain the Ca^{2+} permeable channel Pkd2 and Pkd1l1 exclusively on the left edge of the node resulting in the elevation of intra cellular Ca^{2+} and subsequent asymmetric patterns of gene expression. Adapted from: (Yoshiba & Hamada, 2014).

Furthermore, rescue of *Kinesin family member 3A* (*Kif3a*) expression only in the crown cells of these animals, results in the generation of a weak leftward flow and *Nodal* expression in the LPM. These observations suggest that a weak and transient local flow produced by the few motile cilia crown cells and sensed by the neighboring immotile cilia cells is sufficient to initiate LR asymmetric gene expression.

Although all the evidences mentioned above support the two cilia hypothesis, the bending of sensory cilia has not been verified by in *in vivo* observations and all the cilia in Kupffer's vesicle of Medaka fish were directed visualized as motile cilia, expressing *Lrd*, *Pkd1l1* and *Pkd2* (Babu & Roy, 2013). Furthermore, inhibition of FGF signaling, which is involved in the release of NVP in the morphogen hypothesis, suppresses Ca^{2+} asymmetric intracellular elevation without disturbing the flow, suggesting that FGF signaling and NVPs are needed for the asymmetric distribution and consequent normal LR patterning (Tanaka, Okada, & Hirokawa, 2005). Therefore, these two mechanisms may not be mutually exclusive, but rather may run in parallel or work synergistically with each other.

2.2.1.1.1.4. Readouts of the nodal flow – the role of Cerl2/DAND5

Although the precise mechanism, chemical or mechanical, of how crown cells at the edge of the node senses the flow remains controversial, two main readouts of the flow on the crown cells of the left margin of the node have been reported. These are the elevation of intracellular Ca^{2+}, most likely via Pkd1l1/Pkd2 complex channels, and the elevation of the Nodal activity (Hirokawa et al., 2012).

Nodal activity/signaling (Figure 1.8) is triggered when the transforming growth factor beta TGF-β family ligand Nodal binds to a target cell through interaction with type I (ALK4 and ALK7) and type II (ActRIIa and ActRIIb) serine-threonine kinase receptors in the presence of one EGF-CFC (epidermal growth factor-Cripto-FRL1-Cryptic) family protein co-receptor, Cripto or Cryptic. This leads to the phosphorylation of regulatory Smads (Smad2 and Smad3), their association with Smad4 and translocation of this complex into the nucleus. In the nucleus this complex interact with the transcription factor FoxH1, which is the main responsible for transducing the Nodal signaling and activating downstream target genes like Nodal itself (auto regulation), Lefty2 and Pitx2 [15, 54]. Nodal signaling at the level of the node is inhibited by the Cerl2/Dand5 molecule, which asymmetric expression represents the most immediate molecular response to the flow in the crown cells of this embryonic left-right organizer (Hamada & Tam, 2014). DAND5, also known as Charon in fish, Coco in *Xenopus* and Cerl2 in mouse, belongs to the Cerberus/Dan family, encodes a 20-kDa secreted growth factor with the ability to bind directly to Nodal and to inhibit its signaling pathway (Marques et al., 2004). The level of *Cerl2* mRNA in crown cells is initially bilaterally equal (L=R) at early head-fold (EHF) stage (embryonic day 7.5), when the flow is very weak or not apparent. However, as the velocity of nodal flow increases by late head-fold stage (LHF) (E7.75), the expression of *Cerl2* in the left side began to decline in intensity (downregulated), giving rise to an L<R asymmetric expression pattern among the perinodal cells from 1-somite stage (E8.0) (Figure 1.9A) (Inacio et al., 2013; Kawasumi et al., 2011).

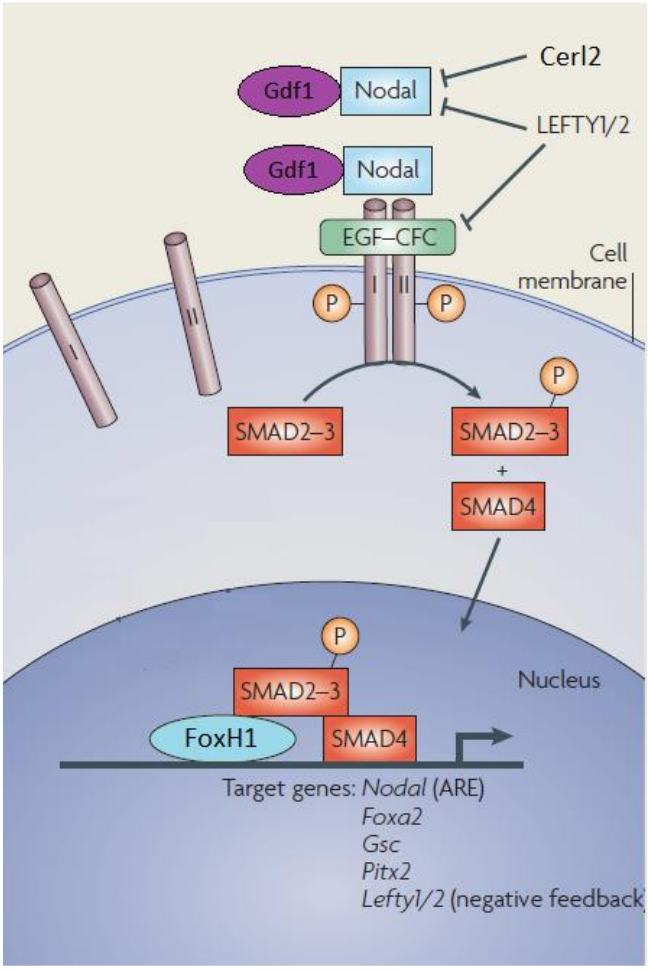


Figure 1.8 – Nodal Signaling.

In the absence of Lefty and Cerl2, Nodal-GDF1 bind type I and type II receptor in conjunction with the co-receptor, EGF-CFC protein. This leads to cross-phosphorylation of receptor I and II and subsequently to phosphorylation of a receptor-regulated SMAD (Smad2/3), allowing this protein to associate with Smad4. This complex moves into nucleus, associates with a DNA-binding partner (FoxH1), binds to specific enhancers in targets genes, and activates their transcription. Adapted from: (Arnold & Robertson, 2009)

This L<R asymmetry of *Cerl2* mRNA is regulated at post-transcriptional level by the degradation of *Cerl2* mRNA via its 3'UTR (untranslated region) in response to the leftward flow (Nakamura et al., 2012). This initial L<R bias in *Cerl2* expression is subsequently amplified by the Wnt-Cerl2 feedback loop, in which Wnt3 up-regulates *Wnt3* expression and promotes *Cerl2* mRNA decay whereas Cerl2 promotes Wnt3 degradation (Nakamura et al., 2012).

Furthermore, a recent study by Sun Brueckner and co-workers, in 2015, revealed another possible important mechanism involved in the left downregulation of *Dand5/Cerl2*. In this study, the authors recorded highly dynamic intraciliary calcium oscillations (ICOs), which coincide with the onset of

cilia motility, in response to leftward flow, predominantly on the left side of zebrafish Kupffer's vesicle (S. Yuan et al., 2015). These oscillations were dependent on Pkd2 and cilia motility, and were followed by cytoplasmic calcium asymmetries in the same left-right organizer (LRO) and in neighboring cells, i.e. en route to the lateral plate mesoderm, the site of Nodal cascade induction. Most interestingly, when these ICOs were suppressed, *Dand5* was not properly downregulated on the left side of the LRO in response to ciliary motility. Therefore, the authors proposed that flow and Pkd2-dependent intraciliary calcium oscillations control *Dand5* asymmetry at the top of the left–right signaling cascade and the mechanism of *Dand5* mRNA degradation might be regulated in a calcium-dependent manner (S. Yuan et al., 2015).

While the *Cerl2* expression in the perinodal region (crown cells) changes from symmetric to asymmetric (L<R) , the mRNA levels of *Nodal* and *Gdf1*, which is co-expressed and forms a heterodimer with Nodal, functioning as a stabilizer of the mature Nodal protein, remain distributed symmetrically in this region from EHF stage to 1-somite stage (Figure 1.9B) (Kawasumi et al., 2011; Shiratori & Hamada, 2014; Yoshiba & Hamada, 2014). Moreover, pSmad2, an indicator of Nodal signaling activity, was already detected asymmetrically (L>R) on the left side of the node at LHF stage (Figure 1.9C) (Kawasumi et al., 2011).

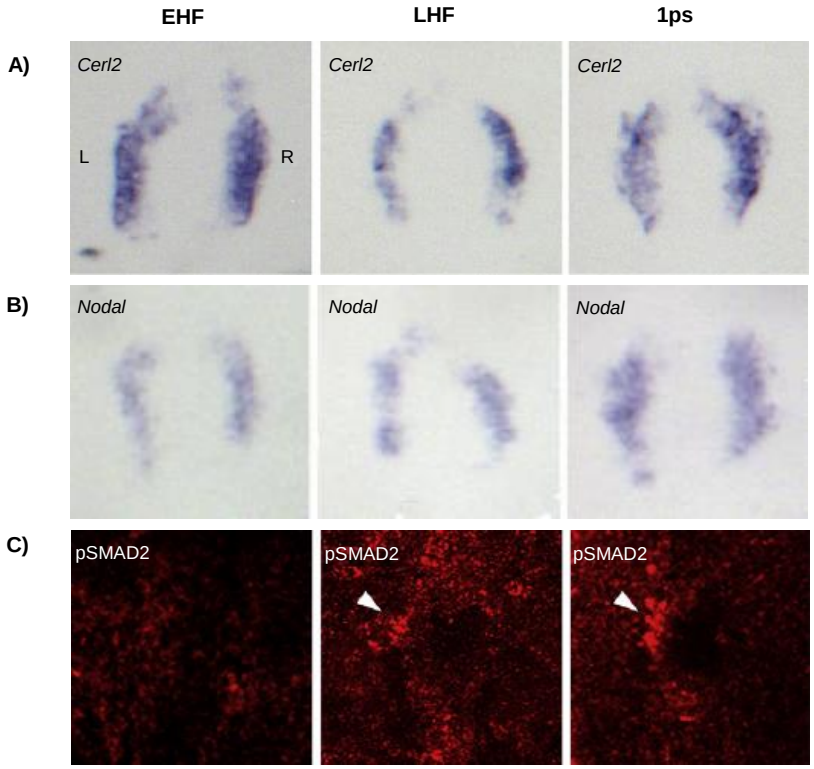


Figure 1.9 – *Cerl2*, *Nodal* and pSMAD2 activity in perinodal cells of the mouse node.

A) *Cerl2* is initial equal expressed in both side of the perinodal cells by EHF stage. At LHF *Cerl2* mRNA starts to be detected asymmetrically (L<R) and by 1-somite stage *Cerl2* is clearly more expressed in the right side of the node. **B)** At these stages, however, *Nodal* expression remain symmetrically distributed in the perinodal cells. **C)** On the other and, pSmad2 activity, a readout of the Nodal signaling, was not detected at the EHF stage but was apparent on the left side of the node at the LHF stage and is clearly evident in the same cells by 1-somite stage. Adapted from: (Kawasumi et al., 2011).

Altogether, these results suggest that *Cerl2* is the major target of the flow and is not the asymmetric *Nodal* expression, but the L<R expression of *Cerl2* the responsible for the pSmad2/3 asymmetry in perinodal cells. So, it was proposed that L<R asymmetric expression of *Cerl2* is responsible for L>R asymmetric Nodal activity in the perinodal cells and thus the role of Cerl2 is to restrict the Nodal activity to the left side of the node, preventing its signaling in the right side.

Although the L<R expression of *Cerl2* resembles the expression pattern of Cerl2 protein until 2-somite stage, the same does not happen from 3 to 6-somite stage (Figure 1.10), suggesting that Cerl2 protein could have another role in addition to the restriction of the Nodal activity to the left side of the node. In fact, at early stages (EHF to LHF stage), Cerl2 protein is expressed in both sides of crowns node cells, in a horseshoe shape, like its mRNA (Inacio et al., 2013). Later, when *Cerl2* expression starts to become L<R, Cerl2 protein begin to accumulate on the right side of the node until 2-somite stage, preventing the activation of Nodal cascade on the right-LPM. By 2/3 somite stage, when *Cerl2* expression remain L<R, Cerl2 protein is translocated to the left side of the node in a nodal flow depend way, suggesting that Cerl2 protein, also due to its small 20-kDa size (fitting the size of a determinant morphogen), may be transported by the flow (Inacio et al., 2013). At 5-somite stage, Cerl2 protein is accumulated on the left side of the node until 6-somite stage, when its signal starts to disappear (Figure 1.10). This dynamic right-to-left translocation of Cerl2 protein, after the asymmetric Nodal expression starts in left LPM, seems to shutdown Nodal activity in the left side of the node at the 6-somite stage, which coincides with the termination of the *Nodal* expression and activity in the left LPM.

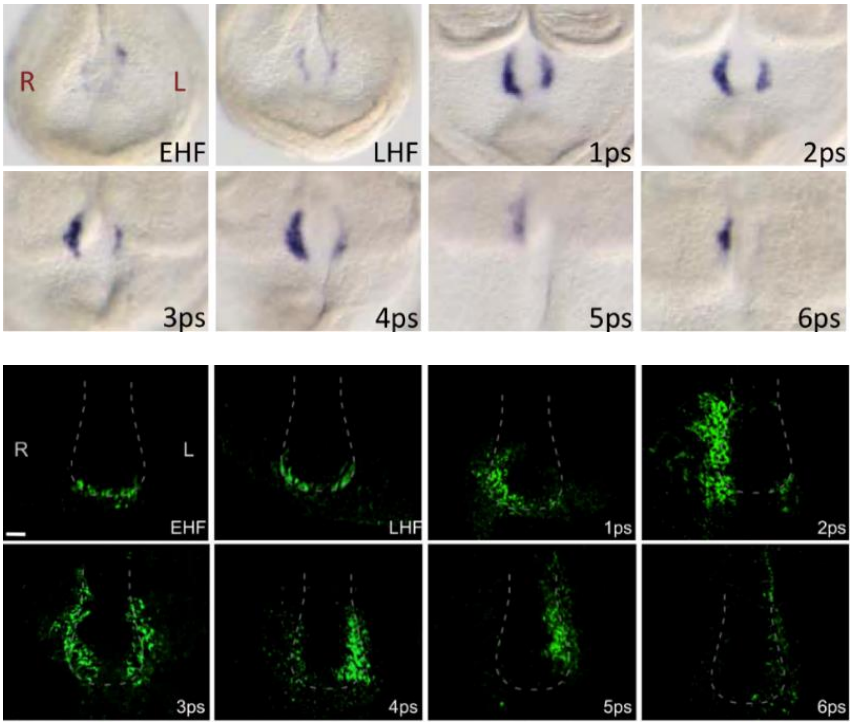


Figure 1.10 – *Cerl2* mRNA and protein localization.

From EHF to 2-somite *Cerl2* expression localizes with the distribution of Cerl2 protein in both sides of the mouse node. However, from 3-somite stage to 6-somite stage, this pattern disappears and Cerl2 mRNA do not localize with Cerl2 protein distribution. At 3-somite stage, when Cerl2 expression remain L<R, Cerl2 protein present equally in both sides of the node and by 4/5-somite stage Cerl2 is accumulated in the left side of the node, contrary to its expression. At 6-somite stage *Cerl2* mRNA and protein start to disappear in opposites sites in the edge of the node. Adapted from: (Inacio et al., 2013).

Therefore, Cerl2/Dand5 is a key molecule in the establishment of the left-right asymmetry, being the only molecule that summons the two main hypothesis mechanisms of flow action (mechanosensor and chemosensor).

2.2.1.2. Transmission of a left–right-biased signal to LPM (second step)

2.2.1.2.1. The Signal

Once the orientation of LR asymmetry has been determined relative to the AP and DV axes it is essential that this asymmetric signal (or signals), generated at the node, must be transferred to the LPM, where it induces the asymmetric expression of *Nodal* and consequent its signaling cascade in order to execute LR asymmetric organogenesis. If the signal is transferred from the node to the left LPM, it must be produced at the node and its expression will need to be

critical for any kind of patterning of LPM, correct or abnormal. Among the several molecules expressed at the node, recent evidence suggests that an active form of Nodal, probably the Nodal-Gdf1 heterodimer protein, in which GDF1 functions as a stabilizer of the Nodal protein (Kawasumi et al., 2011; Shiratori & Hamada, 2014; Yoshiba & Hamada, 2014), is the best candidate signal that is transmitted from the node to the left lateral plate mesoderm. 1 – Specific ablation of *Nodal* expression in the perinodal crown cells prevents *Nodal* expression in either side of LPM. Since *Nodal* is expressed bilaterally at the node (in perinodal crown cells) before the onset of its expression in the left LPM, this suggests that *Nodal* expression in the node are absolutely critical for *Nodal* expression in the lateral plate (Brennan, Norris, & Robertson, 2002). 2 – Asymmetric *Nodal* expression in the LPM is conferred by two Nodal-responsive enhancers (ASEs) and can be induced by Nodal signaling (Adachi et al., 1999; Norris & Robertson, 1999; Saijoh et al., 2005). 3 – Introduction of a Nodal expression vector in the right LPM can induce ectopic endogenous *Nodal* expression on the right side. 4 – Gdf1 is co-expressed with Nodal bilaterally in the perinodal crown cells, and Gdf1 mutant is maintained Nodal expression at the node but it is completely lost in the lateral plate mesoderm [15, 23].

2.2.1.2.2. The Route of the signal to LPM

The transference of the left-right asymmetric signal (or signals) generated at the node to the LPM is still a controversial issue. It has been suggested that if Nodal-Gdf1 heterodimer protein is indeed transported to the left LPM, it can follow at least two transportation routes. First, Nodal-Gdf1 heterodimer protein may be secreted into the node cavity, transported toward the left side of the node by the leftward flow and incorporated, via crown cells, into endodermal cells located at the surface of the embryo. The Nodal signal may then be relayed, across the endodermal cell layer, to the left LPM, where it induces endogenous *Nodal* expression (extraembryonic route) (Figure 1.11B) (Hamada & Tam, 2014; Shiratori & Hamada, 2014; Yoshiba & Hamada, 2014). Alternatively, the Nodal-Gdf1 heterodimer protein may be secreted within the embryo and directly transported along extracellular matrix (ECM) to the left LPM (intra-embryonic route) (Figure 1.11B) (Hamada & Tam, 2014; Shiratori &

Hamada, 2014; Yoshiba & Hamada, 2014), where it induces the asymmetric expression of its own signaling cascade in order to form asymmetric organs. While the extraembryonic route is unlikely, mostly because exposure of endoderm to recombinant Nodal protein does not induce *Nodal* expression in any side of LPM or non-LPM tissues (Oki et al., 2007), some evidences support the intra-embryonic route.

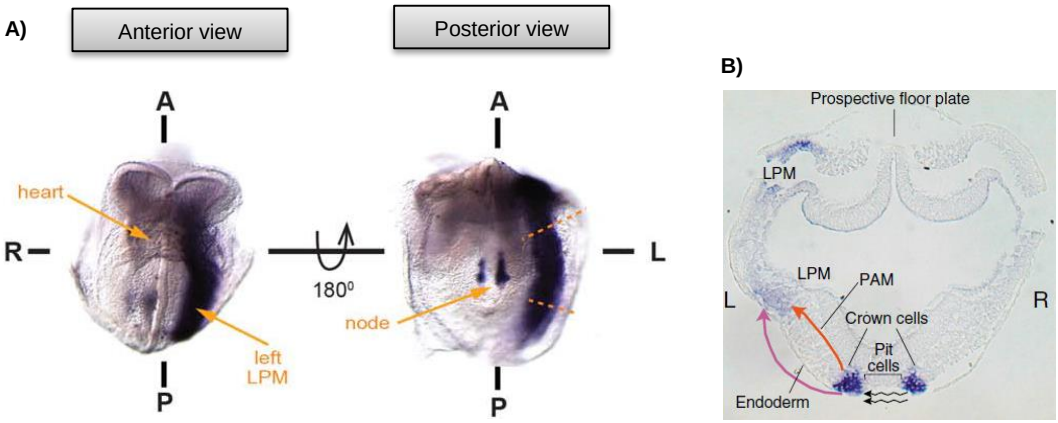


Figure 1.11 – *Nodal* expression and the route of the signal to LPM.

A) Anterior and posterior view of a mouse embryo at ~E8.25 showing Nodal expression in the node and left LPM. **B)** Possible routes of how Nodal are transferred from node to LPM. Asymmetric signal(s) generated by the leftward flow in the node (black wavy lines) might be transferred to the left LPM, either through the endoderm (pink line) – extraembryonic route – or directly transported along extracellular matrix (ECM) through the PAM (red line) – intraembryonic route –. The location of pit cells, crown cells, paraxial mesoderm (PAM), LPM, endoderm and prospective floor plate are shown. Adapted from: (Saijoh, Viotti, & Hadjantonakis, 2014; Shiratori & Hamada, 2006)

First, the injection of the Nodal protein into the right side of the embryo (near the paraxial mesoderm and LPM) induces *Nodal* expression in the right LPM (Oki et al., 2007). Second, transgenic rescue experiments showed that the Nodal co-receptor Cripto/Cryptic is required only in the LPM, not at the node, for left asymmetric *Nodal* expression in the LPM, indicating that the Nodal signal is not relayed indirectly between the node and LPM (Oki et al., 2007). Third, Nodal binds to sulfated glycosaminoglycans (GAGs), which are specifically localized at the extracellular matrix, between the node and LPM, and inhibition of sulfated GAG biosynthesis prevents *Nodal* expression in the LPM, suggesting that Nodal produced at the node might travel directly to the LPM via interaction with sulfated GAGs (Oki et al., 2007). Fourth, ectopic expression of *Nodal* in mouse or Xenopus embryos is capable to diffuse over a long distance like from LPM to

the midline (Kawasumi et al., 2011). Lastly, autocrine-paracrine Nodal signaling in perinodal cells is dispensable for LR patterning of LPM, given that its inhibition by expression of dominant negative forms of Smad3 or ALK4 was still associated with normal (left-sided) *Nodal* expression in LPM (Kawasumi et al., 2011). Although the transport of the heterodimer Nodal-Gdf1 has not been directly detected, secreted Nodal protein was observed undergoing transport in the frog embryo (Marjoram & Wright, 2011). These results collectively suggest that Nodal protein, produced at the node, is transported to the left LPM following an intra-embryonic route. Nevertheless, we cannot exclude the possibility other signaling pathways that act in a paracrine manner, in the mesoderm or endoderm, to transmit the Nodal signal to the LPM (Hamada & Tam, 2014). For instance, a study reported that gap junction proteins in the endoderm are essential to transfer the Nodal signal to the LPM and gap-junction-dependent Ca^{2+} signal can travel from the node to the left LPM (Saund et al., 2012; Viotti, Niu, Shi, & Hadjantonakis, 2012). However, the authors do not propose a mechanism by which Ca^{2+} signal activates *Nodal* expression in the LPM, like Oki et al. proposed in their intra-embryonic route model. Nevertheless, we can think in a combined model of Viotti and Oki hypothesis. This integrated hypothesis, called the Viotti-Oki model (Norris, 2012), proposes that Gap junction-mediated propagation Ca^{2+} signaling may influence the underlying ECM, facilitating ECM-mediated transport of Nodal protein from the node to the LPM (Norris, 2012; Saijoh et al., 2014; Yoshiba & Hamada, 2014).

2.2.1.2.3. Patterning the LPM

After symmetry breaking in the node, asymmetric gene expression arises in the lateral plate mesoderm on the left side. This process is perhaps the best understood step in the establishment of the LR asymmetry and, in mouse, *Nodal* and *Lefty* genes (*Lefty1* and *Lefty2*) play a major role in this process (Shiratori & Hamada, 2006). Although these genes have a similar expression pattern on the left side of the embryo, they have opposite functions. Whereas *Nodal* acts as a left-side determinant, the *Lefty* genes are antagonists of Nodal, restricting Nodal activity to the left side of LPM and also the intensity of *Nodal* expression (Hamada et al., 2002).

The asymmetric expression of *Nodal* and *Lefty2* genes in the left LPM is conserved in all vertebrate species (Nakamura & Hamada, 2012) and its mechanism of action is thought to spread rapidly to the appropriate tissues during its period of competence. To allow the proper engagement of effective programs, the intensity and duration of the signal must be restricted in time and space. This is achieved due to a tighter regulation mediated at two levels, one at the node, due to the action of the Cerl2/Dand5 molecule (explained in section 2.2.1.1.1.4.) and other in the LPM, by the action of *Lefty* genes, which are direct targets of Nodal signaling and act as feedback inhibitors by competitive interaction with the Nodal receptors/co-receptors. Both *Nodal* and *Lefty2* genes are expressed asymmetrically and transiently in the left side of the LPM, starting in a small region adjacent to node, at 2-somite stage, and then expanding along the AP axis, disappearing by the 6-7-somite stage, lasting only several hours (Figure 1.12). Nodal is also thought to travel to the midline or prospective floor plate (PFP), where it induces *Lefty1* expression. *Lefty1* expression is also temporary, beginning at the 2-somite stage, near the node, expanding anteriorly and disappearing by the 5-6-somite stage (Hamada et al., 2002). Such dynamic asymmetric expression is achieved by FoxH1-dependent asymmetric enhancers (ASEs) that are necessary and sufficient for the left side expression patterns of these genes in vivo (Saijoh et al., 2005), ensuring the presence of the Nodal signal at the correct time and place.

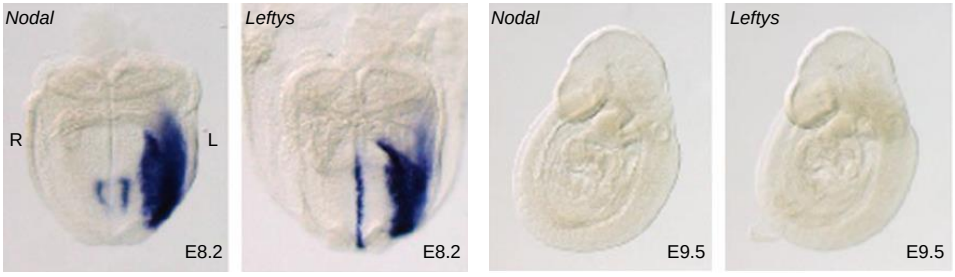


Figure 1.12 – Expression pattern of Nodal and Leftys in wilt-type mice.

At E8.2 *Nodal* is expressed in the node and in left LPM whereas expression of *Lefty* genes localizes in the left LPM (*Lefty2*) and midline (Lefty1). Later, at the time of situs specific morphogens of internal organs, expression of Nodal and Lefty was no longer found in the mouse embryo. Adapted from:(Hamada, Meno, Watanabe, & Saijoh, 2002)

When Nodal protein interact with a target cell in the LPM, triggers the Nodal signaling cascade that in turn leads to the activation of Nodal itself and Lefty2.

The Nodal protein produced will amplify the expression of *Nodal* (positive-feedback loop) and *Lefty2*, whereas the Lefty2 produced will regulate the expression of *Nodal* in left LPM. This negative loop mediated by Lefty2 restricts the intensity and area of Nodal signaling in a highly precise manner, explaining therefore the fact that asymmetric expression of *Nodal* and *Lefty* is indeed transient (Hamada et al., 2002; Saijoh et al., 2005; Shiratori & Hamada, 2006). Lefty1 in turn may function as a midline barrier that prevents Nodal to cross the midline and to induce bilateral expression of left-specify genes like *Lefty2* and *Pitx2*. In fact, *Lefty1* mutant mice present bilateral expression of *Nodal*, *Lefty2* and *Pitx2* and pulmonary left-isomerism (Meno et al., 1998). Moreover, Lefty1 begins to be expressed on the floor plate just before Nodal expression starts on the LPM.

This combination of positive and negative feedback loops between *Nodal* and *Lefty* genes might constitute a self-enhancement lateral inhibition (SELI) system, a type of reaction-diffusion system that is able to amplify a small difference between the left and right of the node into a robust difference in the LPM since the inhibitor diffuses more rapidly than the activator (Nakamura & Hamada, 2012; Shiratori & Hamada, 2006, 2014). Since the Cerl2/Dand5 molecule is responsible for the termination of the Nodal activity at the node, it is thought that this shutdown will also terminate the expression of *Nodal* in the left LPM and consequently the action of the SELI system (Inacio et al., 2013; Nakamura et al., 2006).

2.2.1.3. Asymmetric position and morphogenesis of the internal organs

After the genetic cascade of asymmetric gene expression is accomplished in the left LPM, visceral organs can initiate LR asymmetric morphogenesis synchronized with tissue differentiation. For that, the LR specific orientation initiated around the node and subsequently propagated through the LPM must be interpreted by the different organ primordia and translated into asymmetric gene expression patterns in/or surrounding these primordia. As a result, the cells will undergo asymmetric morphogenesis, which is tightly regulated and

very specific for each organ. This asymmetric morphogenesis can result from at least three different mechanisms: 1 – Directional looping: occurs in organs that are formed from a tube, like heart and gut, and that undergo a series of looping, bending and rotations steps. 2 – Initially symmetric organs develop side specific differences in their size or branching pattern (i.e. lung). 3 – Agenesis of one part: one side of an initial bilateral symmetric structure, such as blood vessels, undergoes regression and disappears, leaving only the other side (Figure 1.13A) (Hamada et al., 2002; Shiratori & Hamada, 2006).

At molecular level, the main readout of Nodal signaling in LPM is the induction of *Pitx2* expression (Shiratori & Hamada, 2014), suggesting that this molecule is the major player in the regulation of asymmetric organogenesis in response to the initial break of symmetry in the node. Pitx2 is a bicoid-type homeobox transcription factor, which, like *Nodal* and *Lefty2*, is asymmetrically expressed in the left LPM (Yoshioka et al., 1998). However, unlike *Nodal* and *Lefty2* genes, the expression pattern of *Pitx2* in the LPM persists until much later stages (Figure 1.13B) and is still apparent in the organ primordia of heart, gut, and blood vessels. This expression is controlled by an intronic enhancer (ASE) in *Pitx2* gene that contains multiple Nodal-responsive, Foxh1-dependent binding sites and an Nkx2-binding site (Shiratori et al., 2001). Thus, the left-side expression of *Pitx2* is initiated by Nodal and maintained by Nkx2, a transcription factor expressed in the developing heart, in the absence of the Nodal signal.

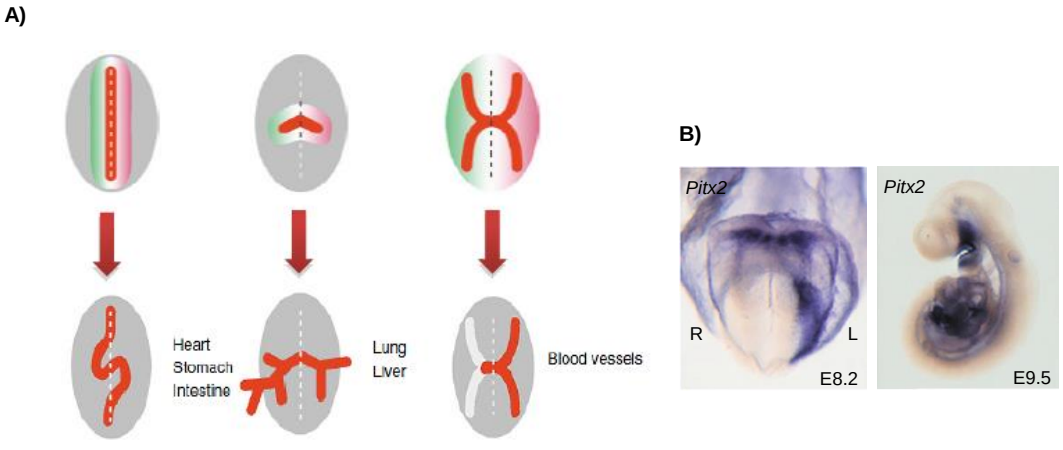


Figure 1.13 – Mechanisms for the generation of morphological asymmetries and *Pitx2* expression pattern. A) Asymmetric morphogenesis of organs can result from directional looping (i.e. heart), differential branching (i.e. lungs) and agenesis of one part (i.e. blood vessels). **B)** Pitx2 seems to be the major player in the regulation of asymmetric organogenesis. *Pitx2* is expressed in the left LPM and heart-forming region at E8.2 and its asymmetric expression persists during situs-specific organogenesis. Red and green represent L/R asymmetric signals. Adapted from:(Hamada et al., 2002)

Pitx2 gene encodes three isoforms, *Pitx2a*, *Pitx2b* and *Pitx2c,* in most mammalian and a *Pitx2d* isoform in humans, however, only the Pitx2c isoforms seems to be cardiac-specific and asymmetrically expressed (Kitamura et al., 1999). Absence of *Pitx2c* in mice results in laterality defects in most visceral organs, like right atrial isomerism and right pulmonary isomerism, and Pitx2-null mutant mice also display severe defects in valve formation, sinoatrial morphogenesis and atrioventricular connections, including double outlet right ventricle and transposition of the great arteries (Kitamura et al., 1999; Liu et al., 2002). *Pitx2c* is expressed in the in the left, not right, side of the developing heart tube prior to looping morphogenesis, suggesting a role for this transcription factor in this process. Although ectopic right-sided expression of *Pitx2c* results in a high frequency of reversed heart orientation (situs inversus), cardiac isomerism and reversed heart looping, the null *Pitx2* and *Pitx2c* knockout mice show normal cardiac looping and a broad spectrum of cardiac defects (Ramsdell, 2005), suggesting that Pitx2c is dispensable for this aspect of the looping process. However, another study in *Xenopus* shows that loss-of-function of *Pitx2c* causes abnormal shifting of the outflow tract (OFT), confirming that Pitx2c can mediate later aspects of looping (Ramsdell, 2005). Furthermore, *Pitx2c* was also expressed in the left atrium, atrioventricular canal, left OFT, interventricular myocardium and second heart field (Furtado, Biben, Shiratori, Hamada, & Harvey, 2011). Together, these evidences suggest that Pitx2 could be involved in the regulation of cardiac looping and morphogenesis of complex cardiac structures. Moreover, errors in the leftward shifting and/or rotation of the OFT are associated with double outlet right ventricle (DORV), transposition of great arteries (TGA), and ventricular septal defects (VSDs), suggesting that the abnormal Pitx2c function could account for these types of Congenital Heart Disease.

Therefore, although the evidence that cardiac looping takes place normally in the absence of Pitx2 suggesting that this molecule do not control the initial direction of looping, its function seems to be necessary and essential for the normal development of the cardiac structures that result from the looping process.

3. How to make a functional heart

During embryonic development, the heart is one of the first organs to form and become fully functional. During this period of heart development, which occurs between days 15-19 and days 56-60 after conception (Sylva, van den Hoff, & Moorman, 2014), many coordinated biological processes take place within the fetus to ultimately result in a four-chambered organ with two parallel circulations. In all higher vertebrates, heart development follows similar basic patterns until the formation of a functional heart. Cardiac development starts with the formation of two bilateral pools of mesodermal precursor cells, the cardiac crescent or so called first heart field, that fuses and folds, giving rise to a simple linear heart tube at the ventral midline (Meilhac, Lescroart, Blanpain, & Buckingham, 2014). After that, the linear heart tube elongates, bends and loops, due to the addition of cardiac progenitor cells by the secondary heart field region (Meilhac et al., 2014; Sylva et al., 2014). Subsequently, growth, differentiation, specialization and septation of the four chambers, valves and great vessels takes place to provide all the necessary components for the physical separation of the systemic and pulmonary circulation of blood. Furthermore, the development of the cardiac conduction system provides the components necessary for the generation of the electrical stimulus as well the means to propagate this impulse to the myocardial cells of the chambers, which results in the heartbeat and the proper sequential contraction of the chambers (Anderson, Spicer, Brown, & Mohun, 2014). Lastly, epicardial and neural crest cells migrate and integrate into the heart. The combination of these changes, along with the generalized maturation of the cardiomyocytes, culminates in the development of the functional four-chambered heart (Meilhac et al., 2014; Sylva et al., 2014).

3.1. Embryonic origins of the heart

Heart formation is a complex process that starts during gastrulation, at day 15 of human development (E6.5 in mouse), when cells that will form the mesoderm ingress through the primitive streak (Schleich, Abdulla, Summers, & Houyel, 2013).

Fate mapping studies have shown that cardiac progenitors that ingress early, at the midstreak stage, become located in the anterior region of the primitive streak (PS) (Figure 1.14A), which comprises newly forming mesoderm (Meilhac et al., 2014). These naıve cardiogenic mesodermal cells migrate away from the primitive streak, anteriorly and laterally on either side of the embryonic midline, to take up a position under the head folds in each side of the splanchnic layer of lateral plate mesoderm, forming two bilateral heart-forming regions (Figure 1.14B), at about day 17-19 in humans and E7.5 in the mouse embryo (Meilhac et al., 2014; Schleich, Abdulla, Summers, & Houyel, 2013). Once they have taken up their position under the head folds, the paired heart-forming regions migrate to the anterior ventral midline and join, giving rise to a crescent-shaped group of cells, where the first and second heart fields can be distinguished. The transition from mesodermal to cardiac progenitor cells requires the induction and specification of the splanchnic mesoderm into cardiac mesoderm, which is mediated by intrinsic signals from the primitive streak and from the surrounding tissues.

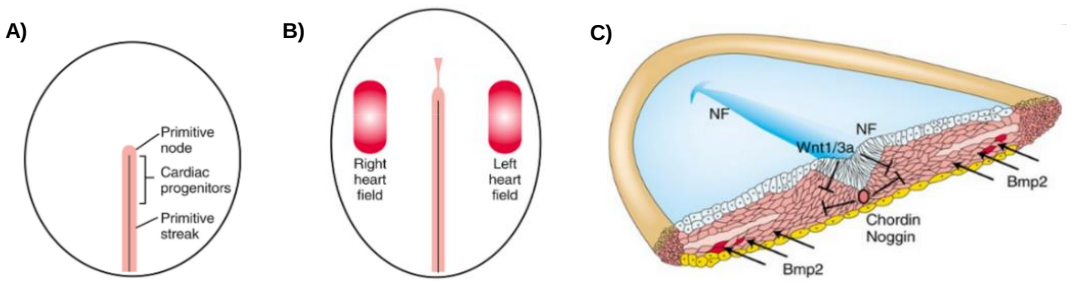


Figure 1.14 – Embryonic origins of the heart.

A) Localization of cardiac progenitor cells after their invagination into the primitive streak. **B)** Migration of the progenitor cardiac cells cranially and laterally results in the formation of two heart-forming regions in each side of the embryo. **C)** Inductive and inhibitory signal released from surrounding tissues results in the specification of the cardiac mesoderm. Adapted from: (Schoenwolf, Bleyl, Brauer, & Francis-West)

During gastrulation, mesodermal cells that ingress through the PS change from expressing *Brachyury T*, a T-box transcription factor, to express mesoderm posterior 1 (*Mesp1*), when they enter in the precardiac mesoderm stage of development (Wu, Chien, & Mummery, 2008). Mesp1, which represents the earliest marker of cardiac progenitors, seems to play a role in the migration of these cells towards the cranial regions of the embryo to form the cardiac crescent (Harvey, 2002). Moreover, Mesp1 induce cardiogenic mesoderm

through the inhibition of genes that promote early mesoderm and endoderm specification and through up-regulation of the core cardiac transcriptional machinery, such as *Gata4*, *Nkx2.5*, *Mef2c*, *Hand2* and *Smarcd3* (*BAF60c*) (Bondue et al., 2008). In addition, inducing signals from the adjacent endoderm (BMP, FGF and Shh) together with inhibitory signals from the notochord (noggin and chordin) and the developing neural plate ectoderm (Wnt11/Wnt3a) refines the site in which the cardiogenic transcriptional program is firstly activated (Figure 1.14C).

3.2. Heart fields and cardiac looping

Around 3 weeks after human development has begun (E8 in mouse), the first cells that migrate to the heart-forming regions, the so-called first heart field (FHF) (Figure 1.15), move ventrally and fuse or coalesce to form a linear heart tube, consisting of an inner layer of endocardial cells and an outer layer of myocardial cells, separated by extracellular matrix or cardiac jelly secreted by the myocardium (Srivastava, 2006). At this stage, the heart is characterized with an anteriorly positioned outflow, or arterial pole, and a posterior inflow, or venous pole, blood starts to courses through the heart tube and heartbeat is initiated (Dunwoodie, 2007). The FHF cardiac progenitor cells, originates from the anterior splanchnic mesoderm, are exposed to cytokines of the BMP and FGF families as well to inhibitors of the Wnt pathway, resulting in the onset of cardiac differentiation, marked by the expression of key regulators of the lineage, namely, *Nkx2.5*, *Gata4* and *Tbx5* (Meilhac et al., 2014). FHF derived cells, ultimately, give rise to the majority of the left ventricle (LV), atrioventricular canal, and parts of the atria (Dunwoodie, 2007). As development proceeds, from day 23 to 28 of human development (E8.5-10.5 in mouse), the heart tube undergoes a dramatic increase in length. This elongation is essentially driven not by proliferation of cells within the linear heart tube, which has been shown to have extremely low cell-cycle times, but by the addition of cardiac progenitor cells from outside the early heart tube to the arterial and venous poles (Miquerol & Kelly, 2013). These cells originated in the pharyngeal mesoderm contiguous and medially to the cardiac crescent have been termed the second heart field (SHF) (Figure 1.15) (Waardenberg, Ramialison, Bouveret, & Harvey, 2014).

SHF cells are characterized by continued proliferation and differentiation delay relative to the cells that gave rise to the linear heart tube and by the expression, along with *Nkx2.5*, of *Isl1*. This expression, which demarks the SHF from FHF derivatives, is dependent on canonical Wnt signaling, and Isl1 function is required for survival, proliferation, migration and commitment of the SHF progenitors into the primitive heart tube. Ultimately, SHF derivatives will form the main parts of the atria, the right ventricle (RV) and the outflow tract myocardium (Meilhac et al., 2014; Waardenberg et al., 2014). Concomitant with the increase in length, the heart tube undergoes rightward looping, in the first morphological manifestation of embryonic laterality (Figure 1.15) (Srivastava & Olson, 2000)

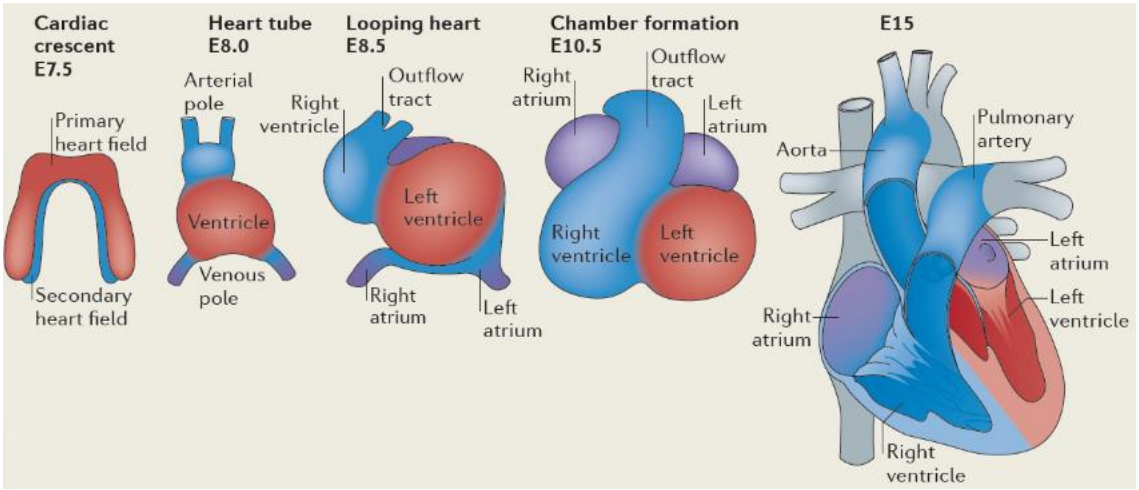


Figure 1.15 – Heart fields regionalization of the heart.

First and second heart field can be distinguished at the time of the cardiac crescent. First heart field derived cells will form mostly the left ventricle and contribute for the formation of the atria. SHF derivatives will form the main parts of the atria, the right ventricle, and the outflow tract myocardium. Adapted from: (Xin, Olson, & Bassel-Duby, 2013).

During cardiac looping, the anterior portion of the heart tube (primitive ventricular region) bends towards its ventral side and simultaneously rotates to the right, forming a C-shaped tube. Next, the C-shaped tube is converted into an S-shaped loop. In this process, the ventricular bend, which was originally cranial to the inflow attachment and atrium, is now caudal (Kirby, 2007). Later looping steps involve the shift of the ventricular position from its originally anterior to its definitive position, posterior to the atria, the shift of OFT from its right lateral position with respect to the atria towards its definitive position ventral to the right atrium and the atrioventricular canal alignment with both

ventricles (Dunwoodie, 2007). Cardiac looping brings the four presumptive chambers of the future heart into their correct spatial relationship to each other and with the vasculature. It is also during the early stages of looping that primitive segments become evident in the heart tube. These include the outflow tract (conotruncus), future right ventricle, future left ventricle, atrioventricular canal, and future atria (Kirby, 2007; Rana, Christoffels, & Moorman, 2013). Rightward cardiac looping is the first morphological indication of asymmetry in the vertebrate embryo and the mechanisms that drive this process are intrinsically related with the correct establishment of the LR axis during gastrulation. In turn, this correct left or right identity depends of a complex network of signals and molecules that are straight regulated and maintained on the lateral plate mesoderm in the form of *Pitx2* expression, like it was described in the section 2.2.1.3.

3.3. Chamber morphogenesis and maturation of the heart

The primitive heart tube is composed of slowly proliferating myocardium, low sarcomeric and sarcoplasmic reticular transcripts, low velocity conduction and poor automaticity and contractility (Rana et al., 2013). As the tubular heart continues to develop, its outer curvature initiates a distinct genetic program by which localized differentiation and proliferation cause expansion of the atrial and ventricular chambers, in a process of ballooning morphogenesis, at ~28-32 of human development (E10.5 of mouse) (Miquerol & Kelly, 2013). While low rates of proliferation are maintained at the inner curvature of the heart tube, including the outflow tract and atrioventricular canal region, the differentiating myocardial cells of the future atrial and ventricular chambers acquire fast conduction velocity, high contractility, high automaticity and well developed sarcomeric structures, which will ensure rapid contractions and efficient blood pumping (Miquerol & Kelly, 2013; Rana et al., 2013). Chamber specification is straightly regulated by the T-box transcription factors. TBX20 and TBX5 promote chamber myocardial development while TBX2 and TBX3 repress the chamber genetic program and promote cushion development in the atrioventricular canal and outflow tract (Miquerol & Kelly, 2013).

Once the chambers are specified and start to grow, its remodeling and maturation processes are initiated. This leads to the formation of the trabeculae, a sponge-like layer of myocytes at the luminal side of the fetal ventricles, that increase the surface area for oxygen uptake in the absence of a coronary circulation (Dunwoodie, 2007). Chamber remodeling and maturation also result in the formation of the ventricular septum at the interface between left and right ventricles, a process that is dependent on a single structure, the interventricular septum (IVS), which has both muscular and mesenchymal components and derives from both right and left ventricular cardiomyocytes (Rosenthal & Harvey, 2010).

It is also during this period that the cardiac conduction system develops from the slow-conducting non-chamber myogenic tissue (Miquerol & Kelly, 2013), consisting of the His-Purkinje system, the sinoatrial node and atrioventricular node. The sinoatrial node (the cardiac pacemaker), located in the right atrium generates a pacemaker impulse, whereas the atrioventricular node delays an electrical impulse for separating the contraction of the atrial and ventricular chambers of the heart (Xin et al., 2013). The His-bundle propagates the signal to the Purkinje fibers, which are embedded in the ventricles, enabling fast and coordinated conduction of impulses throughout the ventricular muscle (Miquerol & Kelly, 2013). The outer layers of the ventricles become highly proliferative, resulting in a thickening of the chamber walls to form the compact layer of the myocardium (Dunwoodie, 2007).

Concomitant with cardiac septation, endocardial cushions in the atrioventricular canal and outflow tract become remodeled to form the cardiac valves, which separate and control the blood flow (Miquerol & Kelly, 2013). In this process, at the level of the OFT and AVC, a cardiac jelly, secreted by the myocardium, between the endocardium and the myocardium, are populated by endocardial cells, which migrate and proliferate to form opposing tissue masses, called endocardial cushions (Miquerol & Kelly, 2013). In the AVC, endocardial cushions are the precursors of the tricuspid and mitral valves, while in the OFT, they will form the aortic and pulmonary valves and a scaffold for the aorticopulmonary septum, which divides the OFT into the aorta and pulmonary artery (Schleich et al., 2013).

3.4. Other sources of cells in heart development

In addition to the myocardial and endocardial populations of cells, epicardial and cardiac neural crest cells are also essential for proper cardiogenesis. Proepicardial cells, localized caudally from the heart, migrate over its outer surface, surround the entire heart, and into the subjacent myocardium, giving rise to the epicardium and contributing to the cardiac-cushion mesenchyme, the cardiac fibroblasts, and to the endothelial and smooth muscle cells of the coronary arteries (Chong, Forte, & Harvey, 2014; Miquerol & Kelly, 2013). Moreover, cardiac neural crest cells (CNC) migrate from the dorsal neural tube to the heart, playing essential roles in septation, contributing to the division of the outflow tract into the base of the ascending aorta, outlet of the left ventricle, pulmonary trunk and outlet of the right ventricle, atrioventricular septum and cardiac valves (Chong et al., 2014). Around week 8 of human development (E16.5-17.5 in mouse), the heart morphogenesis is complete and the fourth chambered heart is now prepared to support the circulatory system where RV – pulmonary ventricle – pumps venous blood to the lungs (pulmonary circulation) at the same time that the LV – systemic ventricle – pumps arterial blood to the rest of the body (systemic circulation) (Figure 1.16).

This overview of the major steps in heart development points to the complicated processes that are highly susceptible to disturbances and can, ultimately, result in congenital heart defects.

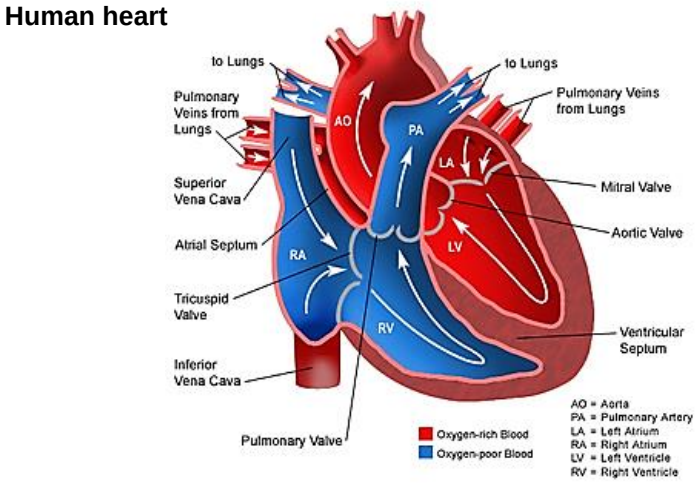


Figure 1.16 – Mature 4-chambered human heart.

Final morphology of the mature four-chambered heart, showing its different compartments and the direction of the blood flow (white arrows).

4. Congenital Heart Diseases

The earliest mention of any Congenital Heart Disease (CHD) dates from Babylonian clay tablets back to 4000 BC. In one of these tablets, ectopia cordis, a type of congenital heart defect, come along with the description “When a woman gives birth to an infant that has the heart open and has no skin, the country will suffer from calamities” (Savona & Grech, 1999). From that period to now, many important breakthroughs were achieved in the identification, characterization and etiology of some forms of Congenital Heart Disease. However, despite all the notable efforts in critical care for affected patients, CHD remains the first cause of both infant morbidity and mortality in the first year of life. Furthermore, the etiology of most CHDs remains unknown, and they are so-called “multifactorial” diseases.

4.1. Epidemiology of Congenital Heart Diseases

With an incidence of approximately 8 per 1000 live births (~1%), Congenital heart defects (CHDs) are the most common type of birth defects, and the leading non-infectious cause of infant mortality (Mozaffarian et al., 2015). Additionally, approximately 3 of every 1000 newborn require some expert cardiologic intervention during the first year of life, which can be debilitating and compromise the quality of life of patients. Therefore, CHDs are also one of the largest cause of infant morbidity and a significant source of economic burden (Cowan & Ware, 2015). Taking into account the high rates of stillbirths with CHDs (10%) (Fahed, Gelb, Seidman, & Seidman, 2013), the presumed cases of spontaneous abortuses, the subclinical abnormalities and the mild lesions that are quiescent for decades, the true overall incidence of CHDs is undoubtedly much greater. Moreover, because prenatal diagnosis, medical and surgical treatments of CHDs become extremely advanced, today, over 75% of children with CHDs that survive in the first critical year of life will live to adulthood (Fahed et al., 2013), making this disease no longer just a childhood disease, but also an important adulthood problem with a high healthcare resource utilization and a significant source of global economic burden.

4.2. Etiology of Congenital Heart Diseases

The etiology of CHD is complex and heterogeneous (Table 1), associated with either genetic and non-genetic causes or a combination of both. Non-genetic causes of CHD include teratogens (dioxins, polychlorinated biphenyls, pesticides), maternal exposures to alcohol, tobacco and drugs, infectious agents (rubella) and known risk factors, like obesity, diabetes mellitus and hypercholesterolemia (Fahed et al., 2013). The genetic landscape of CHD is very heterogeneous and CHD origins are diverse, such as abnormal chromosome structure, coding and non-coding gene alterations, single nucleotide polymorphisms (SNPs), modifier genes, microRNA dysfunction, epigenetic alterations and copy number variants (CNVs). This genetic abnormality can occur as part of a well-defined syndrome, in conjunction with additional extracardiac anomalies, not formally recognized as a syndrome, or, as in the majority of the cases, occurring sporadically in non-syndromic patients (Azhar & Ware, 2016). Syndromic cases of CHD, which accounts for 20-30% of the cases (Azhar & Ware, 2016), includes chromosomal alterations, such as trisomy 21 (Down syndrome), trisomy 13 (Patau syndrome), trisomy 18 (Edwards syndrome), and Turner syndrome aneuploidies, and microdeletions associated with DiGeorge syndrome and Williams syndrome (Azhar & Ware, 2016; Fahed et al., 2013; Fahed & Nemer, 2012). Furthermore, the majority of CHDs with a monogenic mode of inheritance result from well-known syndromes like, Holt-Oram syndrome, Alagille syndrome, and Noonan syndrome (Cecchetto et al., 2010). In contrast, most non-syndromic CHDs occur sporadically, and families with clear monogenic inheritance are infrequent (Wessels & Willems, 2010). This makes the identification of human disease genes involved in non-syndromic difficult and approaches for identify the causes of CHD in non-syndromic patients came mostly from sequencing candidate gene screenings or genome wide association studies (gwas).

Table 1.1. Causes of Congenital Heart Disease.

Cause	CHD attributed (%)	Reference
Environmental	2	[81]
Genetic		
Single gene	3-10	[78]
Chromosomal/aneuploidy	8-10	[78]
Copy number Variation	3–25 (syndromic), 3–10 (isolated	[78]
Multifactorial	Unknown, estimated 80–85	[78]

With the advance in the understanding of the molecular basis of heart development, mostly using animal models, and the recognition that perturbations during embryonic heart development are at the origin of CHD. Studies on large cohorts of isolated CHD identified three characteristics of the disease: genetic heterogeneity, reduced penetrance and variable expressivity (Cecchetto et al., 2010). Therefore, it became evident that not all CHD manifests true Mendelian inheritance and it is possible that multiple common low penetrance genetic variants will have a weak to moderate effect in raising an individual's risk of having a heart malformations when considered together (Lalani & Belmont, 2014). At this time, a model involving both rare and common variants, possibly with gene-gene and gene-environmental interactions, is the most likely explanation for the high frequency of CHDs.

4.3. Laterality defects and Congenital Heart Disease

Laterality defects describe a group of congenital disorders that involve abnormal arrangement and morphology of the abdominal and/or thoracic organs (Cohen et al., 2007). In addition to these anomalies, it is usual to find complex malformations within the heart, often requiring cardiac surgery in the neonatal period. Laterally defects result from abnormalities in the left-right axis establishment during the early stages of development and the spectrum of the disease includes different arrangements of the normal *situs* of the body, called *situs solitus* (Figure 1.17) *(Peeters & Devriendt, 2006)*. This spectrum includes, *situs inversus* and heterotaxy (Figure 1.17). *Situs inversus* (SI) is a defect in the global *situs* orientation and results in a complete mirror image of the normal arrangement of the body and organs (*situs solitus*). This condition, which is

characterized by the maintenance of the normal position of all organs with respect to each other, occurs in 1 per 6000-8000 newborns and presents an incidence of congenital heart defects around 3% (Peeters & Devriendt, 2006). Heterotaxy (HTX), also called *situs ambiguus or* heterotaxy syndrome, refers to any disruption in the normal left-right asymmetry of the thoracoabdominal organs that cannot be strictly classified as *situs solitus* or *situs inversus* (S. J. Kim, 2011). The incidence of heterotaxy is estimated as 1 in 10000 births and its etiology is genetically highly heterogeneous with reduced penetrance, presumably because of monogenic, polygenic or multifactorial causes (H. Deng, Xia, & Deng, 2015; Peeters & Devriendt, 2006). Furthermore, SI and HTX occur among families or in sporadic cases, associated with other syndromes or in isolated cases (Peeters & Devriendt, 2006). In the latter, alterations in genes involved in the Nodal signaling pathway seem to be the most probable cause of

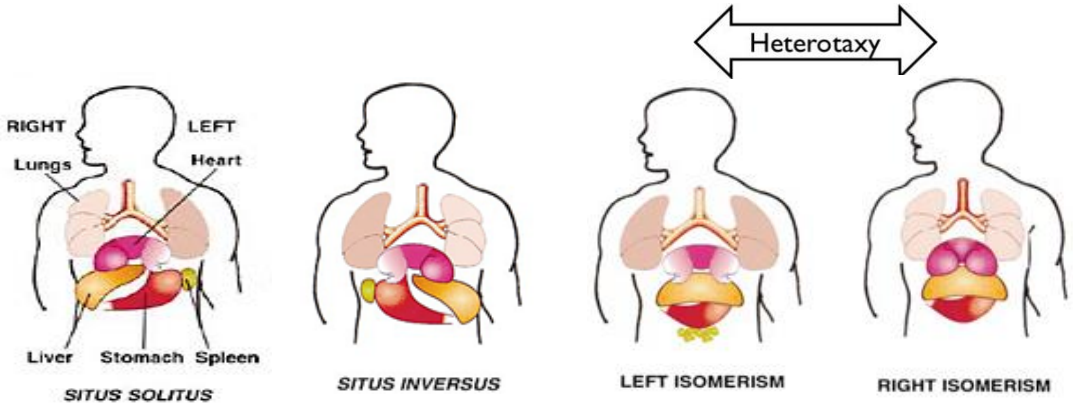


Figure 1.17 – Human laterality defects.

Illustration of normal arrangement (*situs solitus*) of internal organs according left–right asymmetry and the spectrum of laterality defects that affect the lungs, heart, liver, stomach and spleen. Left or right isomerism reflects a duplication of the left or the right side respectively. Adapted from: (Capdevila, Vogan, Tabin, & Izpisua Belmonte, 2000).

the disease. In addition, SI is usually characterized by autosomal recessive inheritance, while HTX displays X-linked inheritance, autosomal dominant inheritance with reduced penetrance, or autosomal recessive inheritance (H. Deng et al., 2015).

Contrarily to SI, which presents an incidence of congenital heart defects around 3%, almost all the patients (90%) with heterotaxy also have complex congenital heart defects like pulmonary atresia/stenosis, transposition of the great arteries

(TGA), double outlet right ventricle (DORV), ventricular septal defects (VSDs), atrial septal defects (ASDs), atrioventricular septal defects (AVSD), hypoplastic left heart, anomalous venous return and coarctation of the aorta (Mohapatra et al., 2009). Furthermore, heterotaxy also includes isomerisms (Figure 1.17), a condition in which asymmetric organs fail to show asymmetry. In left isomerism, the heart has two long narrow atrial appendages and both lungs are bilobed whereas in right isomerism, both atrial appendages are pyramidal in shape and the lungs are trilobed (Peeters & Devriendt, 2006). Both of these conditions present characteristic heart and extra-cardiac alterations, which are not mutually exclusive to one or other phenotype and may overlap. The prognosis of patients with complex cardiac lesions and heterotaxy is poor, with the mortality in the first year of life being >85% for patients with right isomerism and >50% for patients with left isomerism (S. J. Kim, 2011).

Taking into account the vast number of studies, from animal models, in heart morphogenesis and determination of left-right axis during embryogenesis, the evidence that the mortality of laterality disease are almost always attributed to complex congenital heart defects and the fact that the developing heart is extremely susceptible to disturbances in embryonic left-right patterning (Ramsdell, 2005), is fair enough suggest that congenital heart diseases could be due to problems of laterality. Therefore, the genes that control left-right specification could be implicated in the etiology of isolated cases of CHDs. In the same line of thinking, alterations in these genes might also contribute for some phenotypes of CHD patients.

4.3.1. Cardiac left-right asymmetry

The first sign of cardiac molecular asymmetry comes from the cells that form the primitive heart fields. These cells, which migrate from the primitive streak to the left and right embryonic midline, are exposed to different left-right patterning and will contribute to the formation of different regions of the heart. As an example, the cells that contribute to the posterior region of the heart tube are the same subset of cells that previously expressed *Nodal* and *Lefty2* prior to their incorporation into the primary heart tube (Ramsdell, 2005). This suggests

that although these genes are no longer expressed once cardiac cells become incorporated into the heart tube, this subset of posterior cardiac cells possesses a background of molecular asymmetries that distinguishes it from other cells in the heart. Furthermore, the second heart field, which arises from an area that includes the mesoderm surrounding the aortic sac and all of the pharyngeal arches, also presents left-right differences since *Pitx2c* expression is detected in the left, but not right, aortic sac mesoderm and left pharyngeal arch mesenchyme (Rosenthal & Harvey, 2010). The SHF cells, which present an asymmetric *Pitx2c* expression, will contribute to the conotruncus or outflow tract of the heart. This means that as cardiac mesoderm becomes established and as it contributes to the heart tube, these cells are already specified, at least to some extent, for left–right identities.

As cardiac development progresses, the linear heart tube undergoes rightward looping, under influence of Pitx2-dependent or Pitx2-independent Nodal signaling, with inner curvature remodeling and outer curvature proliferation, the endocardial cushions of the inflow and outflow tracts are brought together at the inner curvature and become adjacent to each other (Ramsdell, 2005). If either ends of the heart tube are delayed or impaired in the looping, the chance to occur septal defects increase. In addition, abnormal leftward looping results in reversal ventricular topology such that the morphological right ventricle (pulmonary circulation) is situated to the left of the heart and the morphological left ventricle (systemic circulation) on the right, a phenotype that can be found in *Zic3* and *Cited2* knockout mice, which are molecules also involved in left-right asymmetry establishment (Dykes, 2014).

Subsequently, in the inflow tract, the common atrioventricular canal becomes divided into left and right components by an atrioventricular septum that shifts to the right. This results in alignment of the atrioventricular canal (AVC) with the ventricles. If the inflow tract fails in shift to the right both atria communicate with the left ventricle in a condition known as double-inlet left ventricle (DILV) (Kathiriya & Srivastava, 2000; Ramsdell, 2005). Meanwhile, at the outflow region, the conotruncus shifts to the left, becoming positioned above the AVC, and simultaneously twists 180º to position the aorta properly over the left ventricle and the pulmonary artery correctly positioned over the right ventricle.

Failure in shift to the left, results in double-outlet right ventricle (DORV), a condition in which the right ventricle communicates with both the aorta and the pulmonary artery and the left ventricle has no outlet, whereas defects in the rotation results in transposition of the great arteries (TGA), a condition in which the right ventricle communicates with the aorta and the left ventricle with the pulmonary artery (Kathiriya & Srivastava, 2000; Ramsdell, 2005).

Left and right atria results from common atrium, which division may be related to differences in atrial cells that derive from left vs right primary heart field. Failure in this process is the base of atrial isomerisms (Ramsdell, 2005). Because the right atrium contains the sino-atrial node and other specialized structures of the conduction system, right atrial isomerism is associated with cardiac arrhythmias as a result of node duplication while left atrial isomerism is associated with hypoplastic or absent sino-atrial node. *Pitx2c* expression is required in the left atrium during the late phase of asymmetric differentiation where it appears to actively suppress the development of the conduction system (Dykes, 2014). Moreover, during cardiac cushion formation from the AVC, myocardium in the AVC exhibits left, but not right, side expression of *Pitx2c.* Laterality disturbances of AVC affect the formation of the valves and increases the risk of valvuloseptal defects (Ramsdell, 2005). Ultimately, the left atrium and right atrium are aligned with the left ventricle and right ventricle, respectively.

Therefore, the correct morphogenesis of the various compartments of the heart depends on the correct establishment of left-right axis. If this complex process fails, it might result in many congenital heart defects with or without associated extracardiac laterality defects, depending on the levels and distribution of the left determinants in lateral plate mesoderm and heart.

5. Disease modeling using patient-derived iPSCs

Despite all the advances in medical therapy, heart disease remains a leading cause of mortality and a major worldwide healthcare burden (Ebert, Diecke, Chen, & Wu, 2015). Animal models have provided indispensable insights into the pathogenesis of genetic cardiac diseases, however not all findings from research in animal models cardiomyocytes can be translated into human cardiomyocytes. For example, mouse models are not always able to fully recapitulate the disease phenotype seen in patients, due to species differences at both physiological and genetic levels (Avior, Sagi, & Benvenisty, 2016). Moreover, the limited regeneration capacity of heart myocardial cells (Dell'Era et al., 2015) in conjunction with the fact that heart transplantation is a very limited option due to organ supply paucity, immunosuppression and organ rejection, makes it difficult to study cardiac disorders and emphasize the need for novel human cellular and physiological models of heart diseases (Ebert et al., 2015).

To overcome these limitations several efforts were made in the last 15 years and one of the most exciting was the ground breaking work reported by Yamanaka a decade ago. In 2006-2007, Yamanaka reprogramed for the first time murine and human somatic cells to a pluripotent state, transducing retrovirus-carrying four specific transcription factors (Oct3/4, Sox2, c-Myc, and Klf4 also known as the Yamanaka factors) (Takahashi et al., 2007; Takahashi & Yamanaka, 2006) in those cells. The resulting induced pluripotent stem cells (iPSCs) present an intrinsic capability for indefinite self-renewal and have the potential to differentiate into virtually any cell type, resemble embryonic stem cells (ESCs), the “gold standard” for pluripotency (Bellin, Marchetto, Gage, & Mummery, 2012).

iPSCs technology has captured scientists with the promise and hope that these cells hold a potential that can be used in disease modeling, drug development and regenerative medicine without the ethical and political concerns associated with the use of human embryos for the generation of human embryonic stem cells (C. Kim, 2014).

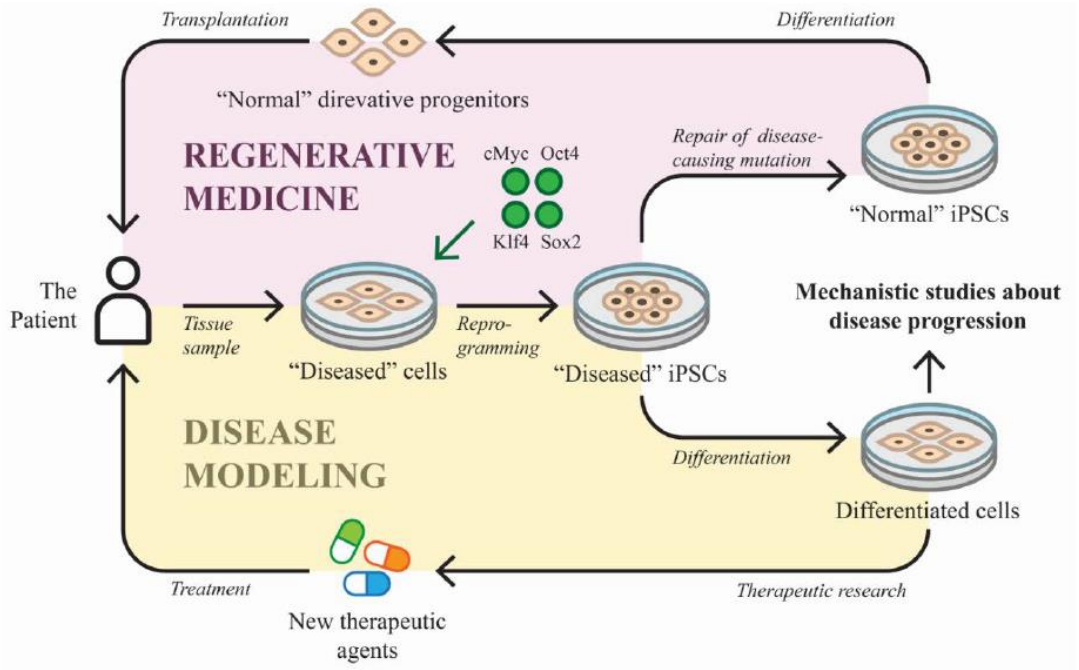


Figure 1.18 – Applications of iPSCs technology

Patient iPSCs are mostly generated using the four Yamanaka factors and can be used in disease modeling, drug development and regenerative medicine.

Since the initial publication, human iPSCs have developed into impetuous, spectacular, and groundbreaking research field and in our days iPSCs can be generated in any period of patient's life from almost any type of differentiated cells (C. Kim, 2014). It is now possible obtain iPSCs-derived cardiomyocytes from healthy people and also from patients suffering of a genetic disease very similar to native cardiomyocytes in morphology, gene expression pattern, electrophysiological rates, and sensitivity to chemical substances (Bayzigitov et al., 2016). The resulting human iPSCs-derived cardiomyocytes will keep the entire genetic information of the patient, allowing the study of the effect of the variant in different backgrounds and the functional effect of the variant compared to the healthy people cardiomyocyte.

Therefore, the use of hiPSC disease model is an alternative platform for studying mechanisms of CVD and evaluating novel therapeutic avenues. In addition, human iPSCs present the unprecedented opportunity to study disease-specific differences in a patient-specific manner.

6. Objectives

Despite all the notable efforts over the last 15 years, recurring to a vast array of both approaches and techniques/technologies, in the attempt to unveil the causes of all forms of CHDs, the etiology and the genetic basis of the majority of congenital defects remains unknown. There is a distinct lack of data concerning the molecular mechanisms that are required for normal human cardiac development and is difficult to make the connection between normal cardiac development and CHDs, mostly because there are no direct methods for determining which developmental step(s) were disturbed in any single affected individual. Furthermore, many CHDs have an ambiguous embryologic origin and they may be interpreted as arising from multiple alternative early events or later and more specific processes.

Therefore, in the sequence of several studies in heart morphogenesis and determination of the left-right axis during embryogenesis, concerning to the Nodal pathway and the role of its involved genes, it becomes clear the need of beginning the translational research, applying the new data resulting from the animal models research to the human cardiac disease, in particular to the disorders of laterality and/or isolated cases of CHDs.

The purpose of this project, in a close collaboration between research and clinical practice is the molecular and functional study of human genes involved in the Nodal pathway, focusing mostly in *DAND5* (human homolog of mouse *Cerl2*), but also trying to find alterations in the other intervening genes of the pathway, namely *CFC1* (human homolog of mouse Cripto), *NODAL* and *PITX2C.*

The objectives of this thesis were: 1- to define the cellular and molecular mechanisms that underlie cardiac asymmetry on the congenital heart disease (CHD) and the laterality defects. 2- to identify the correlation between the observed/described phenotypes in affected children and their genotype. 3- to understand the hereditary mode of transmission in case of pathologic alteration identification and assess the frequency and role of these genetic variations in children with congenital heart defects resulting from left-right abnormalities.

Chapter 2

Methods

1. Ethics

This study was conducted in accordance with the ethical principles of the revised Declaration of Helsinki. The research protocol was approved by the local institutional ethics committee of Hospital de Santa Cruz, Centro Hospitalar de Lisboa Ocidental, E.P.E., Hospital de Faro, Centro Hospitalar do Algarve, E.P.E., Hospital Sta. Marta, Centro Hospitalar de Lisboa Central, E.P.E. (Supplementary Material 1 as example) was obtained from all guardians participants prior to DNA collection. The biological samples collected were anonymized to ensure the confidentiality and privacy of patients and data.

2. Study population

The study sample included 38 Caucasian unrelated pediatric patients with congenital heart defects arising from possible left-right defects recruited from Hospital de Santa Cruz, Centro Hospitalar de Lisboa Ocidental, E.P.E. (11), Hospital de Faro, Centro Hospitalar do Algarve, E.P.E. (24), Hospital Sta. Marta, Centro Hospitalar de Lisboa Central, E.P.E (3). The sample was obtained according to the convenience and availability of physicians in each hospital.

3. Clinical evaluation and inclusion criteria

Patients were evaluated by individual and familial history, review of the medical records and individual anatomic descriptions of the patients' defects were obtained by the review of echocardiogram, echocardiography, magnetic resonance imaging, angiography or direct view during cardiac surgery.

Since *DAND5* (*Cerl2* in mice) gene was recently associated to congenital heart defects arising during embryo development, we decided to base our inclusion criteria firstly in the phenotype features observed on *Cerl2* KO mice. In this animal model it was observed that 35% of the homozygous mutants died within the first 48 h after birth, presenting several cardiovascular malformations,

including incomplete atrial and ventricular septation, conotruncal defects, ventricular hypertrophy, and/or to laterality defects like left isomerism and thoracic heterotaxy or situs inversus. In addition, it was performed an exhaustive bibliographic search of the most commons symptoms and features presented by the animal models used in the study of CHD and laterality disorders, namely the generated KO mice for other genes involved in the Nodal signaling pathway, and the clinical symptoms presented by affected children or adults with CHD associated mainly with laterality defects. Nevertheless, throughout the contact with clinicians, in particular with the paediatric cardiologist, it was defined a list of the main diagnosed CHD phenotypes in each of the involved institutions. Conjugating all of these factors, the inclusion criteria required the presence of congenital heart malformations and/or laterality defects, namely: a) left or right isomerism; b) heterotaxy; c) situs inversus; d) transposition of great arteries (TGA); e) double inlet left ventricle (DILV); f) double outlet right ventricle (DORV); g) tetralogy of Fallot or Fallot like phenotype; h) atrioventricular septal defects (AVSD); i) atrial septal defects (ASD); j) ventricular septal defects (VSD) and j) ventricular hypertrophy (Supplementary Material 2 as example). The patients with known chromosomal abnormalities or syndromic cardiovascular defects, such as Down syndrome, Turner syndrome, Marfan syndrome, Di George syndrome, and Holt–Oram syndrome, were excluded from the study.

4. Patient's phenotypes

The phenotypes of the patients included in the study are described in the table 2.1:

Table 2.1. Patient´s Phenotypes

HStM001	Left Isomerism; pulmonary atresia; VSD; right-sided aortic arch; inferior vena cava interruption; visceral *situs inversus*
HStM004	Left Isomerism; AVSD
HStM006	*Situs inversus totalis*; AVSD; TGA; Univentricular Heart; right-sided aortic arch
HsTC001	VSD; coarctation of the aorta; aortic stenosis; mitral valve stenosis

HStC002	Hypoplastic right ventricle with pulmonary atresia and tricuspid atresia
HsTC003	Hypoplastic right ventricle; pulmonary atresia with intact ventricular septum
HsTC004	VSD; pulmonary atresia; MAPCAs; PLSVC
HsTC005	TGA; pulmonary atresia; tricuspid atresia
HsTC006	AVSD
HsTC007	Left Isomerism; VSD with overriding aorta; pulmonary atresia
HsTC008	TGA
HsTC009	VSD; pulmonary atresia; MAPCAs
HsTC010	*Situs inversus*; levocardia; pulmonary atresia; single ventricle; single atrioventricular valve
HsTC011	TGA; VSD
HCF001	ASD; pulmonary atresia
HCF002	Tetralogy of Fallot
HCF003	Tetralogy of Fallot
HCF004	Right isomerism with asplenia; dextrocardia; TGA; AVSD; pulmonary stenosis; TAPVC; right-sided aortic arch
HCF005	Dextrocardia; VSD
HCF006	Tetralogy of Fallot, TAPVC
HCF007	AVSD
HCF008	VSD; pulmonary atresia
HCF009	TGA
HCF010	Hypoplastic left heart, TAPVC
HCF011	AVSD
HCF012	VSD; pulmonary atresia; TGA
HCF013	Left Isomerism
HCF014	TGA
HCF015	TGA; DILV
HCF017	Pulmonary atresia; VSD
HCF019	DILV
HCF021	Tetralogy of Fallot
HCF023	Tetralogy of Fallot; Dextrocardia
HCF025	Tetralogy of Fallot; Pulmonary valve agenesis
HCF027	Tetralogy of Fallot; Double outlet right ventricle
HCF029	AVSD
HCF030	AVSD
HCF032	Hypoplastic Aortic Arch; VSD; ASD

TGA - Transposition of the great arteries; **AVSD -** Atrioventricular septal defects; **DILV -** Double inlet left ventricle **VSD -** Ventricular Septal defects; **TAPVC -** Total anomalous pulmonary venous connection; **DORV -** Double outlet right ventricle; **MAPCAs** - Major Aorto-Pulmonary Collateral Arteries; **PLSVC** - Persistent left superior vena cava

5. Genetic Screening of Nodal signaling pathway genes

5.1. Sample collection

To have a rapid, reproducible and non-invasive/painless way for biological sample collection in children, which could be easy to transport and to conserve, we collected buccal epithelial cells using buccal swabs. The samples collection was performed during the clinical follow up consultation in the health institution where the affected children are being followed. It was recommended that the patients avoid eating thirty minutes before the samples collections and washed the mouth only with water. The sample protocol was simplified and briefly, we identified the swabs with the patient data and collect the sample scraping the inside of the cheek approximately 6-10 times. Then we let the swaps with the samples dry at room temperature for about 15 minutes and finally the collected samples were sent to our laboratory (directly in the same day of collection, by mail, or other).

5.2. DNA extraction and quantification of DNA concentration

The DNA from the buccal epithelial cells was extracted using the Isolate Genomic DNA mini kit (BIOLINE) according to the manufacturer's instructions for DNA isolation from buccal swabs (Supplementary Material 3).
After extraction, the concentration of DNA sample suspensions was determined by spectrometry using an NanoDrop 2000 UV-Vis Spectrophotometer (ThermoScientific).

5.3. Genes analyzed

The clinical complexity and genetically heterogeneity of congenital heart diseases suggests that some of these defects could arise due to alterations in more than one gene that orchestrate the cardiac program, or in gene pathways

that control essential embryonic processes before the cardiac program starts. Therefore, in order to have a better picture of the genetic cause of a particular CHD phenotype, some of the NODAL signaling pathway genes were sequenced, namely *DAND5*, *NODAL*, *CFC1* and *PITX2C*.

The nucleotide and protein sequences of these genes or gene products were obtained through the genomic databases for human gene sequences and diseases, namely the OMIM, Ensembl, Uniprot and some NCBI sub-databases (the reference of each gene from each of the chosen databases are showed in table 2.2). From the gene nucleotide sequences, it was analyzed the exonic and intronic regions, the coding and non-coding regions, and the main domains already identified for each gene in order to find the more relevant sequences for gene variant screening. Consequently, it was decided to search for variants in the exons and exon-intron boundaries of all the genes mentioned above.

Table 2.2. Reference of the genes analyzed in this study in the different databases.

Gene	Gene ID (NCBI)	Ref Seq accession number	OMIM identifier	Ensemble	Uniprot
DAND5	199699	NM_152654.2	609068	ENSG00000179284	Q8N907
NODAL	4838	NM_018055.4	601265	ENSG00000156574	Q96S42
CFC1	55997	NM_032545.3	605194	ENSG00000136698	P0CG37
PITX2C	5308	NM_000325	601542	ENST00000306732	Q99697

5.4. Oligonucleotide primers design

To perform the DNA amplification by polymerase chain reaction (PCR), considering the genes *DAND5*, *NODAL*, *CFC1* and *PITX2C,* several sets of primers were designed targeting the coding exons and flanking exon-intron boundaries using the online program PRIMER-BLAST. The designed primers were checked with the PRIMER-BLAST software (internet free access), in order to amplify only the target gene region, without homology for other genes/genomic regions that could interfere in the genetic analysis and on

Vector NTI to access the dimerization capability and significant hairpin formation.

The genomic localization, the nucleotide sequence and the amplicon length of the designed primers are shown in Figure 2.1 and table 2.3, respectively:

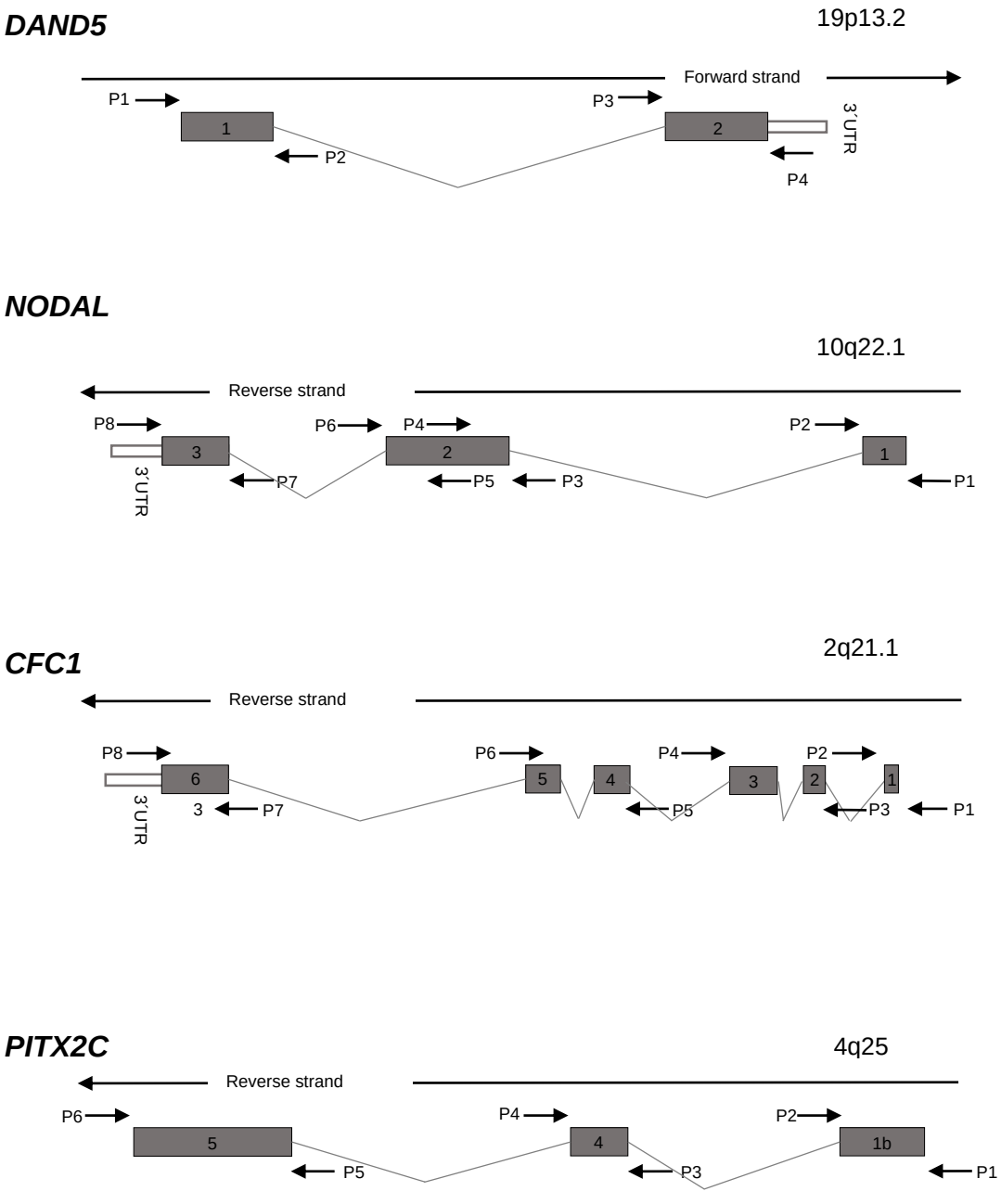


Figure 2.1 – Genomic localization of *DAND5*; *NODAL*; *CFC1* and *PITX2C p*rimers.
Primers are notated with "P" letter; the corresponding nucleotide sequences are shown in table 2.3.

Table 2.3. Primers used for amplification of specific genomic regions

Gene and exon	Forward (5' – 3')	Reverse (5' – 3')	Amplicon size (bp)
***DAND5* Ex1**	P1 - GTCGACTGCTAGTGACCTTGAG	P2 - TCAGGTGGAGGATACAGGACTT	499
***DAND5* Ex2**	P3 - GGAAGTGGACAGGTGATTATCC	P4 - CACGTCTTTCTTGGTCCATCTC	399
***NODAL* Ex1**	P1 - GCTGAGGCCCAGGATATAAG	P2 - CCAACCCACAGCACTTCC	341
***NODAL* Ex2a**	P3 - CCTCACTCACCCTCCTTCC	P4 - CTCCAGCTACCCTGGACATC	383
***NODAL* Ex2b**	P5 - CCTCTCTCCAAGTGGCTGAAG	P6 - GCAAAGCTAGAGCCCTGTCC	489
***NODAL* Ex3**	P7 - CTATTCTGACCTGCCCATC	P8 - CTGGATTAGATGGTTATCATGTC	292
***CFC1* Ex1**	P1 - CATCCACATCTTCTGTACTC	P2 - GCACTACGATTTATTCTGAG	234
***CFC1* Ex2-3**	P3 - GCCTCAGAATAAATCGTAGTGC	P4 - CTCAGGTCCTAAACTCTGAGTCC	646
***CFC1* Ex4**	P5 - GGCTGGAGGCTGATTTTACTG	P6 - GACCCGCAGGTCCCTCAC	289
***CFC1* Ex5**	P7 - CAGGTCCCCGTCTTCTCAG	P8 - CACTGGAAGATGCACTGTGG	229
***CFC1* Ex6**	P9 - GCAGGCGTTTCTATTGCAC	P10 - CCAGTGCTTCAGCTTACGG	334
***PITX2C* Ex1b**	P1 - CAGTAGCCAAGGACTAGTAG	P2 - CAGAATTGCTCGCGCCCTTAG	472
***PITX2C* Ex4**	P3 - AGTGAATGTGCCGCTGCAGT	P4 - TCGGAGAGGGAACTGTAATC	646
***PITX2C* Ex5**	P5 - GTTGGCCTCCGATGGAAGTT	P6 - GCCCACGACCTTCTAGCATAA	728

Note: The primers listed above were used for the DNA amplification of the buccal swab samples.

5.5. DNA amplification

For each patient genomic DNA sample, PCR reactions for each of the genes analyzed (*DAND5, NODAL, CFC1* and *PITX2C*) were performed using the primers listed in table 2.3. PCR reactions were carried out using 10 - 30 ng of genomic DNA in a 25 µL reaction containing 1X Phusion HF Buffer (ThermoScientific), 0.5 µM of each primer (Sigma- Aldrich), 0.2 mM dNTPs and 0.02 U/µL of Phusion High-Fidelity DNA polymerase (ThermoScientific). PCR cycling (Figure 2.2) was performed on C1000 Thermal Cycler (Bio-Rad) using a denaturation cycle at 98ºC for 1 min followed by 35 cycles of denaturation at 98ºC for 15 s, annealing of primers at a specific temperature, depending on the gene/exon, for 30 s and extension at 72ºC for 30 s and a final extension step at 72ºC for 5 min.

After amplification and to determine the quality and yield of the PCR reaction, 5 µL of PCR product of the genomic DNA of interest and its correspondent amplicons length was analyzed by 1-2% (w/v) agarose gel electrophoresis.

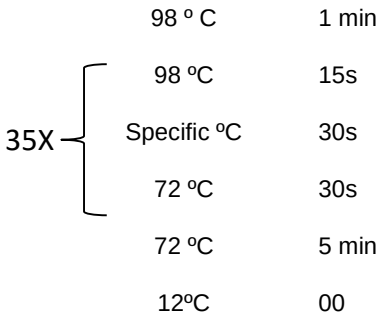


Figure 2.2 – **Standard PCR protocol used for the amplification of *DAND5*, *NODAL*, *CFC1* and *PITX2C*.**

5.6. Purification of PCR products

Purification of the specific PCR products were performed using three different technics. Ethanol and sodium acetate precipitation (Supplementary Material 4) procedure was the first choice during the first two years of PhD but since we experience some problems in the sequencing of some amplicons, we decided to perform the purification of the DNA using an available commercial kit. In some amplicons it was systematically detected the presence of unspecific bands in the agarose gel electrophoresis, and it was decided to extracted the specific band from the gel and perform the purification of the DNA using an available commercial kit.

Kit purification

The PCR product containing only the specific DNA fragment of interest was purified using the commercial kit NZYGelpure kit (NZYTech) and according the manufacturer's protocol. The DNA was ultimately eluted in 30 µL nuclease-free water and kept at -20ºC till further analysis.

Extraction of DNA from Agarose Gels

Following agarose gel electrophoresis, agarose gels were viewed under UV illumination and regions containing relevant DNA fragments were cut from the

gel using a sterile scalpel. DNA products were then recovered using the NZYGelpure kit (NZYTech). The manufacturer's protocol was then followed with DNA eluted in 30 µL nuclease-free water and kept at -20ºC till further analysis.

5.7. Sequencing and sequencing analysis

Direct bidirectional sequencing of each PCR product were performed with an ABI Prism 3130 automated genetic analyzer (Applied Biosystems) and ABI 3730xl DNA Analyzer (Applied Biosystems) according to the protocols from the Sequencing Service of CCMAR (http://www.ccmar.ualg.pt/) and STAB VIDA (http://www.stabvida.com/), respectively. The primers used for the sequencing were the same used to obtain the amplicons of each PCR product.

The chromatogram of each DNA sequence were analyzed using the BioEdit Sequence Alignment Editor Software (http://www.mbio.ncsu.edu/bioedit/bioedit.html) and the sequences were aligned and compared with consensus sequences obtained from the human genome databases in the online program BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi), from NCBI.

The variants found were confirmed by sequencing three independent PCR-generated amplicons from the same patient. Additionally, the identified sequence variation was queried in the single nucleotide polymorphism (SNP) database at NCBI (http:// www.ncbi.nlm.nih.gov/), the human gene mutation (HGM) database (http://www.hgmd.org/), the 1000 Genome Project (1000 GP) database (http://www.1000genomes.org/) and the ExAC browser (http://exac.broadinstitute.org/) to confirm its novelty.

5.8. Sequence alignment of the variant proteins among vertebrate species

Conservation of the amino acid altered by each missense allelic variant was estimated by aligning the human protein of the corresponding gene variant to

the homologous proteins in chimpanzee, rat, mouse, xenopus and zebrafish using Clustal Omega program from Uniprot (http://www.uniprot.org/).

5.9. Characterization of the Causative Potential of the Variants

The potential effect of the p.R152H DAND5 variation and *CFC1* variant were characterized by the software programs Polyphen-2 (http://genetics.bwh.harvard.edu/pph2/) Mutation Taster (www.mutationtaster.org), PROVEAN (http://provean.jcvi.org/), and I- Mutant 3.0 (http://gpcr.biocomp.unibo.it/) . Polyphen-2 utilizes a combination of sequence, function, evolutionary conservation and structure based attributes and uses naive Bayesian classifier to determine the effect of an amino acid alteration. The output levels of damaging in the protein were classified as deleterious (≤ 0.5) and the benign level being classified as tolerated (≥ 0.51). Mutation Taster software's analysis are based on protein structure/function and evolutionary conservation and employs a Bayes classifier to eventually predict the disease potential of an alteration. The outcome of all tests can be either a disease mutation or a harmless polymorphism.

PROVEAN based analysis takes into account the alignment and measurement of similarity between variant sequence and protein sequence homolog and the output classifies the alteration as deleterious (≤ -2.5) or neutral (≥ 2.51). I-Mutant 3.0 is a support vector machine-based tool that predicts the stability change of a mutated protein. The output prediction classifies the alteration in 3 classes: neutral alteration (− 0.5 ≤ DDG ≤ 0.5 kcal/mol), large decrease of stability (≤− 0.5 kcal/mol) and large increase of stability (> 0.5 kcal/mol).

6. Functional analyses of the p.R152H DAND5 protein variant in the Nodal signaling

In order to study the effect of the p.R152H variant in the Nodal signaling pathway we used a Nodal-dependent luciferase assay system as described in the methods, section 6.2. To perform that assay, the human version of the main

intervenient genes in this pathway, namely, *NODAL, CRIPTO, DAND5, DAND5 R152H* variant and *FOXH1***,** was cloned in expression vectors. To ensure the same backbone plasmid for each gene the plasmids were subcloned into the expression vector pCS^{2+}.

6.1. Plasmids

hNODAL

The plasmid containing the *hNODAL* gene cloned in the pcDNA3 expression vector was obtained as a gift from Dr. Michael R. Kuehn from the Center for Cancer Research National Cancer Institute. To subclone the *hNODAL* sequence into the expression vector $pCS2^{+}$, the *hNODAL* sequence was cut using the restriction enzymes XhoI and EcoRI (ThermoFisherScientific). Then, the digested fragments were separated by electrophoresis in a 1% agarose gel, the specific band containing the hNODAL sequence extracted from the gel with a scalpel and purified using the commercial available kit NZYGelpure (NZYTech). Next, we performed the ligation of this DNA fragment to the $pCS2^{+}$, which had been previously linearized and purified in the same way as *hNODAL*, using the T4 DNA Ligase (TihermoFisherScientfic) according to the manufacturer's instructions. The ligation product was then transformed into competent Top 10 bacteria. Three clones were picked, grown in suspension overnight and the plasmids purified using the NZYMiniprep kit (NZYTech). Ultimately, the purified plasmid was sequenced to confirm that the *hNODAL* was successfully integrated into $pCS2^{+}$ expression vector.

hCRIPTO

The plasmid containing the *hCRIPTO* gene cloned in the pEF6/V5-His TOPO cloning vector was obtained as a gift from Dr. Michael R. Kuehn from the Center for Cancer Research National Cancer Institute. To subclone the *hCRIPTO* sequence into the expression vector $pCS2^{+}$ we used the same strategy as for the *hNODAL* gene. The only difference was the enzymes that were used for cut the *hCRIPTO* gene from the initial vector, in which we used the BamH1 and

XbaI. Ultimately, the purified plasmid was sequenced to confirm that the *hCRIPTO* was successfully cloned into pCS2^{+} expression vector.

hDAND5

The encoding full-length human DAND5 (IRCMp5012B019D) was purchased from Source BioScience LifeSciences, Nottingham, United Kingdom, cloned in the pCR4-TOPO vector, which isn't an expression vector. To subclone the *hDAND5* sequence into the expression vector pCS2^{+} we used the same strategy as for the *hNODAL* gene and *hCRIPTO* butwecut the pCR4-TOPO vector and pCS^{2+} with EcoRI (ThermoFisherScientific). To prevent self-ligation dephosphorylation of the pCS^{2+} plasmid with alkaline phosphatase was used. Ultimately, the purified plasmid was sequenced to confirm that the *hDAND5* was successfully integrated into pCS2^{+} expression vector.

hDAND5 c.455G>A

The c.455G>A variant of *DAND5* was generated from the hDAND5 pCS^{2+} plasmid by site-directed mutagenesis using the QuikChange kit (Stratagene, La Jolla, CA, USA) with the following pair of primer: Forward: 5´ GTATGCCTGCTCACAAGCGTTGGG 3´; Reverse: 5´CCCAACGCTTGTGAGCAGGC 3´ according to the manufacturer's instruction. Ultimately, the purified plasmid was sequenced to confirm that the *hDAND5 c.455G>A* alteration was successfully introduced into the pCS2^{+} expression vector.

6.2. Nodal-dependent luciferase assay

Human embryonic kidney 293T cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and seeded overnight at 50–60% confluence in 96-well plates. Next day, the cells were transiently co-transfected with Lipofectamine 2000 (Lifetechnologies) and the following plasmids: pAR3-lux, that contain three copies of the activin response element (ARE) in the front of a luciferase reporter, which is specifically activated by

nodal signaling; hDAND5; hDAND5 R152H; hNODAL; hCRIPTO, a Nodal co-receptor necessary for the proper activation of the Nodal signaling; hFoxH1, the major transcriptional transducer of nodal signaling; CMV-β-Gal plasmid as control, and various amounts of $pCS2^{+}$ vector to maintain a constant amount of total DNA. Transfections were performed in triplicate and in three independent experiments with a total amount of 100ng DNA per well with the following conditions: (1) Cells transfect with 6ng of pAR3-lux; (2) Cells transfected with 6ng of AR3-Lux and 1,5ng of FoxH1 plasmid; (3) Cells transfected with 6ng of AR3-Lux, 1,5ng of FoxH1 plasmid and 17ng of hCRIPTO plasmid; (4) Cells transfected with 6ng of AR3-Lux, 1,5ng of FoxH1 plasmid, 17ng of hCRIPTO plasmid and 17ng of hNODAL plasmid; (5) Cells transfected with 6ng of AR3-Lux, 1,5ng of FoxH1 plasmid, 17ng of hCRIPTO plasmid, 17ng of hNODAL plasmid and 17ng of hDAND5 plasmid; (6) Cells transfected with 6ng of AR3-Lux, 1,5ng of FoxH1 plasmid, 17ng of hCRIPTO plasmid, 17ng of hNODAL plasmid and 17ng of mutant hDAND5 plasmid. Twenty-four hours after transfection, cell lysates were prepared, luciferase activity measured and the results were normalized to b-galactosidase control.

7. *In silico* analysis of the intronic variants found in *NODAL* and *PITX2C* genes.

To study if the intronic alterations found in *NODAL* and *PITX2C* genes could alter the normal process of splicing, we decide to use the online software Human Splicing Finder. This software is a tool that predict the effects of mutations on splicing signals or identify splicing motifs in any human sequence. It contains all available matrices for auxiliary sequence prediction as well as new ones for binding sites of the 9G8 and Tra2-β Serine-Arginine proteins and the hnRNP A1 ribonucleoprotein and includes new algorithms derived from the Universal Mutation Database (UMD) to allow the evaluation of the strength of 5′ss, 3′ss and branch points. In addition, in order to identify cis-acting elements it includes already published algorithms, such as the RESCUE-ESE and ESE-Finder as well as new algorithms designed to use available or newly created matrices. To allow the study of virtually any human sequence, HSF includes all

genes and alternative transcripts as well as intronic sequences that were extracted from the Ensemble human genome database.

8. Generation and characterization of human iPSC line

8.1. Ethics statement and informed consent

All the experimental protocols included in the present study were approved by the Institutional Ethical Committee of the NOVA Medical School according to European Union legislation and written informed consent was obtained from guardian patient prior to the collection of the urine sample (Supplementary Material 5 as example). The biological samples collected were anonymized to ensure the confidentiality and privacy of patients and data.

8.2. Isolation of cells from urine

Isolation of cells from urine was performed as described by (Zhou et al., 2012). Briefly, to obtain urinary cells, approximately 50 to 120 ml of urine were collected and centrifuged 400g for 10 min at room temperature. The supernatant was discarded, leaving only ~1ml, and the cell pellet was washed with 10 ml of washing buffer (PBS containing 100 U/ml of penicillin, 100 µg/ml streptomycin (Lonza) and 500 ng/ml of amphotericin B. Following another centrifugation at 200g for more 10 min, the pellet was resuspended in 2 ml of primary medium (Dulbecco's Modified Eagle Medium (DMEM): nutrient Mixture F-12 (DMEM/F-12) (1:1), supplemented with 10% (vol/vol) Fetal Bovine Serum (FBS) (Sigma), 100 U/ml penicillin, 100 µg/ml streptomycin (Lonza), the REGM SingleQuot kit supplements (Lonza) and 2.5 µg ml − 1 amphotericin B (Lonza)) and seeded in one well of a 12-well plate (Corning) coated with 0,1% (w/t) gelatin (Sigma). In the following three days 1 ml of primary medium was added and at day four after plating, most of the medium was removed and replaced by proliferation medium (Renal Cell Basal Medium (REBM) supplemented with REGM BulletKit (Lonza)). Subsequently, half of the culture medium was changed every day and after 4–6 days small 2–4 cell colonies started to appear. Cells were rounded or elongated which represent type I or type II of renal

epithelial (RE) cells, respectively. Once cells reached 90% density, they were passaged and expanded, using TrypLE Select (Life Technologies), to a 6-well plate (Corning).

8.3. Non-integrative reprogramming and establishment of iPSC lines

Reprogramming was performed using Sendai virus vectors (CytoTune-iPS 2.0 Reprogramming Kit, Life Technologies). Briefly, 7,5 × 10^5 cells, seeded in one well of a 6-well plate coated with gelatin and cultured in RE proliferation medium, were transduced using a 1,5 MOI (multiplicity of infection) of each of the 3 CytoTune® 2.0 vectors, KOS (Klf4–Oct3/4–Sox2), hc-Myc, and hKlf4. After 24h, the medium was replaced by fresh RE proliferation medium and cells were cultured for seven days with medium changes every other day. On day 8 after transduction, cells were passaged using TrypLE Select and seeded with a density of 6 × 10^5 cells/dish in RE proliferation medium onto a 100 mm culture dish (Corning) coated with Geltrex (Gibco, Thermo Fisher Scientific). On the next day medium was changed to Essential 8 (E8) flex medium (Gibco, Thermo Fisher Scientific) supplemented with100 U/ml penicillin, 100 µg/ml streptomycin and replaced every day until hiPSC colonies appeared. 17 days after infection, individual colonies were picked using a pipette tip and expanded to 24-well plates coated with Geltrex in E8 flex medium supplemented with100 U/ml penicillin, 100 µg/ml streptomycin and revitacell (Gibco, Thermo Fisher Scientific). In the next day, the medium of all the colonies picked was replace by flesh E8 flex medium without revitacell.

8.4. Culture of iPSC

The hiPSC were cultured in Geltrex (Gibco, Thermo Fisher Scientific) 6-well coated plates with Essential E8 flex medium (Gibco, Thermo Fisher Scientific). Cells were routinely passaged using 0,5 ml of TrypLE Select (Life Technologies) every two days at a ratio of ~1:10.

8.5. Sequencing analysis

Total DNA from the patient's hiPSCs was extracted using the Isolate Genomic DNA mini kit (BIOLINE) according to the manufacturer's instructions for DNA isolation from eukaryotic cells. Subsequently, amplification by PCR of the exon 2 of *DAND5* gene, containing the c.455G>A alteration, was carried out using the following primers (Table 2.4): DAND5 Ex2F: 5′-GGAAGTGGACAGGTGATTATCC-3′ and DAND5 Ex2R: 5′-CACGTCTTTCTTGGTCCATCTC-3′). Following PCR amplification, quality of amplification products was analyzed using a 1% agarose gel and direct sequencing of amplicon was performed on both strands using a ABI 3730xl DNA Analyzer in STAB VIDA (http://www.stabvida.com/).

8.6. Quantitative polymerase chain reaction (qPCR) analyses

Total RNA was extracted and purified from patient and control iPSC lines, H9 human embryonic stem cell line as a positive control and exfoliated renal epithelial cells using TRizol Reagent (Invitrogen) plus Direct-zol™ RNA MiniPrep kit (Zymo Research) according to the manufacturer's instructions. 1 μg of total RNA was used to synthesize cDNA with the RevertAid RT Reverse Transcription Kit (Thermo Scientific). cDNA was then diluted 1:10, with nuclease-free water, and 2 μl were used to quantify by qRT-PCR the expression of the endogenous pluripotency associated genes (*NANOG, SOX2, KLF4, OCT4 and NODAL*). qRT-PCR was performed on a Applied Biosystems® 7500 Real-Time PCR machine and reactions were run with SYBR® Select Master Mix (Applied Biosystems) plus the primers listed in table 2.4. CT-values were normalized to the geometric mean of the two housekeeping genes GAPDH and β-actin and with H9 human embryonic stem cell (ESC) line as reference (set to 1).

Table 2.4. Primers used to characterize the patient and control derived iPSC lines.

Target		Primer Sequence (5´to 3´)	Product size (bp)
Sequencing analysis			
DAND5 exon 2	Forward Reverse	GGAAGTGGACAGGTGATTATCC CACGTCTTTCTTGGTCCATCTC	399
Pluripotency analysis			
NANOG	Forward Reverse	CATGAGTGTGGATCCAGCTTG CCTGAATAAGCAGATCCATGG	191
OCT4	Forward Reverse	AGTGAGAGGCAACCTGGAGA ACACTCGGACCACATCCTTC	110
SOX2	Forward Reverse	GGGAAATGGGAGGGGTGCAAAAGAGG TTGCGTGAGTGTGGATGGGATTGGTG	151
Klf4	Forward Reverse	ACCAGGCACTACCGTAAACACA GGTCCGACCTGGAAAATGCT	79
NODAL	Forward Reverse	GGGCAAGAGGCACCGTCGACATCA GGGACTCGGTGGGGCTGGTAACGTTTC	234
GAPDH	Forward Reverse	CTGGTAAAGTGGATATTGTTGCCAT TGGAATCATATTGGAACATGTAAACC	81
β-actin	Forward Reverse	GCAAAGACCTGTACGCCAAC AGTACTTGCGCTCAGGAGGA	144

8.7. Fluorescent immunocytochemistry

Undifferentiated patient and control derived iPSCs were grown on Geltrex coated cover slips, fixed at room temperature with 1% PFA (Sigma-Aldrich) in PBS for 15 min and then permeabilized with 0.1% Triton X-100 (Sigma-Aldrich) in PBS for 30 min. Non-specific binding was blocked with 1% BSA for 30 min. The cells were then incubated with primary antibodies overnight at 4 °C. To assess the expression of pluripotent markers, immunofluorescence staining was performed with primary antibodies against OCT4 (1:400), NANOG (1:200) and SSEA4 (1:200) as shown in table 2.5. After three washes with 0.1% Triton X-100 in PBS for 5 min each, the cells were incubated with fluorescence-conjugated secondary antibodies (AlexaFluor 488: Donkey anti-mouse 1:400, Invitrogen and AlexaFluor 488: Donkey anti-rabbit 1:400, Invitrogen) for 60 min, and mounted on glass slides with mounting medium containing DAPI (Life technologies). Images were acquired with Zeiss Axio Imager.Z2 microscope (Carl Zeiss).

Table 2.5. Primary antibodies used for immunofluorescence

	Antibody and host species	Dilution	Company and catalog number
Pluripotency analysis	Rabbit anti-NANOG	1:200	Abcam, ab21624
	Rabbit anti-OCT4	1:400	Abcam,ab19857
	Mouse anti-SSEA4	1:200	Abcam,ab16287

8.8. Karyotyping

Karyotyping was performed according to the protocols from Laboratório de Citogenética e Genómica, Faculdade de Medicina da Universidade de Coimbra.

Chapter 3

Results/Discussion

Part I

Functional study of DAND5 variant in patients with Congenital Heart Disease and laterality defects

Fernando Cristo [1-4], José M. Inácio[1], Salomé de Almeida[5], Patrícia Mendes[6], Duarte Saraiva Martins[7], José Maio[6], Rui Anjos[7] and José A. Belo[1]

[1]Stem Cells and Development Laboratory, CEDOC, NOVA Medical School / Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal.

[2]Center for Biomedical Research (CBMR), Universidade do Algarve, Portugal.

[3]PhD Program in Biomedical Sciences, Universidade do Algarve, Portugal.

[4]Regenerative Medicine Program, Biomedical and Medicine Sciences Department, Universidade do Algarve, Portugal.

[5]Medical Genetics Service, Centro Hospitalar Lisboa Central (CHLC), EPE, Lisboa, Portugal.

[6]Departamento Materno-Infantil, Centro Hospital do Algarve, EPE, Faro, Portugal.

[7]Hospital de Santa Cruz, Centro Hospitalar Lisboa Ocidental, Lisboa, Portugal

(Manuscript submitted to: BMC Medical Genetics)

Authors' contribution: Conceived and designed the experiments: FC, JMI, SA, JB; Diagnosis of patients: PM, JM, RA; Patient recruitment, sample collection and clinical data collection: FC, JMI, SA, PM, JM, RA, DM; Analyzed the data: FC, JMI, SA, JB; Performed the experiments: FC, JMI, SA; Contributed to writing the manuscript: FC, JMI and JB. All authors read and approved the final manuscript.

1. Abstract

Background: Perturbations on the Left-Right axis establishment lead to laterality defects, with frequently associated Congenital Heart Diseases (CHDs). Indeed, in the last decade, it has been reported that the etiology of isolated cases of CHDs or cases of laterality defects with associated CHDs is linked with variants of genes involved in the Nodal signaling pathway. **Methods:** With this in mind, we analyzed a cohort of 38 unrelated patients with Congenital Heart Defects that can arise from initial perturbations in the formation of the Left-Right axis. Genomic DNA was extracted from buccal epithelial cells, and variants screening was performed by PCR and direct sequencing. A Nodal-dependent luciferase assay was conducted in order to determine the functional effect of the variant found. **Results:** In this work, we report two patients with a *DAND5* heterozygous non-synonymous variant (c.455G>A) in the functional domain of the DAND5 protein (p.R152H), a master regulator of Nodal signaling. Patient 1 presents left isomerism, ventricular septal defect with overriding aorta and pulmonary atresia, while patient 2 presents ventricular septal defect with overriding aorta, right ventricular hypertrophy and pulmonary atresia (a case of extreme tetralogy of Fallot phenotype). The functional analysis assay showed a significant decrease in the activity of this variant protein when compared to its wild-type counterpart. **Conclusion:** Altogether, our results provide a new insight into the molecular mechanism of the laterality defects and related CHDs, priming DAND5 as a candidate determinant for CHDs in humans.

Keywords: DAND5, Congenital Heart Diseases, laterality defects, Nodal signaling, allelic variation.

2. Background

Heart morphogenesis is a complex process involving multiple cell types (cardiac and non-cardiac) thus requiring a precise control of all molecular and cellular mechanisms (Srivastava, 2001). Subtle deviations in heart development lead to Congenital Heart Diseases (CHDs), which are the most prevalent form of birth defect (postnatal incidence of 0.8 % of all newborns) and the leading non-infectious cause of death in the first year of life (Andersen, Troelsen Kde, & Larsen, 2014; Koefoed, Veland, Pedersen, Larsen, & Christensen, 2014). The etiology of CHD is incompletely understood, thought to be multifactorial, involving multiple genetic and environmental factors and the likely interactions between those factors (Pediatric Cardiac Genomics et al., 2013). Nevertheless, in the last years, CHDs are seen as disorders resulting from alterations in genes that control basic embryonic processes between the third and eight week of gestation, when the majority of the cardiovascular structures develop and when the integrity of the left-right (LR) axis is established (Cecchetto et al., 2010; Icardo, Garcia Rincon, & Ros, 2002). In fact, in the majority of cases presenting laterality disorders, a complex heart malformation is also observed, suggesting that CHDs could be due to a laterality defect in the formation of the heart (Levin, 2004). In others words, even in children that born with CHDs and that do not present others extracardiac laterality abnormalities, the etiology of these CHDs could arise from subtle faults in the formation of the left-right axis.

The molecular control of LR axis development is mainly achieved by the Nodal signaling pathway (Shiratori & Hamada, 2006). Studies in mice revealed that during gastrulation, *Nodal*, a growth factor from the TGF-ß family, is asymmetrically expressed in the node, and this expression is expanded and amplified in the left-lateral plate mesoderm (L-LPM), but inhibited in the right-lateral plate mesoderm (R-LPM) (Shiratori & Hamada, 2014). Nodal itself induces its intracellular signaling by binding to type I (ALK4 and ALK7) and type II (ActRIIa and ActRIIb) serine-threonine kinase receptors in the presence of one EGF-CFC (epidermal growth factor-Cripto-FRL1-Cryptic) family co-receptor, Cripto or Cryptic. This leads to the phosphorylation of regulatory Smads (Smad2 and Smad3), their association with Smad4, translocation into the nucleus, and to the interaction with the transcription factor FoxH1, which at

the end activates the expression of Nodal target genes, in particular *Pitx2* (Shiratori & Hamada, 2014). Pitx2 is a transcription factor that regulates the fate of cells that will form the visceral organs, including the heart (Nakamura & Hamada, 2012; Shiratori & Hamada, 2006). The impairment of the Nodal signaling in the lateral plate mesoderm, by variants in the genes involved in the pathway such as *ActRIIB* (MIM #602730) (R. Kosaki et al., 1999; Ma, Selamet Tierney, Lee, Lanzano, & Chung, 2012), *LEFTY 1* (MIM #603037) (X. Deng et al., 2014; K. Kosaki et al., 1999), *LEFTY 2* (MIM #601877) (X. Deng et al., 2014; K. Kosaki et al., 1999), *CRYPTIC/CFC1* (MIM #605194) (Goldmuntz et al., 2002; B. Wang, Wang, et al., 2011), *CRIPTO/TDGF1* (MIM #187395) (Roessler et al., 2008; B. Wang, Yan, et al., 2011), *FOXH1* (MIM #603621) (De Luca et al., 2010; B. Wang et al., 2010), *NODAL* (MIM #601265) (Mohapatra et al., 2009; Roessler et al., 2009) and *PITX2C* (MIM # 601542) (J. Wang et al., 2013; F. Yuan et al., 2013) had already been associated to laterality disorders and/or CHDs in humans. For example, the heterozygous variant p.G260R in NODAL was associated with transposition of the great arteries (TGA), atrial and ventricular septal defects, double outlet right ventricle (DORV) and/or heterotaxy (Mohapatra et al., 2009). On the other hand, the NODAL p.S60I variant has been associated with the tetralogy of Fallot (TOF) phenotype (Roessler et al., 2009). Additionally, several variations in Nodal co-receptor CFC1 were linked to TOF, TGA, DORV and/or laterality defects (Bamford et al., 2000; Goldmuntz et al., 2002; Roessler et al., 2008).

Recently, we reported that a DAN family member, the Nodal-antagonist Cerberus-like 2 (Cerl2) – DAND5 in humans (MIM #609068; refseq - NM_152654.2) – controls Nodal signaling at the mouse node (Belo et al., 2009; Oki et al., 2009) and the transmission of LR asymmetry information to the left-lateral plate mesoderm in a precise time window (Inacio et al., 2013; Kawasumi et al., 2011). Furthermore, the absence of *Cerl2* in *Cerl2* KO embryos leads to a range of laterality defects and/or cardiovascular malformations such as incomplete atrial and ventricular septation, TGA, DORV, randomized positioning of the cardiac apex, ventricular hypertrophy and/or heterotaxy of the abdominal organs, and a significant mortality rate within a few hours after birth is observed (Araujo, Marques, & Belo, 2014; Marques et al., 2004).

Here, in order to advance the understanding of DAND5 in the genetic etiology of laterality disorders and associated CHDs, we conducted a genetic screening to identify *DAND5* allelic variants in children affected with CHDs and/or LR asymmetry defects. We identified a *DAND5* allelic variant in two patients, which reduces the inhibitory activity of DAND5. We propose that this variant might contribute to the risk of congenital heart disease and/or laterality defects.

3. Results and discussion

We studied thirty-eight well-defined probands meeting the inclusion criteria for congenital heart defects with a possible left-right abnormality cause (Table 3.1). These defects include atrial septal defects, ventricular septal defects, atrioventricular septal defects, conotruncal defects – pulmonary atresia/stenosis, tetralogy of Fallot, transposition of the great arteries, double inlet left ventricle, double outlet right ventricle, aortic coarctation, overriding aorta – left isomerism, right isomerism, dextrocardia and other cardiac malformations (right sided aortic arch, inferior vena cava interruption, univentricular heart, aortic stenosis, mitral valve stenosis, tricuspid atresia, hypoplastic right ventricle, hypoplastic left heart, major aortopulmonary collateral artery, persistent left superior vena cava, single ventricle and total anomalous pulmonary venous connection). Pulmonary atresia is the more frequent cardiac phenotype, being present in 37% of the cases. Nevertheless, this malformation was rarely found as an isolated case and, in almost all the cases, the patients present another cardiac malformation. Indeed, typically, each subject had complex defects with more than one abnormality, which is in agreement with the complexity, heterogeneity and diversity of this trait and represents, in many cases, severe forms of CHDs. Moreover, seven of the patients presented obvious defects resulting from abnormal left-right patterning, including four patients with left isomerism, one patient with right isomerism, one with *situs inversus totalis* and another with partial visceral *situs inversus*.

Table 3.1. Clinical characteristics of the patients genotyped

Prevalence of different types of CHD	**Nº of Patients (n=38)**
Atrial Septal Defect	2 (~5%)
Ventricular Septal Defect	11 (~29%)
Atrioventricular septal defects	8 (~21%)
Conotruncal defects	
Pulmonary atresia/stenosis	14 (~37%)
Tetralogy of Fallot	7 (~18%)
Transposition of the great arteries	8 (~21%)
Double inlet left ventricle	2 (~5%)
Double outlet right ventricle	1 (~3%)

Aortic coarctation	1 (~3%)
Overriding aorta	1 (~3%)
Left isomerism	4 (~11%)
Right isomerism	1 (~3%)
Dextrocardia	3 (~8%)
Other cardiac malformations	20 (~53%)
Extracardiac abnormalities	
Situs inversus totalis	1 (~3%)
Visceral *situs inversus*	1 (~3%)
Asplenia	1 (~3%)

Other cardiac malformations included right sided aortic arch, inferior vena cava interruption, univentricular heart, aortic stenosis, mitral valve stenosis, tricuspid atresia, hypoplastic right ventricle, hypoplastic left heart, major aortopulmonary collateral artery, persistent left superior vena cava, single ventricle, total anomalous pulmonary venous connection. Note: Almost all the patients have more than one type of CHD.

The DNA sequencing analysis of this unrelated children cohort revealed two patients, patient 1 and patient 2, displaying the same variant in *DAND5* gene.

For validation, all exons were sequenced using forward and reverse internal primers flanking the missense variant. Clinically, patient 1 presented ventricular septal defect with overriding aorta, pulmonary atresia and left isomerism, while patient 2 presented overriding aorta, ventricular septal defect, right ventricular hypertrophy and pulmonary atresia (a case of extreme tetralogy of Fallot phenotype; Table 3.2).

Table 3.2. Clinical and molecular findings of patient 1 and 2.

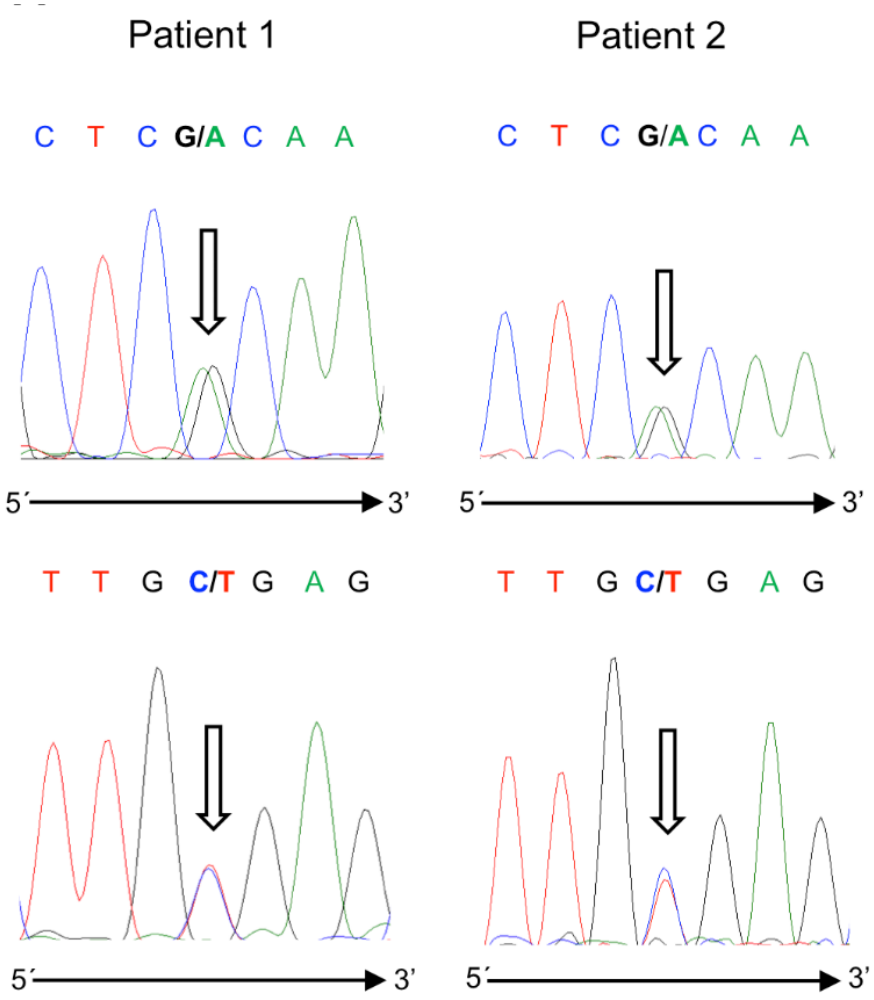
	Genotype	**Protein alteration**	**Phenotype**
Proband 1	*DAND5* c.455GA	DAND5 p.R152H	Left isomerism; VSD with overriding aorta; Pulmonary atresia
Proband 2	*DAND5* c.455GA	DAND5 p.R152H	Tetralogy of Fallot; Pulmonary atresia

VSD – ventricular septal defect
Wild type DAND5 refseq. - NM_152654.2

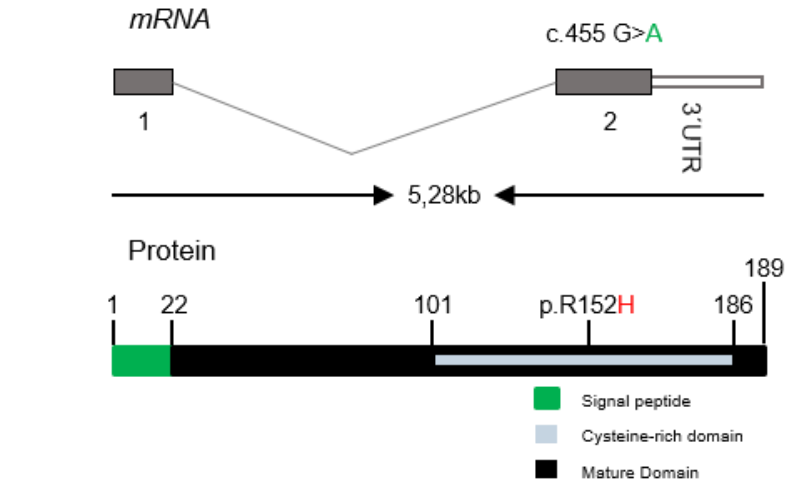
The variant was identified in both patients as heterozygous non-synonymous c.455G>A (Figure 3.1A). The phenotypically normal mother of patient 2 also presented heterozygosity the *DAND5* c.455G>A variant, suggesting incomplete penetrance. Unfortunately, the father of patient 2 and the parents of patient 1 were not available for sequencing.

The c.455G>A variant in exon 2 of *DAND5* results in a substitution of an Arginine (R) residue for an Histidine (H) (p.R152H) located in a highly conserved region of the cysteine-rich DAN domain (Figure 3.1B).

A)



B)




```
DAND5_HUMAN          141  PLVLCNSCMPARKRWAPVVLWCLTGSSAS  169
DAND5_CHIMPANZEE     141  PLVLCNSCMPARKRWAPVVLWCLTGSSAS  169
DAND5_RAT            136  PVVLCNSCVPSRRRWTSVALWCGAGQSAS  164
DAND5_MOUSE          137  PVVFCNSCVPARKRWTSVTLWCGAGQLAS  165
DAND5_XENOPUS        156  LSQPCTSCAPSRSRRISLPLRCRSGHLAW  184
DAND5_ZEBRAFISH      188  QKAQCTRCGPSRARSVLLHLRCGSEVRER  216
                               *. * *:* *   : * * :
```

Figure 3.1 – Genomic localization of *DAND5* 455 allelic variant.

A) Forward and reverse DNA sequence chromatograms of patient 1 and 2 showing the c.455G>A allelic variant (highlighted by arrows). **B)** Schematic diagram of DAND5 structure with approximated localization of the variant identified in this study and cross-species sequence alignment of DAND5 showing the conservation of amino acid R152 (highlighted by red box).

The available protein function prediction algorithms estimated differing functional effect of this variant in DAND5 (Table 3.3). In 2013, Polyphen-2 algorithm (http://genetics.bwh.harvard.edu/pph2/index.shtml) indicated that this variant was possibly damaging for the DAND5 protein but, using an up-to-date version Polyphen-2 we now observed that the DAND5 p.R152H variant had no effect on the DAND5 protein (classified as benign). Moreover, Mutation Taster algorithm (www.mutationtaster.org) predicted that this variant is a polymorphism that might affect DAND5 protein features, namely alterations in the cysteine-rich domain. PROVEAN software (http://provean.jcvi.org/) predicted a deleterious effect (score -2.84) on the normal function of the DAND5 protein, and I-Mutant 3.0 predicted that the a.a. alteration will reduce largely the stability of the protein (a score of −1.28 a reliability index of 9; disease reliability index of 3).

Nevertheless, these algorithms have only an accuracy of 65-80% in the prediction of known disease missense variants, being even less accurate when examining missense variants with milder effect (Richards et al., 2015).

Table 3.3. *In silico* protein prediction effect of DAND5 p.Arg152His alteration

	Score	Prediction effect
Polyphen-2	0,49 (2013) / 0,120 (2016)	Possible damaging / Benign
Mutation Taster	Polymorphism	Might affect protein features
PROVEAN	-2,84 (cutoff= -2,5)	Deleterious
I-Mutant 3.0	-1,28 Kcal/mol	Reduced stability of the protein and related to disease

Taking this in attention, and to clarify the function of this protein variant, we reproduced the p.R152H variant in an expression vector containing the human *DAND5* cDNA, and we evaluated the in vitro effect of p.R152H in the regulation of Nodal activity by using a well-established Nodal-dependent luciferase assay as readout. This assay consists in a luciferase reporter gene under the control of three activin-responsive elements promoter, pAR3-lux, which is transcriptionally activated by Nodal signaling (Mohapatra et al., 2009) In addition, we used the human version of the main intervenient genes, *NODAL*, *CRYPTIC/CRIPTO*, *FOXH1*, all necessary for a reliable transduction of the levels of Nodal signaling in this type of assay. The results, from triplicates of three independent experiments, showed that the p.R152H DAND5 variant causes a significant reduction in the normal NODAL-inhibitory activity of DAND5 (Figure 3.2). Since p.R152H is located in the functional domain of the DAND5 protein and because DAND5 inhibits NODAL by direct binding (Marques et al., 2004), we hypothesize that the protein alteration caused p.R152H variation might interfere with the interaction DAND5-NODAL. Therefore, we further extrapolate that, most probably, during the embryonic development of patients with p.R152H DAND5, which has a DAND5 molecule with reduced activity, the overall activity of NODAL on the node is above the proper levels. This abnormal NODAL activity on the node might be reflected then in an excess of NODAL activity on the left-lateral plate mesoderm and/or an ectopic expression of *NODAL* and *PITX2* on the right-lateral plate mesoderm of the embryo.

Interestingly, the phenotypes observed in these two patients, are consistent with an ectopic expression of *NODAL* and *PITX2* on the right side of the embryo.

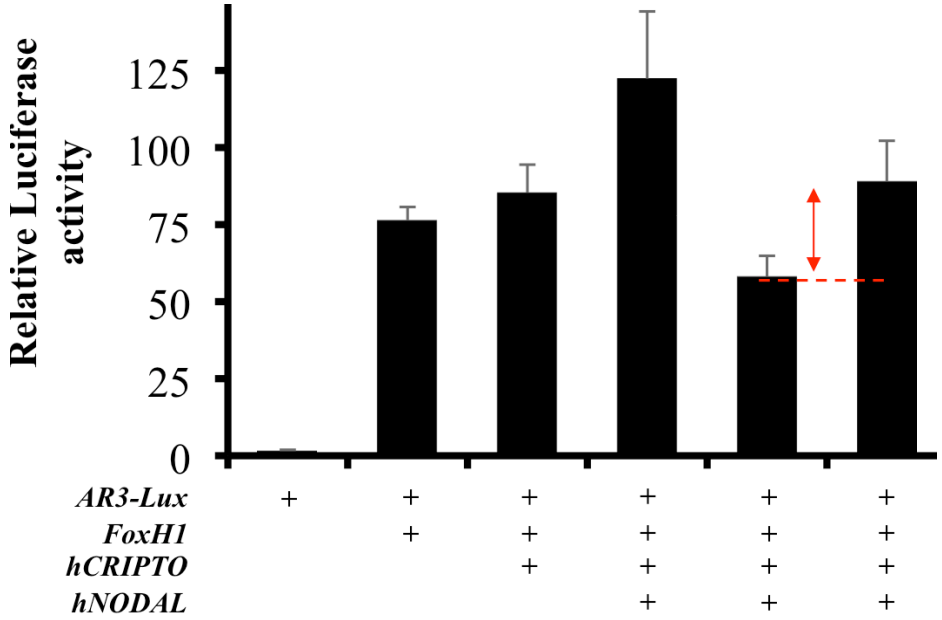


Figure 3.2 – Functional analysis of DAND5 variant.

The identified p.R152H variant was evaluated through a luciferase assay as readout of Nodal signaling. In this assay a luciferase reporter gene under the control of three activin-responsive elements promoter, pAR3-lux, which is transcriptionally activated by Nodal signaling, was used. The results, from triplicates of three independent experiments, showed that the p.R152H DAND5 variant leads to an increase of NODAL signaling when compared to WT DAND5 protein

Tetralogy of Fallot is a malformation of the cardiac outflow tract (OFT) that comprises ventricular septal defect, over-riding of the aortic valve, right ventricular outflow tract obstruction (pulmonary stenosis or in its most severe form, pulmonary atresia), and right ventricular hypertrophy (Goodship et al., 2012). The OFT forms during heart looping as a result of a coordinated process uncompressing proliferation, differentiation, and migration of the pre-cardiac cells of the secondary heart field (SHF). These cells are known to be sensitive to the asymmetrically left expression of the *NODAL* and *PITX2* in the SHF (Nowotschin et al., 2006). Any imbalance in the LR patterning that might lead to an ectopic right-sided expression and loss-of-function of *PITX2* results in reversed heart looping and abnormal shifting of the OFT (Martin, Amendt, & Brown, 2010; Ramsdell, 2005). Suggesting that anomalous LR body patterning has the tendency to develop severe congenital heart defects. Moreover, cases of fully bilateral expression of *NODAL* and *PITX2* in the cells of both left and

right lateral plate mesoderm lead to left isomerism phenotypes, in which both atria present a left morphology, normally associated to atrioventricular septal defects (Shiraishi & Ichikawa, 2012). In agreement, it has been observed that the absence of *Cerl2* (*DAND5* mouse homolog) in *Cerl2* KO mice leads to several cardiovascular malformations, including incomplete atrial and ventricular septation, conotruncal defects, ventricular hypertrophy, and/or to laterality defects like left isomerism and thoracic heterotaxy or *situs inversus* (Araujo et al., 2014; Marques et al., 2004).

Interestingly, the DAND5 p.R152H variant had been already annotated in the NCBI (http://www.ncbi.nlm.nih.gov) and Ensemble (http://www.ensembl.org) databases as Single Nucleotide Variation (SNV - rs45513495), with a minor allele frequency (MAF) of ~ 1% (low-frequency variant; Supplementary. Material 6) and described in ExAC database with an allele frequency of 0.01043 in all populations but never associated with disease, and even present in apparently 16 normal homozygous. Indeed, the mother of patient 2 is phenotypically normal but displays the *DAND5* p.R152H variant. Nevertheless, we cannot exclude the presence of a mild or undiagnosed heart defect on the mother of patient 2, or on the subjects of the genome projects, since detailed examination of these subjects is rarely available. In addition, the ExAC database and the American College of Medical Genetics and Genomics warns to the fact that in these population databases the presence of individuals with subclinical diseases among the supposed healthy individuals leading to minor discrepancies in some variant frequencies cannot excluded (Richards et al., 2015). For example, some phenotypes of CHD such as small atrial septal defects and ventricular defects may be asymptomatic and undetected throughout life (Huang et al., 2010), or cases of apparent normal individuals whose laterality defects were only detected during adulthood (Channabasappa, Mohan, & Sarma, 2013). Besides that, the degree of the functional Nodal pathway impairment might be explained by factors like differential allelic expressivity, an incomplete penetrance, the presence of genetic modifiers, and/or environmental factors. Interestingly, ~40% of the *Cerl2* KO mice become normal adults, without any CHD and/or laterality defect (Marques et al., 2004).This suggests that the distribution and strength of the Cerl2 inhibitory

signal and ultimately the imbalance of the dose-sensitive Nodal signaling might be sensitized by additional perturbations during LR patterning.

Curiously, the sensitization of an imbalanced of Nodal signaling and its correlation to differences in phenotypes was also observed in the other genes of Nodal signaling cascade (Goldmuntz et al., 2002; Mohapatra et al., 2009; Roessler et al., 2008). For example, *cryptic* (CFC1 mouse homolog) full-mutant mice presents randomization of abdominal *situs* and complex cardiac malformations such as atrial and septal defects, dextrocardia, DORV and TGA, but normal configuration of the OFT was seen in one mutant (Gaio et al., 1999). In humans, it was identified and classified as disease variant due to its loss-of-function effect, an heterozygous single-base-pair deletion (G174 del1) in *CFC1* in two patients both with dextrocardia and TGA but one of the patients presented right isomerism, whereas the other patient presented left isomerism. This *CFC1* variant was also reported in one subject with sporadic DORV, ventricular septal defect, aortic arch hypoplasia (Goldmuntz et al., 2002) and in a patient with tetralogy of Fallot (Roessler et al., 2008). Interestingly, likewise we observed in p.R152H DAND5 variant, the mothers of these *CFC1* variant patients were found to carry the G174del1 alteration but none presented any clinical phenotype, suggesting incomplete penetrance. Accordingly, similar cases of reduced penetrance have been reported in CHD by others (Bamford et al., 2000; Megarbane et al., 2000).

Although the DAND5 p.R152H did not attest a specific genotype-phenotype link, our functional results support a model in which the proper levels of Nodal signaling, on both left and right lateral plate mesoderm, are particularly sensitive to gene dosage effects for all genes involved in the pathway. The imbalance in dosage-sensitive Nodal signaling is a final common way for laterality defects and associated CHDs.

In conclusion, our findings firstly prime DAND5 as one of multiple candidate determinants involved in the complex trait mechanism that cause CHDs, associating the human vulnerability for CHDs with the reduced activity of heterozygous *DAND5* p.R152H variant.

Part II

Genetic Screening of other Nodal signaling pathway genes

Fernando Cristo [1-4], José M. Inácio[1], Salomé de Almeida[5], Patrícia Mendes[6], Duarte Saraiva Martins[7], José Maio[6], Rui Anjos[7] and José A. Belo[1]

[1]Stem Cells and Development Laboratory, CEDOC, NOVA Medical School / Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal.

[2]Center for Biomedical Research (CBMR), Universidade do Algarve, Portugal.

[3]PhD Program in Biomedical Sciences, Universidade do Algarve, Portugal.

[4]Regenerative Medicine Program, Biomedical and Medicine Sciences Department, Universidade do Algarve, Portugal.

[5]Medical Genetics Service, Centro Hospitalar Lisboa Central (CHLC), EPE, Lisboa, Portugal.

[6]Departamento Materno-Infantil, Centro Hospital do Algarve, EPE, Faro, Portugal.

[7]Hospital de Santa Cruz, Centro Hospitalar Lisboa Ocidental, Lisboa, Portugal

Authors' contribution: Conceived and designed the experiments: FC, JMI, SA, JB; Diagnosis of patients: PM, JM, RA; Patient recruitment, sample collection and clinical data collection: FC, JMI, SA, PM, JM, RA, DM; Analyzed the data: FC, JMI, SA, JB; Performed the experiments: FC, JMI, SA; Contributed to writing the manuscript: FC, JMI and JB. All authors read and approved the final manuscript.

1. Abstract

Perturbations in the establishment of the Left-Right axis lead to laterality defects, which in most of the times have associated complex CHDs comorbidities. Therefore, genes of the Nodal signaling pathway have been reported to be implicated in the etiology of CHDs and/or laterality defects. Considering the high conservation of genetic pathways regulating cardiac development in mouse and human, the study of the human genes involved in this pathway will bring us new data of the cellular and molecular mechanisms that underlie cardiac asymmetry on the CHD and the laterality defects. In the first chapter of results/discussion we identified and characterize a DAND5 variant that seems to function as a risk allele for CHDs and/or laterality defects. Nevertheless, we cannot make a clearly genotype-phenotype correlation and the genetic data points to an incomplete penetrance, suggesting the presence of modifier variants or other alterations in genes which DAND5 normally interacts or in genes of the Nodal signaling pathway. Therefore, in addition to the *DAND5* gene, we analyzed the *NODAL*, *CFC1* and *PITX2C* genes in the same cohort of 38 patients presenting CHD and/or laterality defects. We have found two coding change alterations, one in *NODAL* and one in *CFC1* and three intronic variants in the genomic sequence of *PITX2C* and *NODAL*. *In silico* analysis of the intronic variants using the splicing prediction software Human Splicing Finder indicates that these variants could somehow affect the splicing process of *NODAL* and *PITX2C* mRNA. Regarding the coding sequence alterations, the *NODAL* exonic variant was already described by other authors and classified as a potential modifier variant that can increase the risk of disease, depending on other genetic alteration in other genes. The *CFC1* variant, analyzed using prediction softwares to verify the potential effect of a mutation in the final protein, was predicted to have no functional effect in the CFC1 protein. Altogether, our data indicates that abnormal Nodal signaling caused by the presence of gene variants in various components of the pathway may be an important factor influencing or increasing the risk of developing CHDs during development.

Keywords: Nodal Signaling, CHD, laterality defects, allelic variants.

2. Background

Most congenital heart defects arise during embryonic development as consequence of failures in critical developmental processes that lead to the formation of the heart. This results in most cases from the disruption of a certain signaling pathway caused by genomic alterations in a gene or in various genes involved in that specific pathway. In this study, and as reported in the previous chapter, we found a DAND5 alteration that can function as a risk allele for the development of CHDs/ and or Laterality defects. Nevertheless, and since that the genotype-phenotype seems to represent a case of incomplete penetrance, we cannot exclude the hypothesis that the effect of this DAND5 p.R152H variant could be affected/modified by the presence of other alterations in genes involved in the Nodal signaling pathway. Moreover, alterations in these genes might also contribute for abnormal levels of the pathway and thus for the phenotypes of the patients. Therefore, even though our main focus was on the role of DAND5 in the NODAL signaling pathway, we extended the genetic screening to other genes from this cascade in order to understand how genomic alterations in NODAL signaling genes could impact the DAND5 alteration and congenital heart defects. In addition to the two exons of the *DAND5* gene (Chapter 3, Part I), we also screened for alterations all the exons and exon-boundaries of *NODAL*, *CFC1* and *PITX2C* genes in the children cohort described anteriorly. *NODAL* gene, which is located at chromosome 10q22.1, contains three exons and encodes a secreted protein belonging to transforming growth factor (TGF-β) superfamily that is critical for establishing the left-right asymmetry during early development [120]. *CFC1* is a six exons gene at chromosome 2q21.1 that belongs to the epidermal growth factor-CFC1 family. This gene encodes an extracellular protein that functions as a membrane-associated Nodal co-receptor and plays a key role in intercellular left-right signaling pathway [98]. The paired-like homeodomain transcription factor 2 (*PITX2*) gene is located at chromosome 4q25 and its isoform c results from splicing mechanisms that uses an alternative promoter located upstream of exon 4 [121]. *Pitx2c* isoform plays a critical role as a late effector in left–right asymmetry, being the major downstream effector of the Nodal pathway during

heart morphogenesis, presenting differential expression in various cardiac precursor cells and components the heart.

3. Results and discussion

From the sequences analyzed using sanger sequencing of the thirty-eight patients, we identified two non-synonymous sequence alterations and three intronic variants in the genes *NODAL*, *CFC1* and *PITX2C*. The missense alterations were found in the exons of the *NODAL* (H165R) and *CFC1* (A175T) genes whereas the non-coding sequence variants were found in the first intron of both *NODAL* (c.193+12 C>T) and *PITX2C* genes (c.205+88G>C; c.205+71A>G) (Table 3.4).

Table 3.4. Localization and frequencies of variants found.

Gene	Nucleotide/ Protein	Affected
NODAL	494 A>G/H165R	33/38
NODAL	c.193+12 C>T (intron)	20/38
CFC1	523G>A/A175T	1/38
PITX2C	c.205+71A>G (intron)	8/38
PITX2C	c.205+88G>C (intron)	17/38

To understand if these sequence variants were newly identified variants, we searched on NCBI (http://www.ncbi.nlm.nih.gov/), Ensemble (http://www.ensembl.org), ExAC (http://exac.broadinstitute.org) and Human Gene Mutation Database (www.hgmd.cf.ac.uk) databases if those alterations had already been described by others or found in genome sequencing projects. According to our analysis, the *NODAL* and *PITX2C* variants were found in different sequencing projects and annotated in the NCBI, Ensemble and ExAC browser as Single Nucleotide Polymorphisms (SNPs). Regarding the *CFC1* genetic variation, we could not find any data related to that particular alteration, neither in the literature, nor in the available polymorphism/mutations databases,

suggesting that it is likely a new variant that had not been studied so far. The results from this search are detailed below.

3.1. NODAL H165R variant

One of the alterations found in this study was a substitution of the ancestral nucleotide adenine to a guanine (Figure 3.3) in the 494th coding nucleotide (c.494A>G) in the exon 2 of the *NODAL* gene. At protein level, this alteration results in the missense replacement of a histidine for an arginine in the 165th position of the pro-domain of the NODAL protein (Figure 3.3).

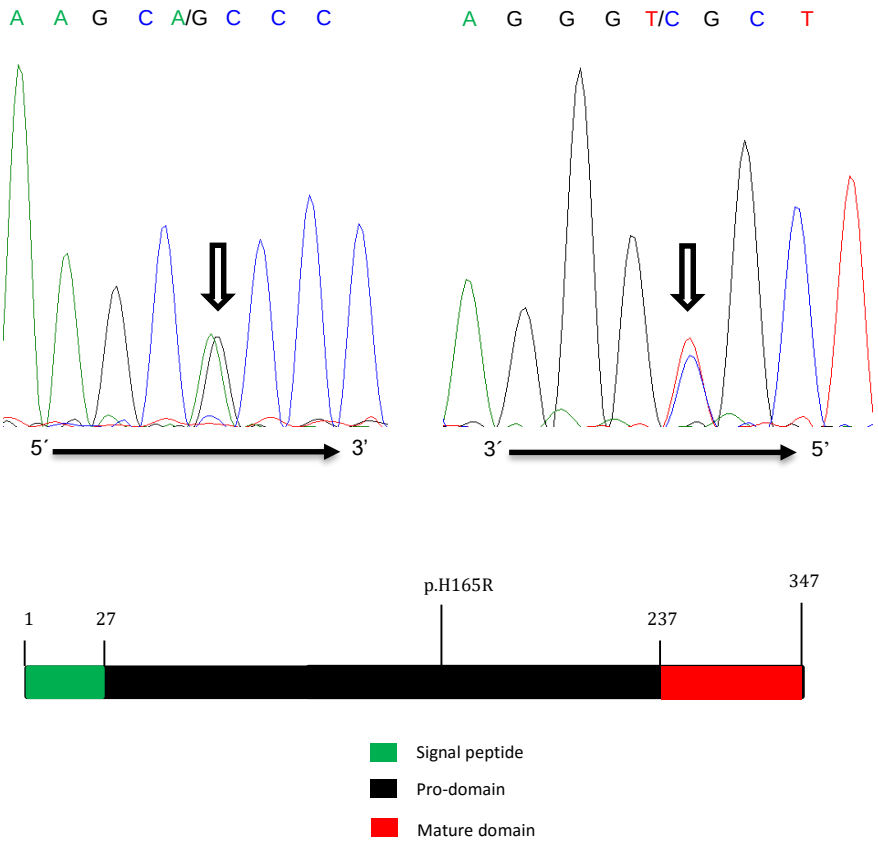


Figure 3.3 – Sequencing and genomic localization of the NODAL 494 A>G allelic variant.

and schematic diagram of NODAL structure with approximated localization of the variant identified in this study and cross-species sequence alignment of DAND5 showing the conservation of amino acid R152 (highlighted by red box).

This genomic alteration was identified in thirty-three of our thirty-eight patients (Table 3.4), being homozygous in seven probands and heterozygous in twenty-six. This alteration had already been described by others and/or in any polymorphism/mutations databases as an extremely common sequence variant, corresponding to the SNP rs1904589, with a heterozygosity of almost 50% and

a minor allele frequency (MAF) (allele A/T) of 46% in the European population of the 1000 genomes project (Figure 3.4). Moreover, it was also identified in two other studies, in 2008 and in 2009, conducted by Maximilian Muenke and co-workers [103, 107].

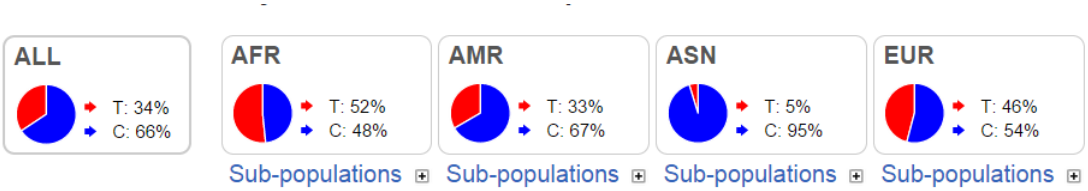


Figure 3.4 – *NODAL* 494 G(C)>A(T) allele frequency according to 1000 genome project.

The NODAL 494 A>G variant present a minor allele frequency (MAF) of allele A = 0.46 in the European Population. AFR - African population; AMR - American population; ASN - Asian population; EUR - European population.

In these studies, Maximilian Muenke and co-workers examined the genomes of individuals with CHD, laterality defects, and Holoprosencephaly (HPE) defects and found that the NODAL p.H165R alteration was present in the CHDs cohort patients [103, 107]. They also functionally characterized this variant using a Nodal-dependent luciferase assay showning that this common variant had less than 50% activity when compared to the wild-type NODAL protein (Figure 3.5) [107].

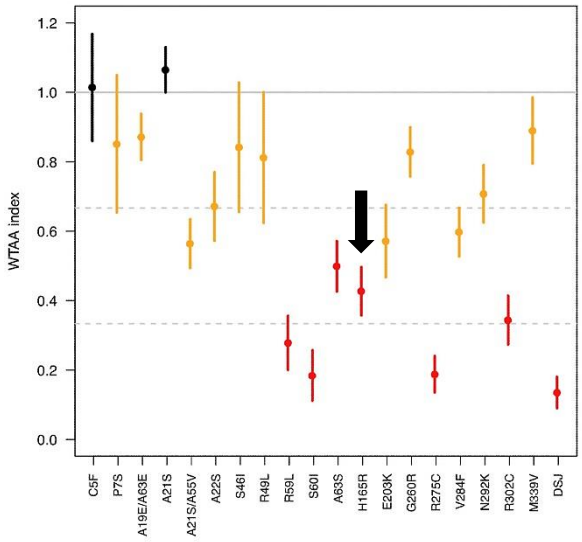


Figure 3.5 – Relative activity of human NODAL p.H165R variant.

The p.H165H variant was evaluated through a luciferase assay as readout of Nodal signaling. The result showed that this NODAL variant leads to a decrease of about 50% of NODAL signaling (black arrowhead) when compared to wild-type NODAL protein (adapted from Roessler et al. 2009).

Furthermore, and since the variant allele (A/T) is the more frequent in the population, as shown in Figure 3.4, the author evaluated if the NODAL p.H165R variant had been observed in other vertebrates. As shown in Figure 3.6, the p.R165 variant only exists in chimpanzees and humans, not in other higher mammals, indicating that the p.H165 is the ancestral version and correctly used as the normal version of the gene/protein [107]. Therefore and due to its functional effect, the common p.R165 variant may serve as a common modifier associated with an increased risk of disease in the CHD patients.

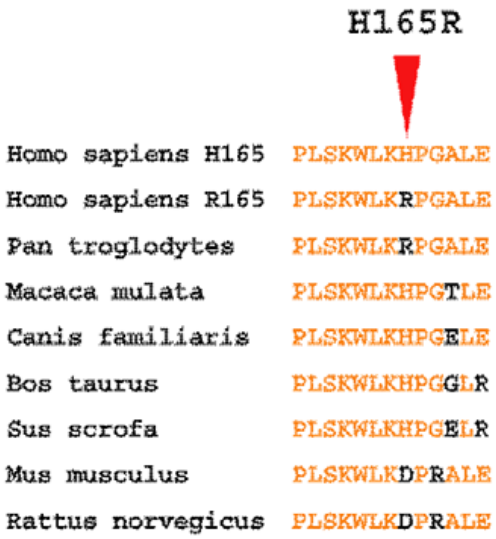


Figure 3.6 – Conservation of the NODAL p.H165R variant between species.

Cross-species sequence alignment of p.H165R NODAL protein variant showing the conservation of amino acid R165 (highlighted by red inverted triangle) only in the homo sapiens and pan troglodyte (adapted from Roessler et al. 2009).

3.2. *NODAL* intronic mutation

The other *NODAL* gene alteration identified in this study is localized in the first intron and consists of a substitution of a cytosine for a thymine (Figure 3.7), twelve nucleotides after the end of the exon 1 (c.193+12 C>T). This variant was observed in twenty patients of our cohort (Table 3.4), being heterozygous in seventeen and homozygous in three. This c.193+12 C>T allelic variant is also annotated in the public databases as a SNP – rs10999338 – (Ensemble, NCBI and ExAC). In fact, this SNP has been described as an extremely common sequence variant with a heterozygosity of 45% and a minor allele frequency (T)

of 44% in the European population sequenced in the 1000 genomes project (Figure 3.7).

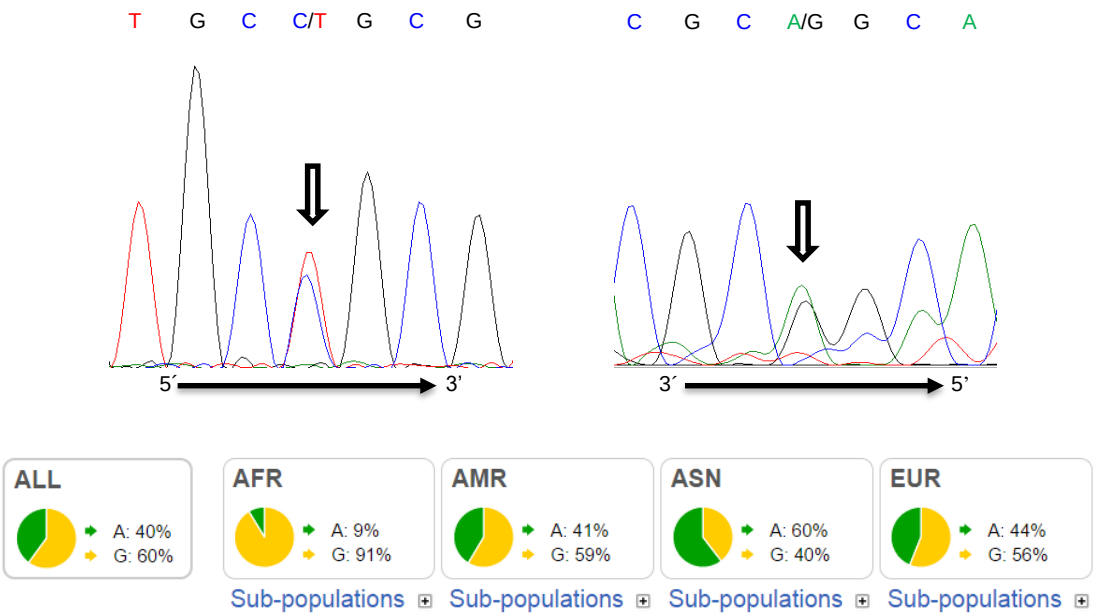


Figure 3.7 – Sequencing and allele frequency of the c.193+12C>T variant.

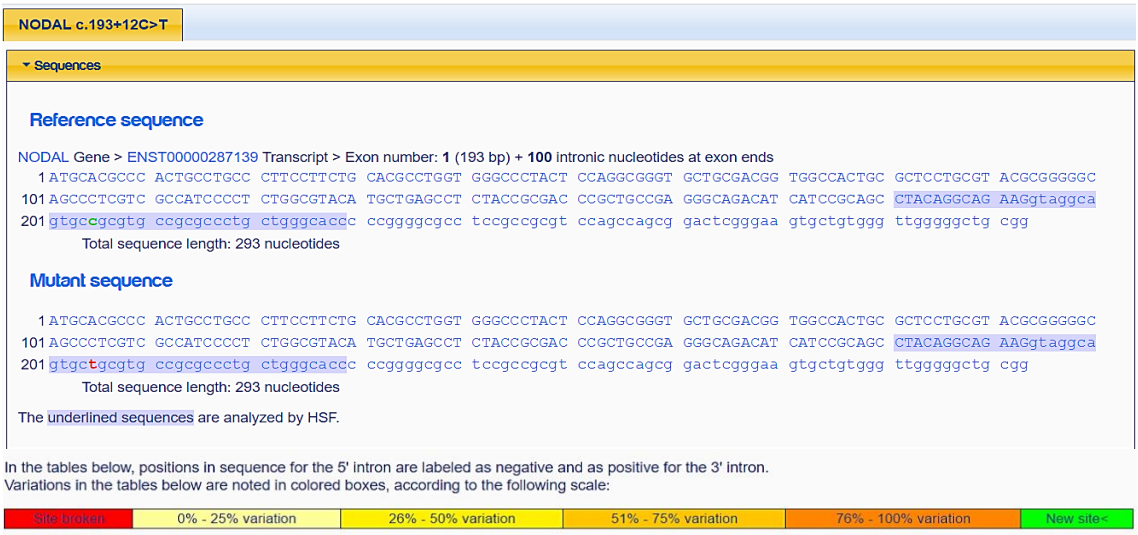
Forward and reverse DNA sequence chromatograms of one patient showing the c.193+12 C>T allelic variant (highlighted by arrows) and representation of allele frequency of C and T alleles in the populations sequenced in the 1000 genomes project. AFR - African population; AMR - American population; ASN - Asian population; EUR - European population

3.2.1. *In silico* analysis

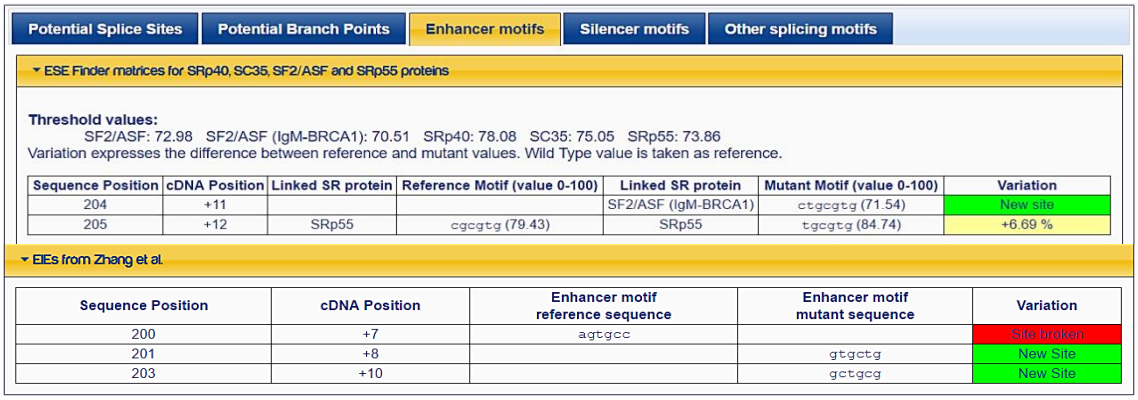
Although alterations in the coding regions of the genome are suggested to be the main and more logical cause of disease in humans, today, it is estimated that ~15% of all point mutations causing human disorders disrupt splice-site consensus sequences, particularly at intronic positions [122]. Therefore, due to the close localization of the c.193+12 C>T allelic variant to the 5' splice site and in order to understand its relevance, we used the online software Human Splicing Finder to predict *in silico* if this intronic variation could be interfering with the normal splicing process. As shown in Figure 3.8, the prediction results indicate that this C>T substitution may lead to the generation of 3 new enhancer locals, including one for the ligation of the SF2/ASF splicing factor and other that allows the biding of the SRp55 protein to its motif with more affinity and may disrupt the "AGTGCC" enhancer motif (Figure 3.8B). The SRp55 and SF2/ASF proteins are involved in pre-mRNA splicing and in the specificity of

alternative splice site selection. Furthermore, the presence of T nucleotide instead of the normal C in the c.193+12 position seems also to create new silencer motifs (Figure 3.8C). Altogether, these *in silico* results indicate that the c.193+12C>T can be creating new or interfering with existing splice sites, interfering with the normal splicing of the *NODAL* gene.

A)



B)



C)

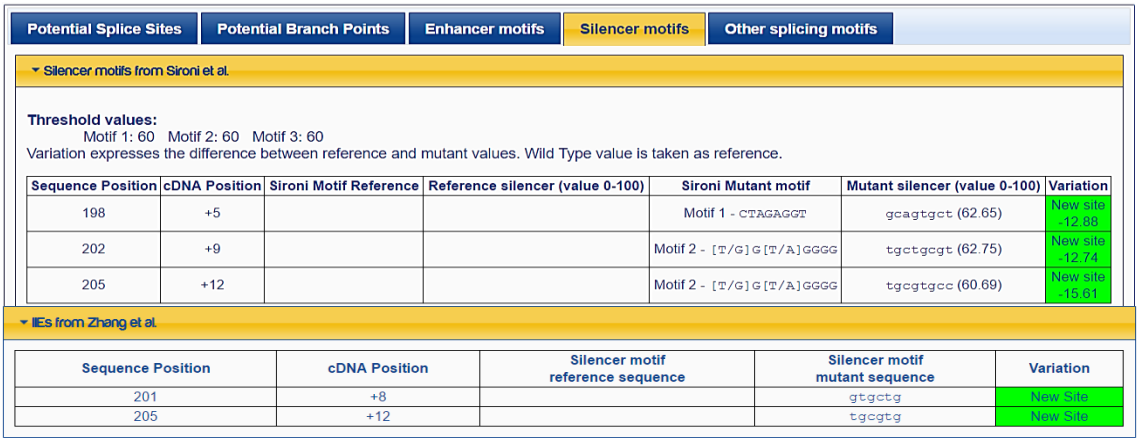


Figure 3.8 – *In silico* analysis of NODAL intronic variant using the Human Splicing Finder online software A) Wild-type and variant genomic sequences analyzed by the HSF software. **B)** Predicted enhancer motifs created or broken due to the c.193+12C>T variant. **C)** Potential new silencer motifs.

3.3. *PITX2C* variants

The genetic screening in our patient cohort also revealed two variants in the intron of the *PITX2* isoform C gene. These alterations localize close to each other in intron 1B, more specifically in the genomic positions c.205+71 and c.205+88, and lead to a substitution of an A>G and a G>C, respectively. The c.205+71A>G allelic alteration was identified in eight of the thirty-eight patients cohort, being heterozygous in six and homozygous in two whereas the c.205+88 G>C was present in seventeen patient, of which four are homozygous and thirteen heterozygous.

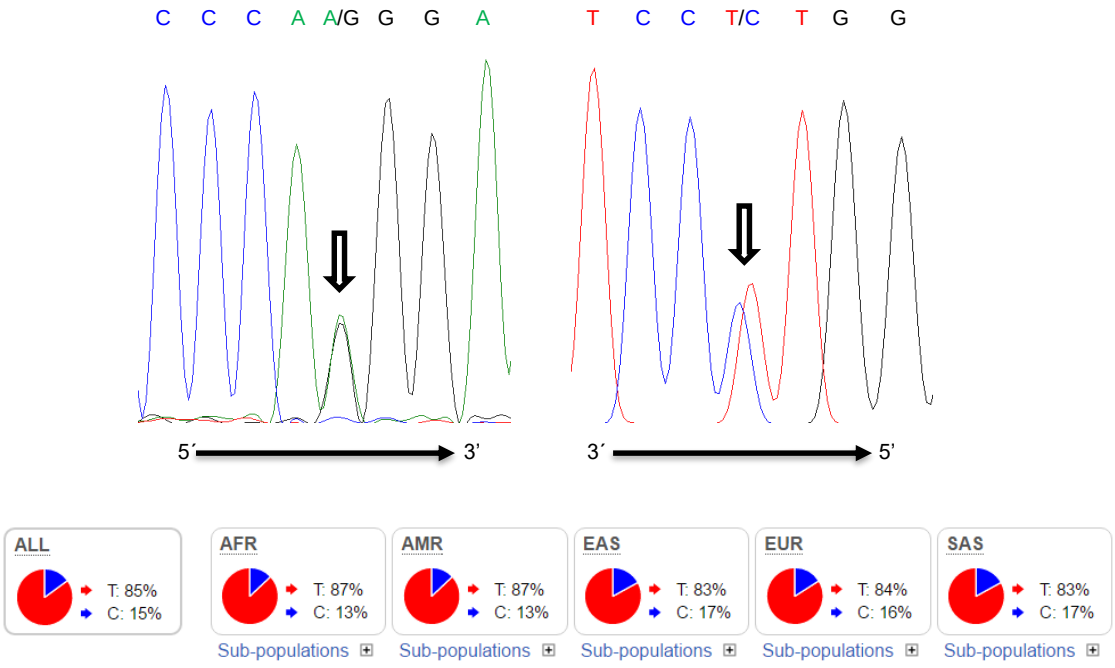


Figure 3.9 – Sequencing and allele frequency of the *PITX2C* c.205+71A>G substitution.

Forward and reverse DNA sequence chromatograms of one patient showing the c.205+71 A>G allelic variant (highlighted by arrows) and representation of the allele frequency of T(A) and G(C) alleles in the populations sequenced in the 1000 genomes project. AFR - African population; AMR - American population; ASN - Asian population; EUR - European population

In the same line of thought as for the alterations described above, we also checked in the public databases, NCBI, Ensemble, ExAC and HGMD for any

reference to these two *PITX2C* intronic variants. We verified that the intronic c.205+71A>G substitution (Figure 3.9) in the *PITX2C* gene was reported in the public databases of NCBI and Ensemble as the rs28510307 SNP, but not in ExAC. Moreover, in these two databases, we also found some information about the frequency of the normal and variant nucleotide. As shown in Figure 3.9, this common variant has a Minor Allele Frequency (allele G/C) of 16% in the European population and was detected as heterozygous in 27% of the same cohort of persons that participated in the 1000 genome project. Regarding the c.205+88 G>C variant (Figure 3.10), which we found to correspond to the SNP rs762401950, even though it was also described in the NCBI and Ensemble databases, it was not detected in the 1000 genome project or other big sequencing projects. Therefore, we could not have any relevant information about its frequency in the population our about the study context in which the variant was detected.

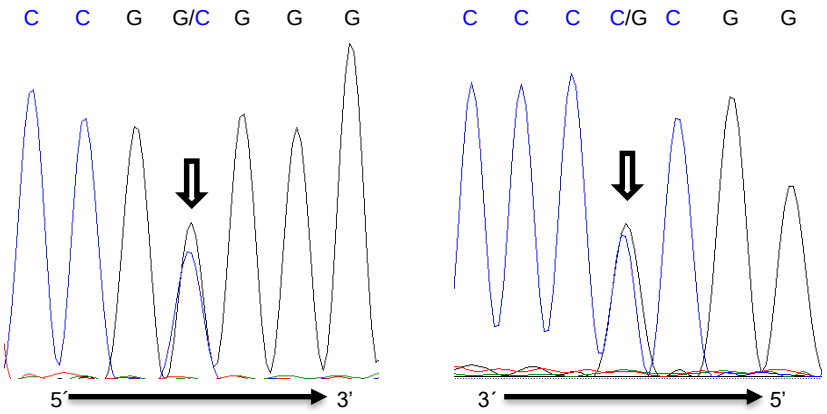


Figure – 3.10 – DNA Sequencing of *PITX2C* gene.

Forward and reverse DNA sequence chromatograms of one patient showing the c.205+88 G>C allelic variant (highlighted by arrows).

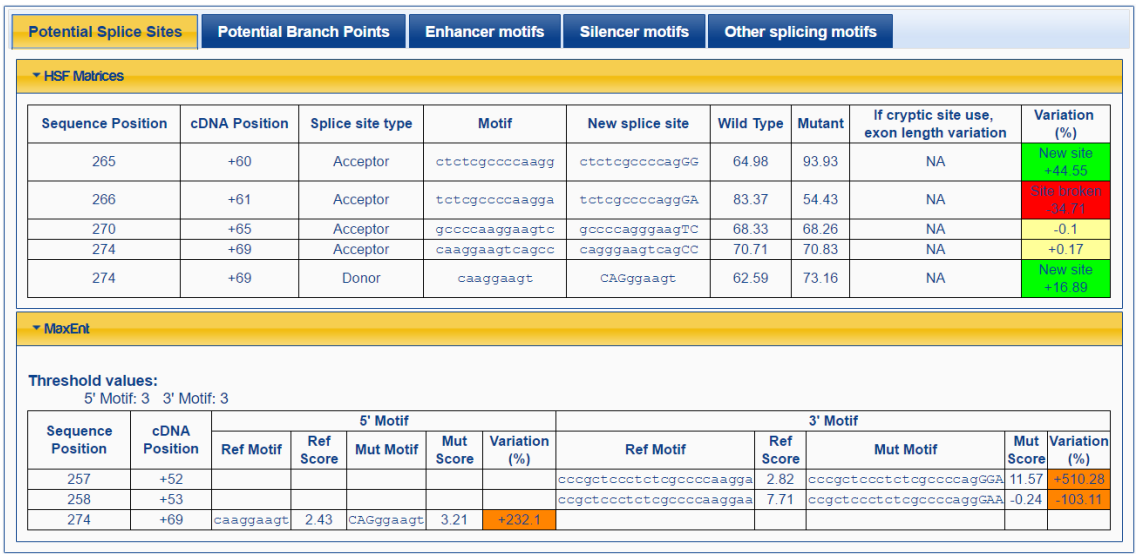
3.3.1. *PITX2C* c.205+71A>G *in silico* analysis

In order to understand the possible impact of the c.205+71 A>G allelic variant in the splicing process of the *PITX2C* gene isoform, we performed an *in silico* analysis of this variant using the previously mentioned Human Splicing Finder (HSF) online software. As presented in the Figure 3.11, the HSF predicted that the A>G nucleotide alteration could lead to the creation or disruption of splice site acceptors and creation of a new splice site donor (Figure 3.11B) in the

genomic sequence of *PITX2C* gene. Moreover, this alteration could also disrupt a branch point site (Figure 3.11C), which could be involved in the selection of the next exon to incorporate into the mature RNA. But the impact of this variant in the splicing event of the *PITX2C* gene could not be just in the main interviewing sequences (donor and acceptor sites and branch point site) responsible for this process. As shown in Figure 3.11D and E, this nucleotide change seems also to alter several possible splicing enhancer and silencers motifs, being the more important ones the disruption of the splicing enhancers motifs responsible for the ligation of the SF2/ASF and 9G8 splicing factor proteins and the disruption of the splicing silencer motif "AAGGAA" responsible for the biding of the hnRNP A1 protein.

A)

B)



Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
265	+60	Acceptor	ctctcgccccaagg	ctctcgccccagGG	64.98	93.93	NA	New site +44.55
266	+61	Acceptor	tctcgccccaagga	tctcgccccaggGA	83.37	54.43	NA	Site broken -34.71
270	+65	Acceptor	gccccaaggaagtc	gccccagggaagTC	68.33	68.26	NA	-0.1
274	+69	Acceptor	caaggaagtcagcc	cagggaagtcagCC	70.71	70.83	NA	+0.17
274	+69	Donor	caaggaagt	CAGggaagt	62.59	73.16	NA	New site +16.89

Sequence Position	cDNA Position	5' Motif					3' Motif				
		Ref Motif	Ref Score	Mut Motif	Mut Score	Variation (%)	Ref Motif	Ref Score	Mut Motif	Mut Score	Variation (%)
257	+52						cccgctccctctcgccccaagga	2.82	cccgctccctctcgccccagGGA	11.57	+510.28
258	+53						ccgctccctctcgccccaaggaa	7.71	ccgctccctctcgccccaggGAA	-0.24	-103.11
274	+69	caaggaagt	2.43	CAGggaagt	3.21	+232.1					

C)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

Branch Points calculation is performed using a new algorithm.

Sequence Position	cDNA Position	Branch Point motif	CV for reference sequence (0-100)	CV for mutant sequence (0-100)	Variation
270	+65	gccccAa	82.72	86.61	
271	+66	ccccaAg	76.42	46.79	Site broken

D)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins

Threshold values:
SF2/ASF: 72.98 SF2/ASF (IgM-BRCA1): 70.51 SRp40: 78.08 SC35: 75.05 SRp55: 73.86
Variation expresses the difference between reference and mutant values. Wild Type value is taken as reference.

Sequence Position	cDNA Position	Linked SR protein	Reference Motif (value 0-100)	Linked SR protein	Mutant Motif (value 0-100)	Variation
269	+64	SF2/ASF	cgccccaa (76.65)	SF2/ASF	cgccccag (76.65)	0 %
271	+66	SRp40	ccccaag (80.06)	SRp40	ccccagg (83.47)	+4.26 %
271	+66	SRp40	ccccaag (80.06)	SF2/ASF (IgM-BRCA1)	ccccagg (71.38)	-10.84 %
272	+67	SF2/ASF (IgM-BRCA1)	cccaagg (74.54)	SF2/ASF (IgM-BRCA1)	cccaggg (87.62)	+17.54 %
272	+67	SF2/ASF (IgM-BRCA1)	cccaagg (74.54)	SF2/ASF	cccaggg (82.18)	+10.25 %
273	+68	SF2/ASF (IgM-BRCA1)	ccaagga (77.38)			Site broken -100

RESCUE ESE hexamers

Sequence Position	cDNA Position	Enhancer motif reference sequence	Enhancer motif mutant sequence	Variation
275	+70	aaggaa		Site broken
276	+71	aggaag		Site broken

Predicted PESE Octamers from Zhang & Chasin

Sequence Position	cDNA Position	Reference motif	Motif value (0-100) reference sequence	Mutant motif	Motif value (0-100) mutant sequence	Variation
270	+65			gccccagg	52.95	New site
275	+70	aaggaagt	38.75			Site broken -100

EIEs from Zhang et al.

Sequence Position	cDNA Position	Enhancer motif reference sequence	Enhancer motif mutant sequence	Variation
271	+66	ccccaa	ccccag	
272	+67	cccaag		Site broken
273	+68	ccaagg		Site broken
274	+69	caagga		Site broken
275	+70	aaggaa		Site broken
276	+71	aggaag	gggaag	

ESE motifs from HSF - *Experimental*

Threshold values:
Tra2: 75.964 9G8: 59.245
Variation expresses the difference between reference and mutant values. Wild Type value is taken as reference.

Sequence Position	cDNA Position	Linked ESE protein	Reference Motif (value 0-100)	Linked ESE protein	Mutant Motif (value 0-100)	Variation
275	+70	9G8	aaggaa (61.34)			Site broken -100

E)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

Silencer motifs from Sironi et al.

Fas-ESS hexamers

Sequence Position	cDNA Position	Reference sequence	Set	Mutant sequence	Set	Variation
+68	+68	ccaagg	2			Site broken

hnRNP motifs - *Experimental*

Threshold values:
hnRNP A1: 65.476
Variation expresses the difference between reference and mutant values. Wild Type value is taken as reference.

Sequence Position	cDNA Position	Linked hnRNP protein	Reference Motif (value 0-100)	Linked hnRNP protein	Mutant Motif (value 0-100)	Variation
274	+69	hnRNP A1	caagga (67.62)	hnRNP A1	caggga (84.76)	+25.35 %
275	+70	hnRNP A1	aaggaa (76.43)			Site broken -100

Figure 3.11 – *In silico* analysis of the *PITX2C* c.205+71 A>*G* variant.

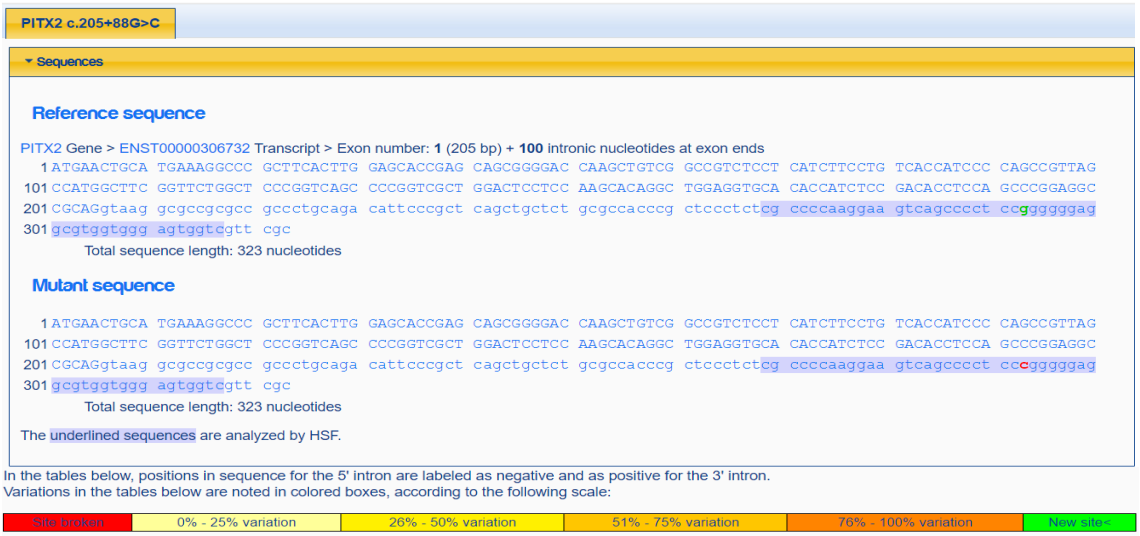
A) Wild-type and variant genomic sequences analyzed by the HSF software. **B)** Potential donor or acceptor splice sites created or broken by the c.205+71 A>G alteration. **C)** Potential branch point disruption caused by the allelic variant. **D-E)** Predicted created or broken splicing enhancer and silencers motifs.

These results, from an *in silico* analysis performed using the online software HSF, indicate that the c.205+71A>G intronic variant could result in abnormal splicing of the isoform C of the *PITX2* gene, mainly due to the disruption/creation of essential donor, acceptor or branch point sites and due to the broken of splicing enhancer/inhibitor motifs responsible for the biding of the spliceosome proteins.

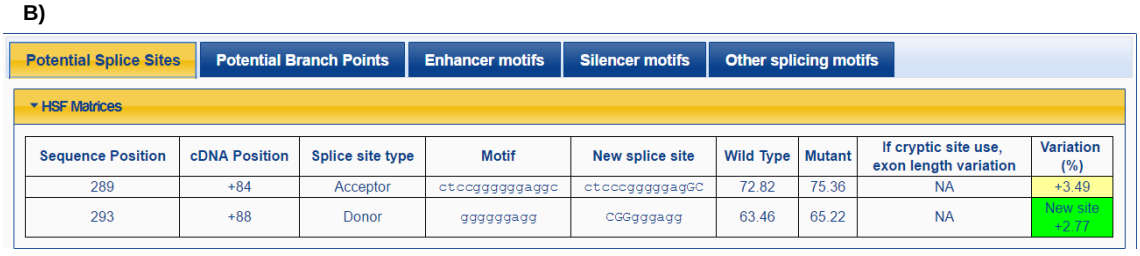
3.3.2. *PITX2C* c.205+88 G>C *in silico* analysis

As for the *NODAL* and the other *PITX2C* intronic variants described above, we also performed the *in silico* analysis of the *PITX2C* c.205+88 G>C nucleotide change using the HSF prediction software. This variant was predicted to create a new splice donor site, new enhancer motifs for the biding of the splicing factor SF2/ASF and disrupt some splice inhibitor motifs (Figure 3.12).

A)



B)



Sequence Position	cDNA Position	Splice site type	Motif	New splice site	Wild Type	Mutant	If cryptic site use, exon length variation	Variation (%)
289	+84	Acceptor	ctccgggggaggc	ctcccgggggagGC	72.82	75.36	NA	+3.49
293	+88	Donor	gggggagg	CGGgggagg	63.46	65.22	NA	New site +2.77

C)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

Branch Points calculation is performed using a new algorithm.

Sequence Position	cDNA Position	Branch Point motif	CV for reference sequence (0-100)	CV for mutant sequence (0-100)	Variation
287	+82	ccctcCg	67.65	70.37	

D)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

ESE Finder matrices for SRp40, SC35, SF2/ASF and SRp55 proteins

Threshold values:
SF2/ASF: 72.98 SF2/ASF (IgM-BRCA1): 70.51 SRp40: 78.08 SC35: 75.05 SRp55: 73.86
Variation expresses the difference between reference and mutant values. Wild Type value is taken as reference.

Sequence Position	cDNA Position	Linked SR protein	Reference Motif (value 0-100)	Linked SR protein	Mutant Motif (value 0-100)	Variation
288	+83	SRp40	cctccgg (81.14)	SRp40	cctcccg (79.28)	-2.29 %
289	+84	SF2/ASF (IgM-BRCA1)	ctccggg (86.38)	SF2/ASF (IgM-BRCA1)	ctcccgg (88.00)	+1.87 %
289	+84	SF2/ASF (IgM-BRCA1)	ctccggg (86.38)	SF2/ASF	ctcccgg (83.46)	-3.39 %
289	+84	SF2/ASF	ctccggg (79.91)	SF2/ASF (IgM-BRCA1)	ctcccgg (88.00)	+10.13 %
289	+84	SF2/ASF	ctccggg (79.91)	SF2/ASF	ctcccgg (83.46)	+4.45 %
291	+86	SF2/ASF (IgM-BRCA1)	ccggggg (71.46)	SF2/ASF (IgM-BRCA1)	cccgggg (74.38)	+4.09 %
292	+87	SF2/ASF (IgM-BRCA1)	cgggggg (78.92)	SF2/ASF (IgM-BRCA1)	ccggggg (71.46)	-9.45 %
293	+88			SF2/ASF (IgM-BRCA1)	cgggggg (82.23)	New site
293	+88			SF2/ASF	cgggggg (75.48)	New site

RESCUE ESE hexamers

Predicted PESE Octamers from Zhang & Chasin

Sequence Position	cDNA Position	Reference motif	Motif value (0-100) reference sequence	Mutant motif	Motif value (0-100) mutant sequence	Variation
286	+81	cccctccg	3.30	cccctccc	8.16	
288	+83			cctcccgg	43.34	New site
291	+86			cccggggg	9.69	New site

E)

Potential Splice Sites | Potential Branch Points | Enhancer motifs | Silencer motifs | Other splicing motifs

Silencer motifs from Sironi et al.

Threshold values:
Motif 1: 60 Motif 2: 60 Motif 3: 60
Variation expresses the difference between reference and mutant values. Wild Type value is taken as reference.

Sequence Position	cDNA Position	Sironi Motif Reference	Reference silencer (value 0-100)	Sironi Mutant motif	Mutant silencer (value 0-100)	Variation
288	+83	Motif 3 - TCTCCCAA	cctccggg (64.93)	Motif 3 - TCTCCCAA	cctcccgg (82.12)	+26.47 %
290	+85	Motif 2 - [T/G]G[T/A]GGGG	tccggggg (70.71)			Site broken -16.13
291	+86	Motif 2 - [T/G]G[T/A]GGGG	ccgggggg (60.84)	Motif 2 - [T/G]G[T/A]GGGG	cccgggggg (60.84)	0 %
292	+87	Motif 2 - [T/G]G[T/A]GGGG	cgggggga (77.74)	Motif 2 - [T/G]G[T/A]GGGG	ccgggggga (60.84)	-21.74 %
293	+88	Motif 2 - [T/G]G[T/A]GGGG	gggggggag (73.71)	Motif 2 - [T/G]G[T/A]GGGG	cgggggag (66.91)	-9.22 %

ESS decamers from Wang et al.

Fas-ESS hexamers

Sequence Position	cDNA Position	Reference sequence	Set	Mutant sequence	Set	Variation
+87	+87	cggggg	2			Site broken
+88	+88	gggggg	2, 3	cggggg	2	

PESS Octamers from Zhang & Chasin

Sequence Position	cDNA Position	Reference motif	Motif value (0-100) reference sequence	Mutant motif	Motif value (0-100) mutant sequence	Variation
291	+86	ccgggggg	2.24			Site broken -633.12

IIEs from Zhang et al.

Sequence Position	cDNA Position	Silencer motif reference sequence	Silencer motif mutant sequence	Variation
292	+87	cggggg		Site broken
293	+88	gggggg	cggggg	

Figure 3.12 – *In silico* analysis of the *PITX2C* c.205+81 A>*G* variant.

A) Wild-type and variant genomic sequences analyzed by the HSF software. **B)** Potential donor or acceptor splice sites created or broken by the c.205+81 A>G alteration. **C)** Potential branch point disruption caused by the allelic variant. **D-E)** Predicted created or broken splicing enhancer and silencers motifs.

Therefore, the predicted results illustrated in the Figure 3.12 indicated that this variant seems to have a milder effect on splicing compared with the other *PITX2C* variant, mostly because of the lack of effect in the motifs responsible for the biding of the 9G8, Tra2 and hnRNP A1 splicing factors.

3.4. CFC1 missense alteration

The sequencing analyses of our patient cohort revealed one proband displaying one alteration in the *CFC1* gene. Clinically, the patient presented ventricular septal defect, coarctation of the aorta, aortic and mitral valve stenosis. The variant was identified as heterozygous non-synonymous and results in a substitution of guanine for an adenine in the 523th nucleotide of the coding sequence of the *CFC1* gene. At protein level, this alteration results in the missense replacement of an alanine, a non-polar side chain amino acid, for a threonine, a polar side chain amino acid, in the 175th position of the hydrophobic C terminus - membrane-anchoring signal of the CFC1 protein (Figure 3.13).

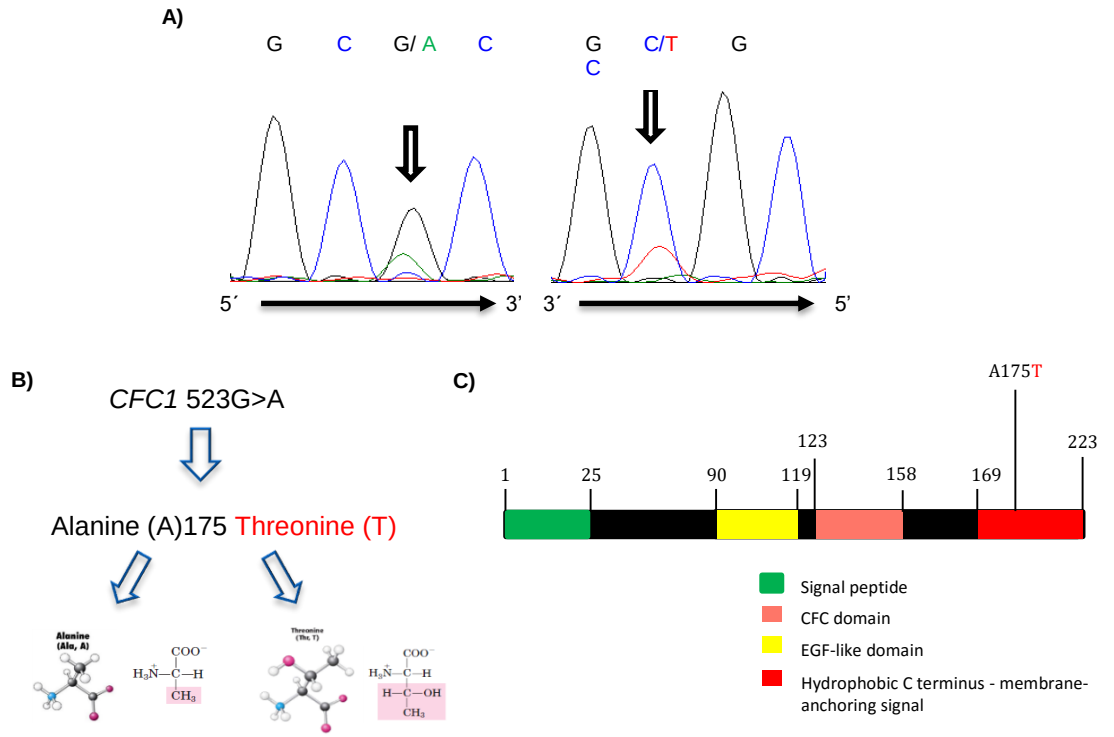


Figure 3.13– *CFC1* allelic variant.

A) Forward and reverse DNA sequence chromatograms showing the c.523 G>A variant (highlighted by arrows). **B)** The alteration result in a substitution of an a non-polar by a polar side chain amino acid **C)** Schematic diagram of *CFC1* structure with approximated localization of the p.A175T variant identified.

After searching in the public databases if this alteration was already described, we could not find any match related to that particular alteration, neither in the literature. This suggested that this alteration is likely a new variant. In order to understand the potential functional effect of this alteration in the structure and function of the CFC1 protein, we used the Polyphen-2, Mutation Taster, PROVEAN and I-Mutant 3.0 prediction softwares. As shown in table 3.5, all softwares predicted that the p.A175T amino acid alteration had no functional effect in the CFC1 protein.

Table 3.5. *In silico* protein prediction effect of CFC1 p.A175T alteration

	Score	**Prediction effect**
Polyphen-2	0,010	Benign
Mutation Taster	Polymorphism	No protein features affected
PROVEAN	-0,23 (cutoff= -2,5)	Neutral
I-Mutant 3.0	-0,36 Kcal/mol	Decreased stability of the protein

3.5. Conclusions

Altogether, the results of the sequencing of the genes *NODAL*, *CFC1* and *PITX2C* involved in the Nodal signaling pathway, which is essential for the correct establishment of the left-right axis and posterior heart morphogenesis, allowed us to identify 4 reported SNPs and a likely new *CFC1* variant in our thirty-eight patient cohort. These variants have impact at the level of the functional protein, as in the case of the NODAL p.H165R variant, or in the splicing process of the RNA transcript as showed by the *in silico* analysis performed with the HSF software. Even though all the SNPs that we identified are relatively frequent in the apparently normal population, we cannot exclude the possibility that they can confer risk for disease especially if more than one SNPs are found in the same individual. Therefore, functional studies are needed in order to understand the real impact of these variants in Nodal signaling, ideally the combined impact of the different variants. Nonetheless, our data indicates that abnormal Nodal signaling caused by the presence of gene

variants in various components of the pathway may be an important factor influencing or increasing the risk of developing CHDs during development.

Part III

Disease Modelling using iPS derived cells

Fernando Cristo[1-4], José M. Inácio[1] and José A. Belo[1]

[1]Stem Cells and Development Laboratory, CEDOC, NOVA Medical School / Faculdade de Ciências Médicas, Universidade Nova de Lisboa, Lisboa, Portugal.

[2]Center for Biomedical Research (CBMR), Universidade do Algarve, Portugal.

[3]PhD Program in Biomedical Sciences, Universidade do Algarve, Portugal.

[4]Regenerative Medicine Program, Biomedical and Medicine Sciences Department, Universidade do Algarve, Portugal.

Authors' contribution: Conceived and designed the experiments: FC, JMI, JB; Sample collection: FC, JMI; Analyzed the data: FC, JMI, SA, JB; Performed the experiments: FC, JMI; Contributed to writing the manuscript: FC, JMI and JB. All authors read and approved the final manuscript

1. Abstract

A 15-year old female child was clinically diagnosed with left isomerism, ventricular septal defect with overriding aorta and pulmonary atresia, defects that can be associated with early left-right establishment impairment. Genetical analysis indicate that this patient carries a heterozygous non-synonymous variant in exon 2 of *DAND5* gene (c.455G>A), causing an amino acid change of p.R152H in the functional domain of the DAND5 protein, a master regular of Nodal signalling. In the work presented here, we generated a *DAND5* 455G>A, and a control iPSC lines using the CytoTune®-iPS 2.0 Reprogramming kit (Life Technologies, Invitrogen), which includes the reprogramming factors SOX2, OCT3/4, c-MYC and KLF4. For this purpose, exfoliated renal epithelial cells isolated from a urine sample were obtained from the patient and healthy control to serve as template cell for reprogramming.

After reprogramming, iPSC colonies from each line were picked, expanded and ultimately characterized. Both cell lines displayed a typical small, round shape, and tightly packed ESC-like morphology with a high nucleus/cytoplasm ratio with prominent nucleoli. DNA Sanger sequencing confirmed the presence of a c.455G>A substitution in one of the alleles of exon 2 in the patient *DAND5* gene corresponding to a R152H protein alteration and no such alteration in the same genomic region of the control cell line. Gene expression and immunofluorescence analysis confirmed the expression of pluripotency markers at mRNA (*OCT3/4*, *NANOG*, *SOX2*, *KLF4* and *NODAL*) and protein (NANOG, OCT4 and SSEA4) levels, characteristic of pluripotent ES cells and illustrating the purity of the both iPSC lines. Karyotyping shows a normal 46, xx alignment of the chromosomes in the patient-derived iPSC line.

Keywords: DAND5 allelic variant, Congenital Heart Diseases, laterality defects, iPSCs, Disease modelling.

2. Background

Numerous animal models carrying defined mutations in their genome have been generated and used for the analysis of gene function and also to mimic human variants *in vivo*, either by gene *knockin* approaches or study of *null* mutations [122]. This approach give rise to a tremendous amount of data helping to increase our knowledge about the molecular mechanisms that lead to human congenital heart diseases. However, in most of the cases it fails in the effective recapitulation of the human phenotype mostly due to the considerable differences that exist between the human and mouse genomes (not all gene sequences or variants are known or sufficiently conserved between species), with many genetic modulators being human specific [122, 123]. In addition, wherein heterozygous alteration of human disease genes is often well tolerated in mice, homozygous alteration could cause in mice severe and even lethal phenotypes not observed in humans [122]. Also, interspecies differences in cell and organ physiology or cellular and molecular composition may preclude the phenotype from being revealed [123]. Therefore, the creation of *in vitro* models for human congenital heart abnormalities that accurately reflect the patients disease phenotypes are important not only for understanding the genetic mechanisms leading to the disease itself, but also for drug screening efforts aimed to the identification of potential treatment therapies.

In 2006, the remarkable report of Yamanaka and colleagues showed that human adult somatic cells could be reprogrammed into a pluripotency state using a combination of four-transcription factors (Oct3/Oct4, Sox2, Klf4 and c-Myc) (Takahashi & Yamanaka, 2006). This groundbreaking observation has revolutionized the stem cell and disease modelling fields, and, in our days, the generation of patient-specific induced pluripotent stem cells (iPSCs) is a well-established technique and the gold standard for cellular models of diseases without the ethical and political concerns associated with the use of human embryos for scientific purposes (C. Kim, 2014).

iPSCs can be generated from different sources of tissues, being the skin fibroblast or blood cells the more commonly used. Nevertheless, skin biopsy has many drawbacks since it is an invasive procedure, accompanied by

discomfort and risk of bleeding, infection, permanent scars on donors, the resulting cells expand slowly in culture, and the reprogramming efficacy is usually low (Jouni et al., 2015). Recently, it was demonstrated that the cells present in a simple urine sample could be an easy and non-invasive source for generating human iPSCs, overcoming the issues from the skin biopsies samples (Zhou et al., 2012).

Here, to have a clear and robust insight into the mechanism of disease displayed by patients carrying the *DAND5* 455G>A allelic variant (chapter 3, part 1), we took advantage of the iPSCs technology to generate a patient-specific cell line of pluripotent stem cells (Figure 3.14). In addition, we generated a iPSC line derived from a healthy adult donor without any alteration in the *DAND5* gene to serve as control cell line (Figure 3.14).

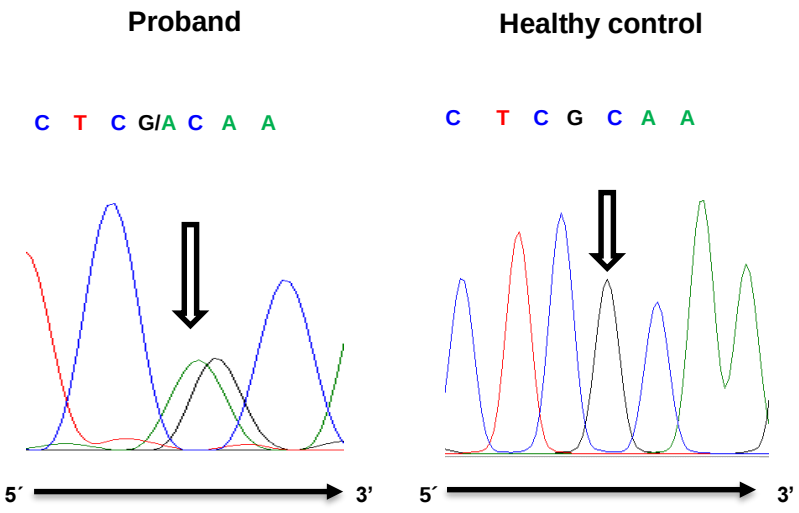


Figure 3.14 – Genotype of patient and control samples used to generated iPSC lines.

DNA sequence chromatograms of a specific region of *DAND5* exon 2 showing the c.455G>A allelic variant (highlighted by arrows) in the proband with Congenital heart defects associated with laterality defects) and a normal nucleotide sequence in the healthy control.

3. Results and discussion

3.1. Isolation and generation of Patient-Specific human iPSC lines from a Urine Sample.

To evaluate the potential effect of the identified *DAND5* 455G>A allelic variant and to model the disease presented by the patient with this alteration, we isolated exfoliated renal epithelial cells isolated from a 15-year old female child patient and from an healthy control urine samples, as described in material and methods, section 8.2. The patient was clinically diagnosed with left isomerism, ventricular septal defect with overriding aorta and pulmonary atresia, defects that can be associated with early impairment of left-right axis establishment. Genetically, the patient carries a heterozygous non-synonymous variant in exon 2 of *DAND5* gene (c.455G>A), causing an amino acid change of p.R152H in the functional cysteine-rich DAN domain of the DAND5 protein (Table3.6).

Table 3.6. Clinical and molecular findings of patient and healthy control.

	Genotype	**Protein alteration**	**Phenotype**
Proband	*DAND5* c.455GA	DAND5 p.R152H	Left isomerism; VSD with overriding aorta; Pulmonary atresia
Healthy control	*DAND5* c.455GG	DAND5 p.R152R	Without CHDs

After isolation, the cells in culture were mostly squamous cells, some blood cells and barely detectable exfoliated renal epithelial cells (Figure 3.15A). With the daily replacement of the medium, the squamous cells completely disappear around day 4 and the first colonies of exfoliated renal epithelial cells become more evident. At this time, the cell culture is composed of two types of exfoliated renal epithelial cell colonies with distinct morphology. As shown in Figure 3.15B-C, type I colonies have a more regular appearance with smooth-edged contours and cobblestone-like cell morphologies, whereas type II colonies are more elongated and randomly arranged, grow more quickly and are more abundant than type I cells. After 7 days in culture, the exfoliated renal epithelial cells reached 70-80% confluency (Figure 3.15D) and were passaged to a six-well plate well for further expansion and reprogramming.

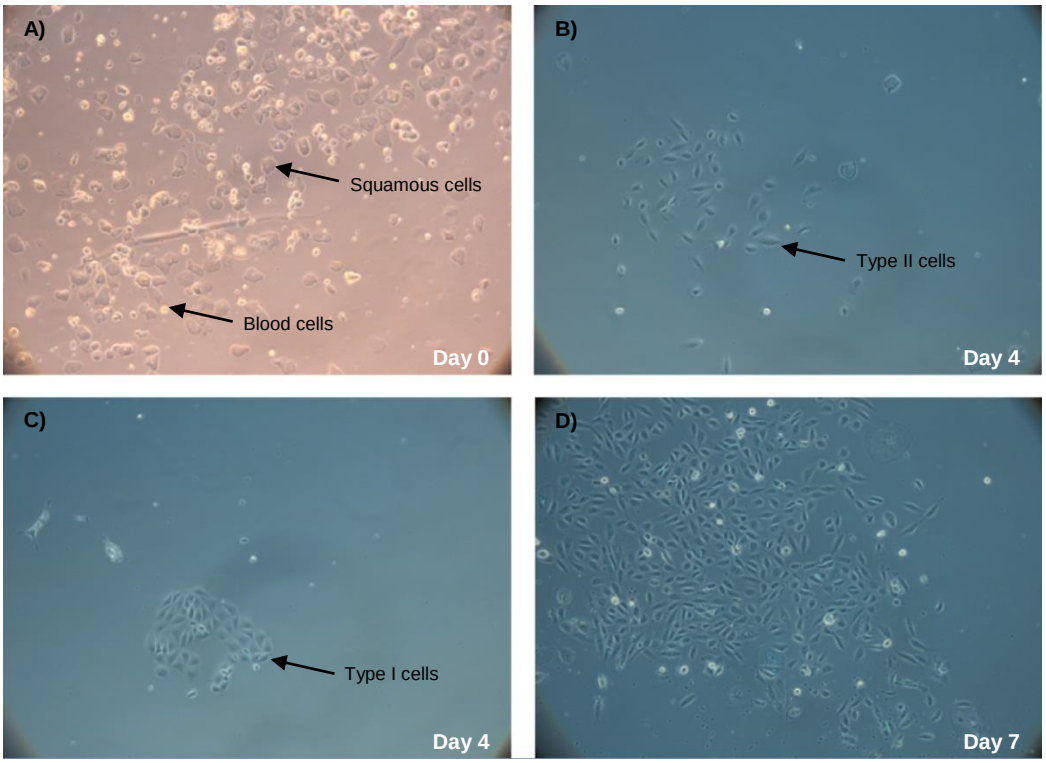


Figure 3.15 – Morphology of the proband urinary cells at different time points after collection.

A) Fresh urine samples mainly consist of squamous cells and a few blood cells (arrowheads). **B,C)** Type II and type I urinary cell colonies (arrowheads) at days 4. **D)** Urinary cell colonies at days 7.

The reprogramming of the cells was performed by using the CytoTune®-iPS 2.0 Reprogramming kit (Life Technologies, Invitrogen), which includes the reprogramming factors SOX2, OCT3/4, c-MYC and KLF4 (Takahashi & Yamanaka, 2006) and is based on a modified non-integrative, and non-transmissible form of Sendai virus (SeV). After transduction with the viruses, the cells passed through a period of morphological changes, going from elongated type II urine cells at the time of infection into a cobblestone-like pattern of cells 120 hours after infection, having higher nucleus to cytoplasmic ratio (Figure 3.16). At 7 days post-transduction, clusters of reprogrammed urinary cells with stem cell-like morphology intercalated with elongated cells that were not reprogrammed could be observed (Figure 3.16G).

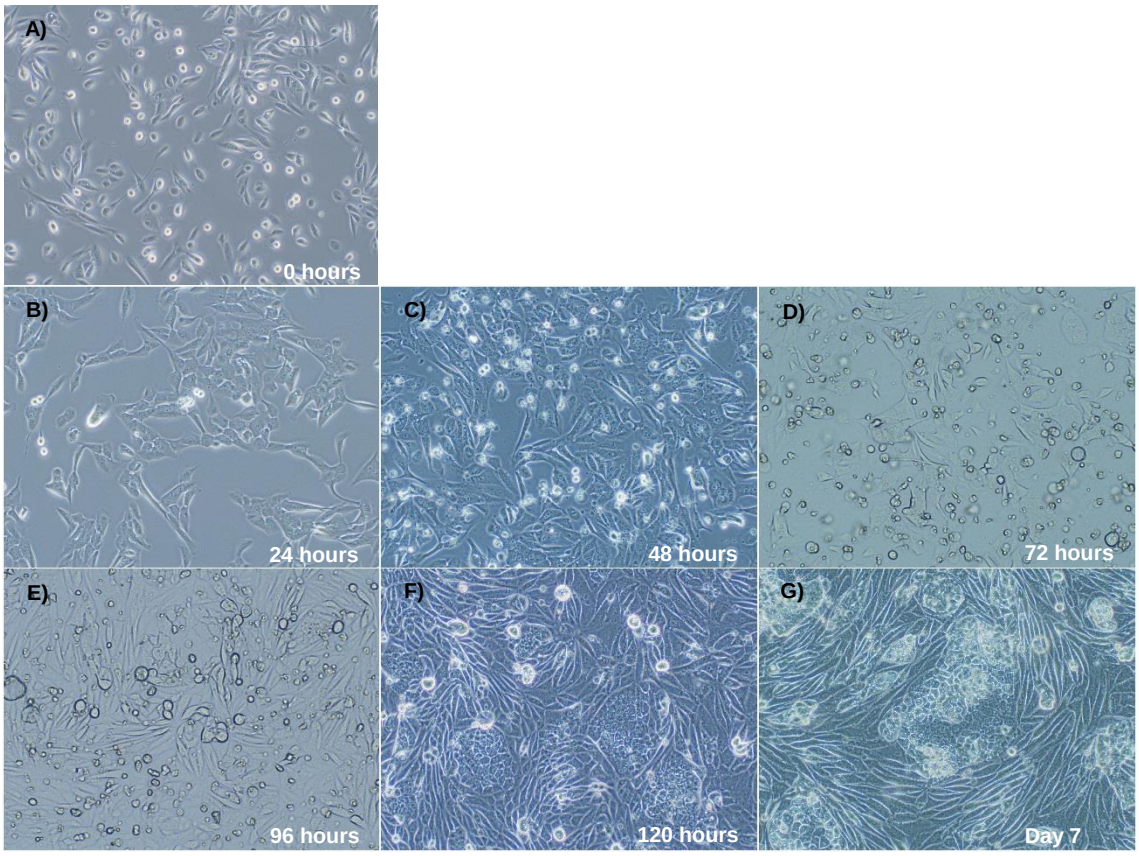


Figure 3.16 – Morphology of the proband-derived cells at different time points during transduction with Sendai virus carrying the Yamanaka reprogramming factors. A) Elongated type II urinary cells at the day of infection. **B – G)** Morphological transition of urinary cells after transduction, resulting in a cobblestone-like pattern of cells at 120 hours post-infection.

In order to isolate the iPSCs and to allow the complete reprogramming of the stem cell-like colonies without the influence of the non-reprogrammed elongated cells, the cultured cells were expanded to a larger 100mm culture dish and cultured in the presence of stem cell E8 medium. As shown in Figure 3.17A, one day after passaging to a 100mm dish, the large clusters of iPSC-like cells were split and the culture composed mainly by individual cells. At this time there was also some cell death of the cells that did not attach to the plate, which were likely mostly the non-reprogrammed cells. With the replacement of the stem cell medium every day, the dead cells in suspension were removed from the culture and the adherent cells started to proliferate and form colonies (Figure 3.17B-D). At day 16, the reprogrammed cells assumed a typical stem cell colony morphology, with defined borders, containing dozens of small round cells each colony, with spaces between them, and large nuclei with notable nucleoli (Figure 3.17E).

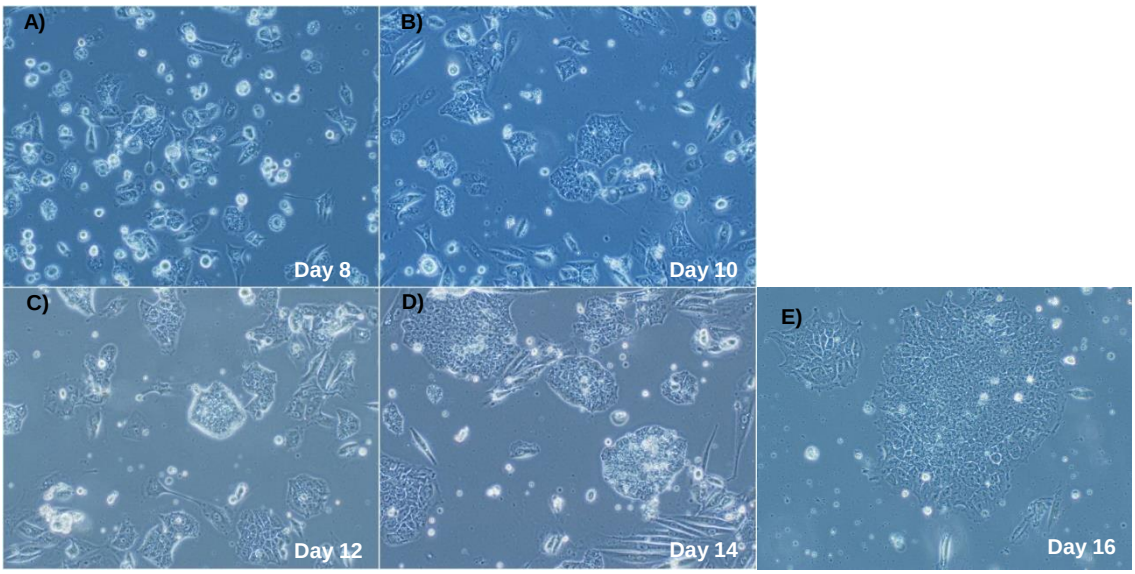


Figure 3.17 – Morphology of proband iPSC-like cells at different time points after expansion.

A) One day after passaging (8 days post-infection) most of the cells attach to the plate whereas some die. **B – E)** Recovery of the iPSC morphology, proliferation and clumping of the cells that survive and attach to the plate.

To obtain homogenous and clonal iPSC lines, several single cell-derived iPSC colonies (Figure 3.17E) were manually picked and transferred to new plates at day 17 post-transduction. This clonal isolation and expansion allows the generation of iPSCs with a higher stemness and enriches the purity of the iPSC line.

Although the age and gender of the donors were different, the same procedure was performed to generate control iPSC lines and equivalent results and timings were obtained for the isolation and generation of these control cells.

3.2. Characterization of *DAND5* allelic variant and healthy control human IPSCs lines.

3.2.1. Morphology

After clonal expansion (~15 clones of each cell line), both the healthy control and *DAND5* allelic variant iPSC lines, continued to display a typical small, round shape, and tightly packed ESC-like morphology with a high nucleus/cytoplasm ratio with prominent nucleoli as presented in Figure 3.18. Suggesting that the generated iPSC lines were successfully reprogrammed.

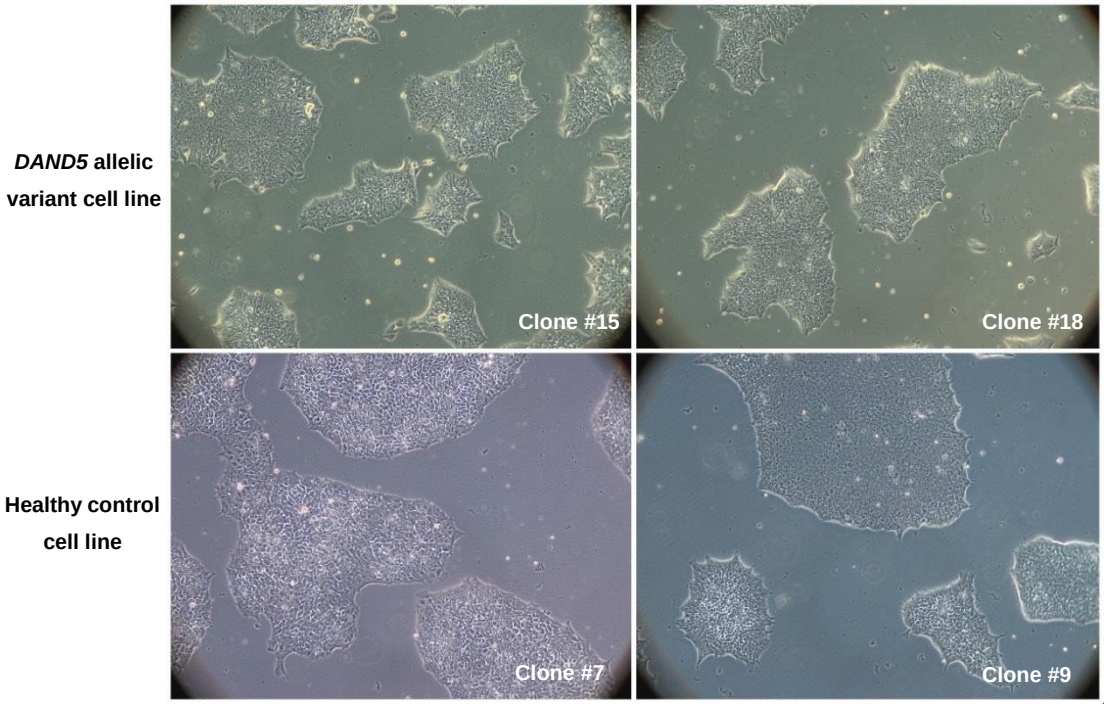


Figure 3.18 – Morphology of two picked clones of each iPS derived cell line.

Both cells lines displayed a typical small, round shape, and tightly packed Embryonic Stem Cell-like morphology with a high nucleus/cytoplasm ratio.

3.2.2. Genotyping

DNA Sanger sequencing confirmed the monoallelic presence of the c.455G>A substitution in the exon 2 of *DAND5* gene, corresponding to a R152H amino acid protein alteration, in one clone of the iPSC line derived from the disease proband (Figure 3.19). As expected, this alteration was absent in the healthy control at the same nucleotide genomic localization (Figure 3.19).

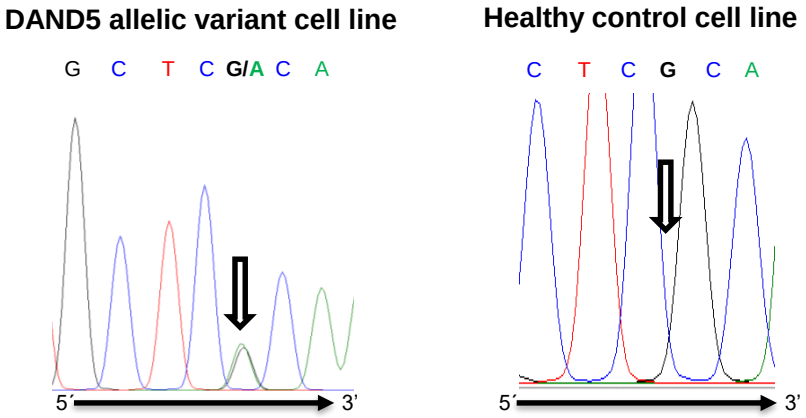


Figure 3.19 – Confirmation of *DAND5* allelic variant and healthy control cell line genotype.

DNA sequence chromatograms of *DAND5* exon 2 showing the c.455G>A allelic variant (highlighted by arrows) in the DAND5 allelic variant derived iPS cell line and a normal nucleotide sequence in the iPS cells from the healthy control.

3.2.3. Pluripotency analysis

To confirm the expression of pluripotency markers at mRNA level, gene expression analysis through quantitative RT-PCR was performed in two clones of each control and patient derived iPSC lines (we just show the values for one clone of each cell line because they are equivalent). According to our analysis, the expression of the endogenous pluripotency genes *OCT3/4*, *NANOG*, *SOX2*, *KLF4* and *NODAL* was present at comparable or higher level in both proband-derived and control iPSC lines (Figure 3.20) when compared to the expression of those genes in a common human embryonic stem cells line (H9). Furthermore, these pluripotency genes were absent in exfoliated renal epithelial cells (Figure 3.20).

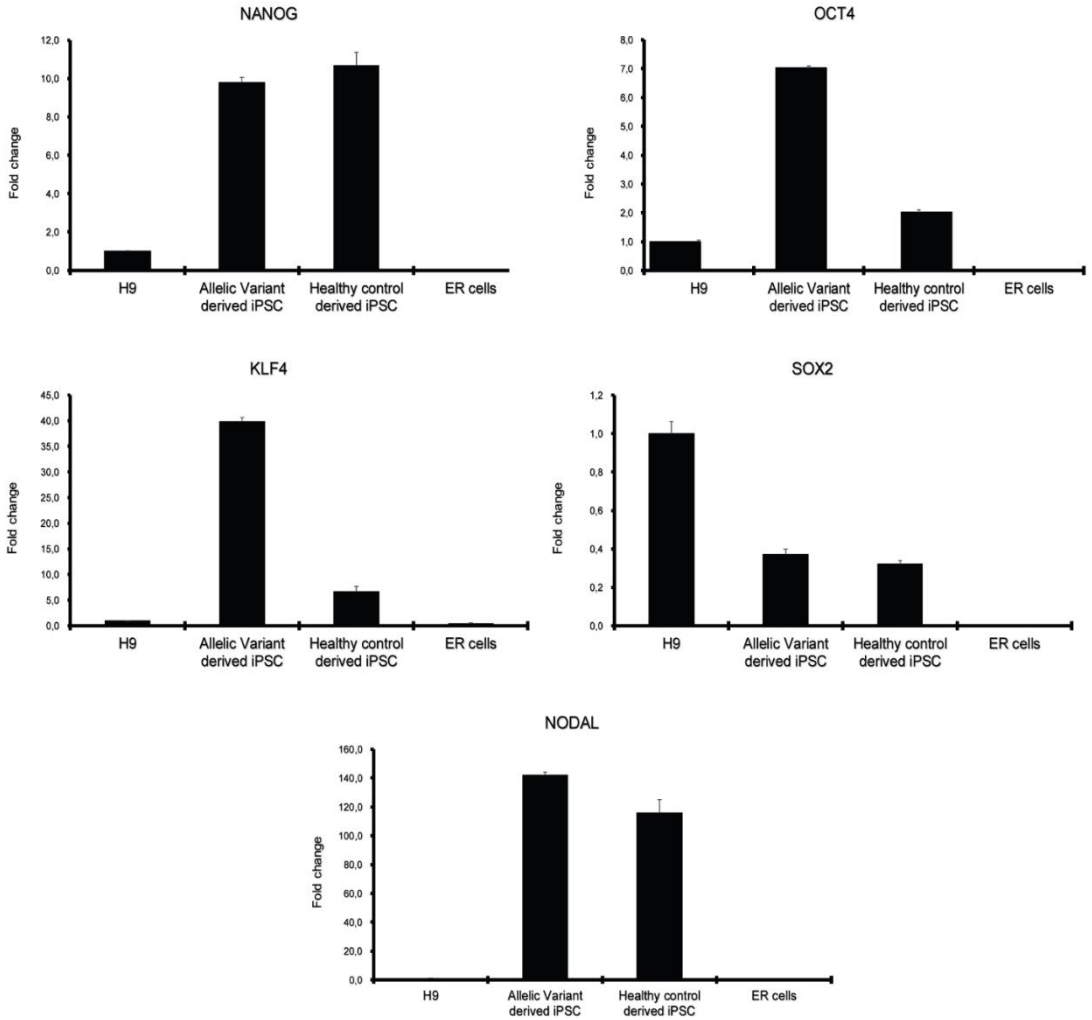
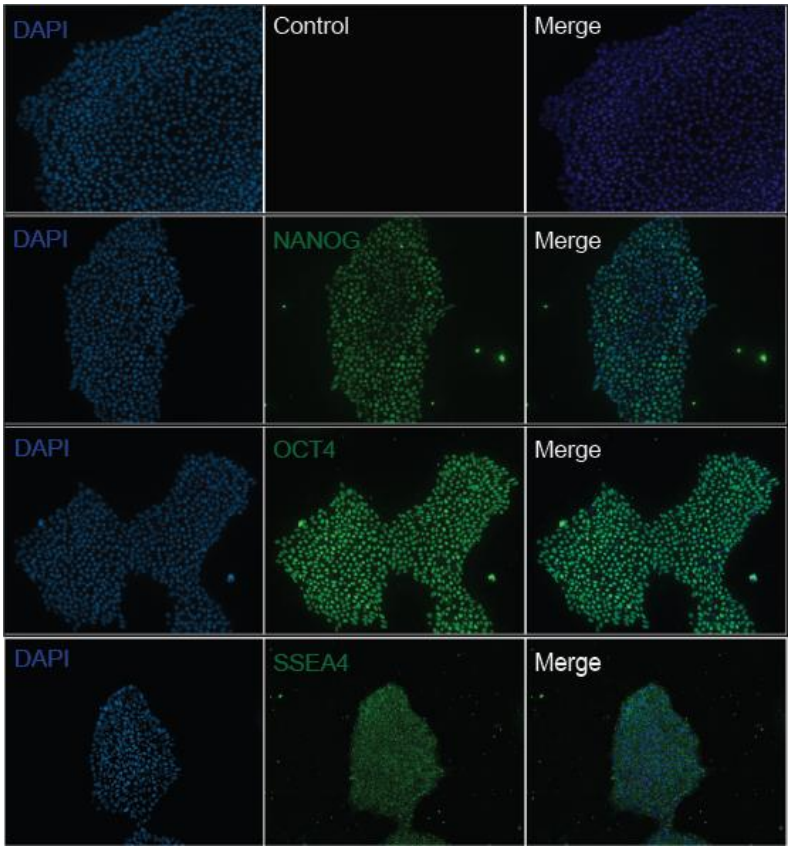


Figure 3.20 – Pluripotency analysis at mRNA level.

Expression of endogenous pluripotency-associated markers *NANOG*, *OCT3/4*, *SOX2*, *KLF4*, and *NODAL* in both control and patient-derived iPSC lines were confirmed by qPCR. H9 and exfoliated renal epithelial cells were used as positive and negative controls, respectively. CT-values were normalized to the geometric mean of the two housekeeping genes *GAPDH* and *β-actin* and with H9 human embryonic stem cell line as reference (set to 1).

At the protein level, immunofluorescence analysis, of the same clones tested for pluripotency gene expression, confirmed the presence of the self-renewal transcription markers NANOG, OCT4, and the surface marker SSEA4 characteristic of pluripotent embryonic stem cells in all cells from the colonies, illustrating also the homogeneity and purity of the newly-generated iPSC lines. These markers were equally expressed in both control and patient-derived iPSCs without obvious differences between both cell lines (Figure 3.21).



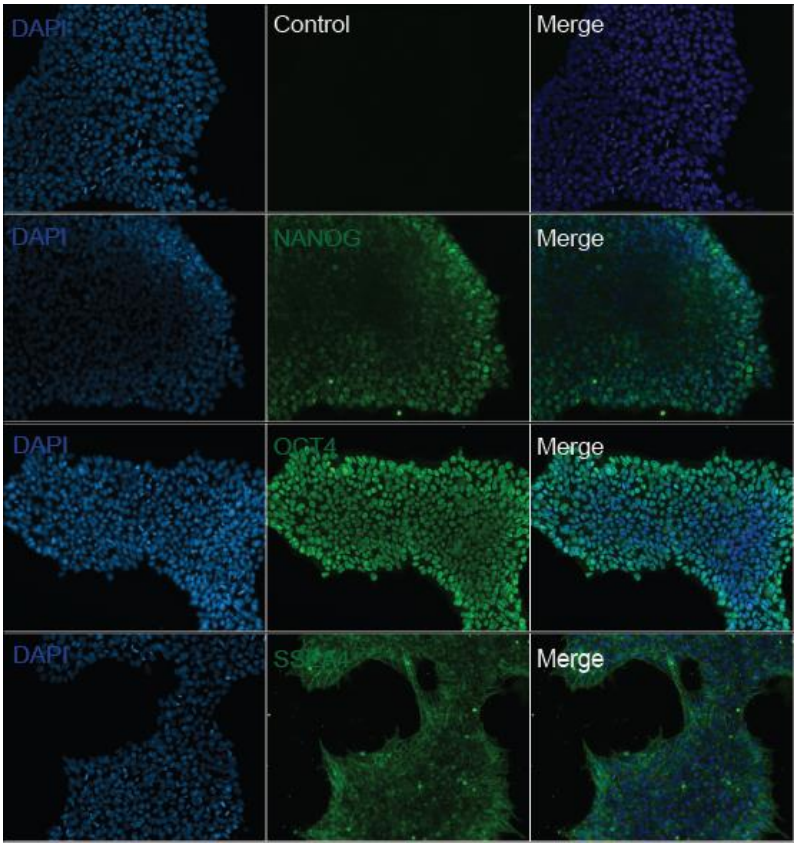


Figure 3.21 – Pluripotency analysis at protein level.

Immunofluorescence showing positive detection of the pluripotency-associated markers NANOG, OCT4 and SSEA4 at protein level in both patient and control derived IPSC lines.

3.2.4. Karyotype

To confirm that the iPSCs obtained through the transduction of the CytoTune® vectors, KOS (Klf4–Oct3/4–Sox2), hc-Myc, and hKlf4 into the renal epithelial cells maintain a proper genomic content, we performed the karyotype of the patient-derived iPSCs at passage 12. As presented in the figure 3.22 the karyotype of the patient-derived iPSCs seems to be normal, presenting 22 pairs of autosomal chromosomes and 1 pair of sex chromosomes. Nevertheless, a more detailed analysis of the chromosomes banding is currently ongoing to exclude possible subtle deletions, inversions, insertions, translocations, fragile sites and other more complex rearrangements.

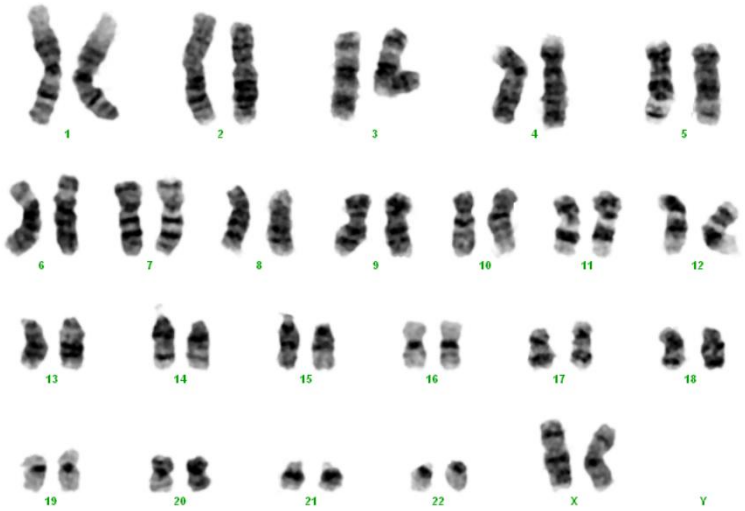


Figure 3.22 – Karyotype of patient-derived iPSCs.

Karyotype showing the normal 46, XX chromosomal arrangement of the patient-derived iPSCs.

The differentiation potential towards the all three germ layers from both control and patient iPSC lines is also important for the complete characterization of the iPSC lines. However, we were unable perform this analysis so far due to time constrains. We believe we should be able to complete this analysis in the near future. Despite that, the data obtained so far for the characterization of the iPSC lines indicate that we were able to successfully generate iPSC lines starting from urinary samples from control and the patient containing the *DAND5* genomic variant identified during this work. Moreover, this battery of characterization assays allows us to select the clones that will be used to functionally determine the effect of the DAND5 alteration in a cardiomyocyte specific context. This strategy eliminates the issues regarding the invasive fibroblast collection procedures, with the advantage of facilitating the study of controls and patients' samples, broadening the potential for iPSC–guided personalized medicine.

Chapter 4

General Discussion

The formation of the heart involves an incredible and precisely orchestrated series of molecular and morphogenetic events that goes from commitment of unique lineages that will form distinct regions of the heart to decisions of cell migration, proliferation and death that ultimately result in the formation of a mature four-chambered heart. The complexity of the molecular cascades involved in these events may explain the sensitivity of the heart to perturbations in these processes, which have catastrophic consequences in the form of congenital heart defects. Congenital heart diseases impose a massive burden on societies around the world due to its incidence and prevalence, together with the fact that almost all the children that survive need continuous treatment during all his life (Fahed et al., 2013). In addition, the limited capacity of the heart to regenerate and the difficulty to uncover the etiology of most of CHD represents the major challenges for current medical research, practice and geneticists. Therefore, it is urgent to better understand the mechanisms and genetic causes underlying heart development and congenital heart defects in order to generate genetic testing, effective therapies and to adopt genetic counseling (Cowan & Ware, 2015).

Although the major cardiovascular structures develop between the third and eighth week of gestation (Cecchetto et al., 2010), we cannot exclude that CHDs may be originated earlier, when the cardiac progenitor cells undergo their commitment and migration to the cardiac fields. During this migration, the cardiac progenitor cells are submitted to different signals in the right or left lateral plate mesoderm. Thus, the mechanism and molecules that regulate normal left-right establishment, including this asymmetric signaling in the lateral plate mesoderm, must also be considerate as potential causes or contributors to the CHDs presented by some patients. In fact, several phenotypes observed in children with complex heart disease are similar to those found in mice models for laterality disorders (Peeters & Devriendt, 2006), which led us to hypothesize that mutations in these genes could be responsible for some cases of human *situs* abnormalities and/or CHD.

Left-right axis establishment is a highly conserved feature in which the Nodal pathway seems to be one of the major players (Shiratori & Hamada, 2014). Genetic alterations, in one or more genes, that alter the normal threshold of the

Nodal pathway and ultimately disrupt the normal left-right axis establishment, can cause a range of outcomes in the normal *situs* of the body. This includes *situs inversus*, a complete inversion of the axis, heterotaxy, an arrangement of internal organs somewhere between *situs solitus* and *situs inversus*, and also *isomerism*, meaning a bilateral sidedness of the organs (Ramsdell, 2005). These situations get worse when a wide spectrum of complex heart defects is associated, what normally happens (S. J. Kim, 2011).

The work presented in this dissertation aimed to contribute for the understanding of the mechanisms and genetic causes underlying congenital heart defects arising from left-right disturbances in the initial developmental stages of life, before the beginning of heart morphogenesis. Our first focus was to translate the knowledge and observations obtained with the generation and analysis of the *Cerl2*/*Dand5* KO mice (Araujo et al., 2014; Marques et al., 2004), which represents a model of laterality defects associated with a vast array of CHD, into new discoveries that allow to answer to the etiology of some CHDs in humans. To do that, we performed a genetic screening in the *DAND5* genomic region of a 38 pediatric children cohort. DAND5 is a secreted growth factor that binds directly to NODAL and inhibits its signalling pathway during the formation of the embryonic left-right axis, being therefore an essential molecule in the establishment of this process (Marques et al., 2004). *Cerl2*/*Dand5* KO mice embryos show a variety of laterality anomalies, including *situs inversus*, heterotaxy and right and left isomerism. Additionally, the mice presented a vast array of heart malformations, including randomized positioning of the cardiac apex, incomplete atrial and/or ventricular septation, inversion of the loop (leftward loop) and failure of the outflow tract rotation leading to TGA and DORV and severe hyperplasia of the ventricular myocardium. Moreover, a significant mortality rate within a few hours after birth is observed (Araujo et al., 2014; Marques et al., 2004).

In addition to the *DAND5* gene, we also extend the genetic screening to other three genes involved in the Nodal signaling pathway, *NODAL*, *CFC1* and *PITX2C* in the same patients cohort. The cohort was composed with Caucasian non-familiar and isolated cases of CHD and/or laterality defects. Only seven patients presented clear laterality defects (left or right isomerism and *situs*

inversus) associated with cardiac malformations. Moreover, almost all the subjects manifest more than one cardiac defect, which is in agreement with the complexity, heterogeneity and diversity of CHD trait (Gelb & Chung, 2014).

The sequencing analysis of the 38 patients cohort revealed three exonic variants, one of them in the NODAL antagonist – *DAND5/Cerl2*, one in the member of the TGF beta superfamily – *NODAL* and another in the NODAL co-receptor – *CFC1* and three intronic sequence variants one in the *NODAL* gene and two in the Nodal signaling effector transcription factor – *PITX2* isoform C.

DAND5 allelic variant

The *DAND5* 455G>A allelic variant was identified in two of our 38 patients cohort. One of the patients presented ventricular septal defect with overriding aorta, pulmonary atresia and left isomerism, while the other presented overriding aorta, ventricular septal defect, right ventricular hypertrophy and pulmonary atresia (a case of extreme tetralogy of Fallot). The functional effect of the variant, which leads to a substitution of an arginine for an histidine in the position 152 of the highly conserved cysteine-rich DAN domain, was further evaluated using a Nodal-dependent luciferase assay. The results showed a substantial decreased in the function of the DAND5 variant protein when compared to the wild-type DAND5 protein [see Chapter 3, Part I]. Although the phenotypes of both patients have a great degree of similarity, we cannot make a clearly phenotype-genotype correlation mostly because this variant was also annotated in public databases as single nucleotide variation and the apparently phenotypically normal mother of one of the patients presented the variant in its genome. Nevertheless, we cannot exclude the presence of a mild or undiagnosed heart defect on the mother of patient 2, or on the supposed healthy individuals of the genome projects, since detailed examination of these subjects is rarely available (Richards et al., 2015).

Attending that these complex disease phenotypes rarely were completely predicted by the genotype, mostly due to its genetic heterogeneity, incomplete penetrance, variable expressivity and dose-sensitive genes (Cooper, Krawczak,

Polychronakos, Tyler-Smith, & Kehrer-Sawatzki, 2013), we cannot exclude that a functionally deficient common low penetrance genetic variant may act as modifier, influencing the phenotypic expression of a CHD alteration gene, depending on the individual genetic background (Cooper et al., 2013). Modifiers are genes capable of modifying the manifestation of a mutant gene without having an obvious effect on the normal condition (Grüneberg, 1963). So, the presence of a functionally deficient common modifier in human population would provide the opportunity for an individually rare deleterious sporadic mutation to manifest itself (Roessler et al., 2009). In other words, the offspring of a clinically normal mutation carrier and a common context-dependent modifier gene carrier might has an enhanced risk for malformations. In this view, neither parents are directly affected by being carriers of a deleterious gene; it is only the combination of alleles in the offspring that results in a recognizable clinical outcome.

NODAL H165R variant

The *NODAL* 494 A>G exonic variant, found in almost 90% of our patients, is localized in the pro-domain of the NODAL protein and leads to a amino acid substitution of an Histidine for one Arginine in the residue 165 (H165R). Taking into account that the probands with this alteration have a broad and different range of congenital heart defects and because the variant is referred in the public databases (NCBI, Ensemble and ExAC) as an extremely common SNP (rs1904589) we cannot make a genotype-phenotype correlation, neither associate the 494 G allele as a disease allelic variant.

However, in 2008 and 2009, Maximilian Muenke and co-workers reported this variant in probands with a vast spectrum of cardiac malformations and/or laterality defects (Roessler, et al., 2008). In these studies, the author confirmed that the H165R variation was extremely common, being present in heterozygosity in approximately 50% of the normal controls. Moreover, they also verify the effect of the variant using a NODAL luciferase assay as a readout of the Nodal signaling activity and calculated the relative risk association between conotruncal defects and the H165H, H165R and R165R

genotypes. Their analysis showed that the variant allele leads to a great decrease (50%) in the functional activity of this mutant NODAL protein compared to the NODAL wild-type protein and that the H165H genotype were associated with decrease risk of conotruncal defects as compared to the H165R and R165R. Nevertheless, these relative risk was not statistically significant, suggesting that if the variant alleles (H165R/R165R) were associated with disease risk, it is a very week risk or only in a subset of conotruncal defects (Roessler et al., 2009). Altogether, these results indicate that this variant can alter the function of the NODAL protein, but besides being common in the general population, this alteration alone appears to be neutral to selection pressure since it is present both in chimpanzees and humans, as shown by their evolutionary analysis (Roessler et al., 2009). Therefore, once this alteration is so common among the normal population, and since there are not many amino acids conserved in the pro-domain where it is located they proposed that this alteration may serve as modifier of variants that occur in others Nodal signaling pathway components, with which NODAL normally interact. Modifier genes is one of the mechanisms responsible for reduced penetrance as shown in different inherited diseases such as pancreatitis, breast cancer, Gaucher disease and hypertrophic cardiomyopathy among other (Cooper et al., 2013). This idea is also supported by the evidence that differences in phenotype are observed when identical gene alleles are crossed in different strain background mice, suggesting that these phenotypic differences can be attributed to specific-strain modifiers (Nadeau, 2001; Roessler et al., 2008).

Intronic Variants

The splicing of introns out of pre-mRNA to create mature mRNA molecule requires hundreds of proteins and several small ribonucleoproteins, known as splicing machinery, that must act together to excise precisely the right intron (Ward & Cooper, 2010). Due to its complexity, this process is prone to genomic alteration that can result in the disruption of the splicing process of certain gene. Indeed, ~15% of known human disease-causing alterations are thought to affect

splicing (Fraser & Xie, 2009). Therefore, despite exonic variants are the major cause of alterations in the functional activity of a mature protein due to amino acid substitution, the deletions, insertions, and/or intronic variants in the exon-intron boundaries could alter the proper maturation of the mRNA and should not be underestimated (Cartegni, Chew, & Krainer, 2002).

In the strategy defined for this study, the primers designed to amplify the exons of the genes were localized in the intronic regions near the exon-intron boundaries. This strategy was performed with the purpose of identify new potential disease intronic variants and allowed us to identify three intronic variants, one in the *NODAL* gene and two in the Nodal signaling effector transcription factor – *PITX2* isoform C. As shown on Chapter 3 - Part II, the *NODAL* intronic variant, which is located twelve nucleotides downstream to the last coding triple (c.193+12C>T), was found in a total of twenty patients. Since it was present in patients with a vast array of phenotypes, from a simple case of isolated atrial septal defect to a complex right isomerism with asplenia, transposition of the great arteries, atrioventricular septal defect and pulmonary stenosis total anomalous pulmonary venous connection and was also annotated in the public databases (Ensemble, NCBI and ExAC) as an extremely common SNP – rs10999338 – with a heterozygosity of 45% and a minor allele frequency (T) of 44% in the European population sequenced in the 1000 genomes project, we cannot conclude that this variant is a disease allele neither establish a genotypic-phenotype correlation. However, and because this alteration is located very close to a splice site, we cannot discard the hypothesis that this c.193+12 C>T variant can be affecting the process of splicing, namely the interaction between the proteins of spliceosome and the RNA. In fact, an intronic alteration 12 nucleotides after the end of one exon of the BRCA2 gene was identified and related to splicing aberrations in a breast cancer study (Whiley et al., 2011). The *in silico* analysis presented in the Chapter 3, Part II, showed that c.193+12 C>T alteration may lead to the creation of an important enhancer motif for the ligation of the SF2/ASF splicing factor and a creation of a motif that allows the biding of the SRp55 protein with more affinity. SF2/ASF and SRp55 protein facilitates and promotes 5´ and 3´ splice sites recognition and exon definition by interacting and recruiting other spliceosome proteins

(Richardson et al., 2011). Thus, the creation of new motifs for the biding of these proteins may alter the definition of the exons, resulting in abnormal mRNA transcripts and ultimately the final protein as well its functional activity. We are currently designing a functional splicing assay in order to elucidate if this alteration really alters the splicing process and to understand its functional effect on the NODAL final protein and their correlation with the observed phenotypes.

The *PITX2C* intronic variants were found 71 and 88 nucleotides after the end of the coding region of the exon 1B of the *PITX2* isoform C gene. As shown in Chapter 3, part II, both variants were classified as SNP by the NCBI and Ensemble databases, and so far they were not associated to any disease or risk for develop a certain disease. In this study, we observed both variants in our patients but we cannot make a genotype-phenotype correlation or associate the variants as disease or disease risk variants. To make such correlation we will perform a functional assay to verify the truly effect of the variants in the splicing patterns of the *PITX2C* gene. Nevertheless, the *in silico* analysis performed using the Human Splicing Finder software gave us a clue of what could be effect of these variants in the splicing process of *PITX2C* gene. The analysis presented in Chapter 3, part II, shown that the c.205+71 A>G allelic alteration could result in the disruption of splice site acceptor and branch point site sequences and could lead to the creation of a new splice donor site. In addition, the *in silico* analysis also revealed that this variant could disrupt the splicing enhancer motifs responsible for the ligation of SF2/ASF and 9G8 and disrupt the silencer motif in which the hnRNP A1 protein binds.

The splice site acceptor, splice site donor and branch point site represent three core sequences elements that are essential for splicing, being recognized multiple times by components of the spliceosome during the processes of intron removal and exon joining (Ward & Cooper, 2010). Although these three components are essential for the splicing process, they are not sufficient for defining the exon-intron boundaries (Ward & Cooper, 2010). For that purpose, the splicing machinery requires additional cis-acting sequences, named as

intronic or exonic splicing enhancers (ISE or ESE) or silencers (ISS or ESS) and located within exons and introns, that recruit trans-acting splicing factors to ensure inclusion of constitutive exons or modulate the efficiency of splice site recognition, promoting alternative splicing (Ward & Cooper, 2010). Enhancer elements are bound by positive regulators that increase exon inclusion, such as the SR protein family, in contrast, silencer elements are bound by negative regulators that decrease exon inclusion, such as the abundant hnRNP proteins (Ward & Cooper, 2010). Therefore, if one of these sites could be potentially broken or newly created, it was reasonable to consider that the normal process of splicing could be alternated, either by exon skipping, activation of a cryptic splice site, or intron retention (Baralle & Baralle, 2005), all resulting in an abnormal mRNA molecule and an abnormal final *PITX2C* protein, ultimately leading to expression of different mRNA variants that affect disease susceptibility and severity. A more detailed analysis using functional assays is needed in order to elucidate the effect of this alterations in the isoform C of the *PITX2* final protein and their correlation with the phenotypes of the patients.

The experimental and *in silico* analysis for the variants identified in this study were performed taking into the account each variant isolated. However, many of the patients presented more than one of the variant alleles in their genomes, as shown in table S2 (Supplementary Material 7). For instance, the patients that presented the *DAND5* variant also carry two more alterations in the *NODAL* gene, the c.193+12 C>T and the p.H165R, and one in the *PITX2C* gene. These NODAL alterations may also be interfering with the interaction between NODAL and DAND5 proteins and if we consider that the common polymorphism NODAL p.H165R, which decreases the bioactivity of NODAL to 50%, acts as a modifier, as proposed by Roessler et al. 2009 (Roessler et al., 2009), the genetic interaction between the two allelic variants (NODAL (p.H165R) and DAND5 (p.R152H)) might increase even more the risk of disease. Moreover, the intronic alterations in *NODAL* and *PITX2C* may have a cumulative effect in the imbalance of the proper levels of the NODAL signaling. Therefore, it is important to appraise that the cumulative impairment of the Nodal pathway due to variants in the different genes might increase the susceptibility to CHD more

than the contribution of a single gene. This situation was also verified in others studies, namely the study conducted by Ressler and colleagues, in which they attribute the cumulative impairment of the Nodal signaling due to alterations in several genes of this pathway as a possible cause for the phenotype of some patients (Roessler et al., 2008) and the study performed by Lowe et.al, in which they verify that conditional holomorphic Nodal variants shown distinct phenotypes, depending on the Nodal dosage (Lowe, Yamada, & Kuehn, 2001).

Therefore, while isolated deleterious gene alterations are rare, collectively, common variants within a pathway of interacting genes add up to a considerable genetic burden. While simple heterozygosity for one of these loss-of-function alleles might not be considered sufficient, it can be considered necessary for the observed phenotype(s), which explains their common incidence.

Taken together, this evidences with our data from the DAND5 variant, indicate that the imbalanced dosage-sensitive in the Nodal signaling, due to alterations in several of their components, is a final remark conferring risk for laterality defects and related CHD.

Patient iPSC line generation

Cerl2, or its homolog in humans DAND5, is an extracellular protein belonging to the family of TGF-β/Nodal signaling antagonists Cerberus/DAN. Dand5/Cerl2-mediated antagonism of Nodal signaling requires Dand5/Cerl2 binding to the ligand Nodal, which consequently prevents the interaction of Nodal with the receptor and subsequent signaling activation. In mice, Cerl2 has an important role in the establishment of the left-right axis, regulating the Nodal signaling at the node and the transmission of the LR asymmetry information to the left-lateral plate mesoderm (LPM) in a precise time window (Inacio et al., 2013; Oki et al., 2009). In agreement with this, absence of *Cerl2* in *Cerl2* KO embryos leads to a wide range of asymmetry defects defined as situs ambiguus, and a significant mortality rate within a few hours after birth due to cardiac defects, being some of them laterality-associated ones (Marques et al., 2004).

Interestingly, it was also observed a massive increase of the ventricular myocardial walls in *Cerl2*/*DAND5* KO mice that cannot be explained by laterality defects (Araujo et al., 2014). This increased ventricular mass is caused by increased cardiomyocyte mitotic index associated with increased levels of phospho-Smad2, suggesting prolonged TGF-β/Nodal signaling in the heart (Araujo et al., 2014). Since, TGF-β signaling pathways are known to influence growth and early cardiac specification, proliferation and differentiation (Araujo et al., 2014; Klaus et al., 2012), Cerl2/DAND5, the only member of this family expressed in the heart, seems to be specifically required in the heart to control cardiomyocyte proliferation through the control of TGF-β/Nodal signaling (Araujo et al., 2014). Therefore, and once that the *DAND5* 455G>A, identified in two of our patients cohort, lead to an increasing in the Nodal signaling, we hypothesized that, like *Cerl2/DAND5* KO mice, this human variant most likely can also affect the proliferation and maturation of ventricular cardiomyocytes. To test this hypothesis, Patient-specific and control induced pluripotent stem cell lines were generated to model the disease of one of the patients that was identified carrying the DAND5 p.R152H variant. The iPSCs were obtained through a laborious protocol starting from cells isolated from urine samples of each individuals that were further reprogramed back to a stem cell state using the four Yamanaka factors, SOX2, OCT3/4, c-MYC and KLF4. The results obtained from the initial characterization of the iPSC lines, analyzing the morphology, pluripotency state and the karyotype of the iPSC derived clones, indicate that the cell lines were generated successfully. Nevertheless, a final characterization recurring to the analysis of the differentiation potential towards the all three germ layers from both control and patient iPSC lines is needed in order to really validate both cell lines as iPSC.

As soon as this final characterization is completed, we will proceed to the differentiation of these iPSC into cardiomyocytes to determine whether the proliferation is affected by the *DAND5* variant. To do that, we will perform a comparative analysis of the differentiation and proliferation potential of control and patient-derived iPSCs. Furthermore, we will perform the functional characterization of the differentiated cardiomyocytes, culturing these cells on a PDMS flexible substrate and paced by field stimulation. Single cell force

measurement will be performed with traction force microscopy using the changes in cellular dimensions and the substrate stiffness to calculate force generated at single cell resolution. Calcium cycling will also be concurrently accessed via fluorescence imaging with the calcium dye Fluo-4AM. Alternatively, action potential morphology will be accessed via fluorescence imaging with the transmembrane voltage reporter FluoVolt [28, 29]. Confocal imaging will be used to quantify the force generated, action potential morphology, and calcium cycling in these human stem cell derived CMs. If we observe an ectopic function or proliferation of the variant derived cardiomyocytes, we can functionally address and recover the cause of the defect(s) using drugs that may target these processes.

Conclusion

Altogether, the basic research approach with a translational medicine perspective presented in this thesis, allowed us to identify 5 variants in different genes of a pediatric unrelated CHD and/or laterality defect cohort. Although we cannot make a clearly genotype-phenotype correlation for each variant alone, we verify that the DAND5 p.R152H variant disrupts the proper levels of Nodal signaling in the lateral plate mesoderm and thus might function as a risk allele for CHDs and/laterality defects. Despite the fact that the phenotypes presented by the two carriers of the DAND5 variant were not exactly equal, they result from errors during the LR axis formation, representing a case of incomplete penetrance or variable expressivity. This is in agreement with the complexity of the CHD/ laterality trait and may reflect the action of modifier gene alterations or other genetic variants within the pathway that exacerbate or ameliorate the effect of the DAND5 alteration and the patients phenotypes. According with this hypothesis, it seems that the functional abnormal p.H165R NODAL variant could act as a modifier gene variant and thus might play a role in the phenotypes of the patients. Furthermore, we cannot exclude the possibility that the *PITX2C* variants and the *NODAL* intronic variant, which were predicted to alter the splicing process of the mRNA transcripts, act in the same way.

Therefore, the allelic variants found in this study might not function as disease alleles per si but probably increase the susceptibility to disease. Nevertheless, since most of the patients presented more than one alteration, the cumulative effect of each variant within the pathway seems to enhance even more the risk for disease. This, prompt us to evaluated the potential causes of CHDs through a multi-gene approach, testing several genes involved in the same pathway. Moreover, we start to look further, and using the currently well-established iPSC technology we successfully generated patient and control iPS derived cells. Those iPSCs will be used in the future to understand the cellular and molecular mechanisms of disease behind a simple nucleotide variant, allowing us to study the role of other interacting variants with the final purpose of developing a personalized medicine for each patient.

Chapter 5

References

References

Adachi, H., Saijoh, Y., Mochida, K., Ohishi, S., Hashiguchi, H., Hirao, A., & Hamada, H. (1999). Determination of left/right asymmetric expression of nodal by a left side-specific enhancer with sequence similarity to a lefty-2 enhancer. *Genes Dev, 13*(12), 1589-1600.

Alarcon, V. B., & Marikawa, Y. (2003). Deviation of the blastocyst axis from the first cleavage plane does not affect the quality of mouse postimplantation development. *Biol Reprod, 69*(4), 1208-1212. doi: 10.1095/biolreprod.103.018283

Amack, J. D. (2014). Salient features of the ciliated organ of asymmetry. *Bioarchitecture, 4*(1), 6-15. doi: 10.4161/bioa.28014

Andersen, T. A., Troelsen Kde, L., & Larsen, L. A. (2014). Of mice and men: molecular genetics of congenital heart disease. *Cell Mol Life Sci, 71*(8), 1327-1352. doi: 10.1007/s00018-013-1430-1

Anderson, R. H., Spicer, D. E., Brown, N. A., & Mohun, T. J. (2014). The development of septation in the four-chambered heart. *Anat Rec (Hoboken), 297*(8), 1414-1429. doi: 10.1002/ar.22949

Araujo, A. C., Marques, S., & Belo, J. A. (2014). Targeted inactivation of Cerberus like-2 leads to left ventricular cardiac hyperplasia and systolic dysfunction in the mouse. *PLoS One, 9*(7), e102716. doi: 10.1371/journal.pone.0102716

Arnold, S. J., & Robertson, E. J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat Rev Mol Cell Biol, 10*(2), 91-103. doi: 10.1038/nrm2618

Avior, Y., Sagi, I., & Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. *Nat Rev Mol Cell Biol, 17*(3), 170-182. doi: 10.1038/nrm.2015.27

Azhar, M., & Ware, S. M. (2016). Genetic and Developmental Basis of Cardiovascular Malformations. *Clin Perinatol, 43*(1), 39-53. doi: 10.1016/j.clp.2015.11.002

Babu, D., & Roy, S. (2013). Left-right asymmetry: cilia stir up new surprises in the node. *Open Biol, 3*(5), 130052. doi: 10.1098/rsob.130052

Bamford, R. N., Roessler, E., Burdine, R. D., Saplakoglu, U., dela Cruz, J., Splitt, M., . . . Casey, B. (2000). Loss-of-function mutations in the EGF-CFC gene CFC1 are associated with human left-right laterality defects. *Nat Genet, 26*(3), 365-369. doi: 10.1038/81695

Baralle, D., & Baralle, M. (2005). Splicing in action: assessing disease causing sequence changes. *J Med Genet, 42*(10), 737-748. doi: 10.1136/jmg.2004.029538

Bayzigitov, D. R., Medvedev, S. P., Dementyeva, E. V., Bayramova, S. A., Pokushalov, E. A., Karaskov, A. M., & Zakian, S. M. (2016). Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes Afford New Opportunities in Inherited Cardiovascular Disease Modeling. *Cardiol Res Pract, 2016*, 3582380. doi: 10.1155/2016/3582380

Beddington, R. S., & Robertson, E. J. (1999). Axis development and early asymmetry in mammals. *Cell, 96*(2), 195-209.

Bellin, M., Marchetto, M. C., Gage, F. H., & Mummery, C. L. (2012). Induced pluripotent stem cells: the new patient? *Nat Rev Mol Cell Biol, 13*(11), 713-726. doi: 10.1038/nrm3448

Belo, J. A., Silva, A. C., Borges, A. C., Filipe, M., Bento, M., Goncalves, L., . . . Marques, S. (2009). Generating asymmetries in the early vertebrate embryo: the role of the Cerberus-like family. *Int J Dev Biol, 53*(8-10), 1399-1407. doi: 10.1387/ijdb.072297jb

Bondue, A., Lapouge, G., Paulissen, C., Semeraro, C., Iacovino, M., Kyba, M., & Blanpain, C. (2008). Mesp1 acts as a master regulator of multipotent cardiovascular progenitor specification. *Cell Stem Cell, 3*(1), 69-84. doi: 10.1016/j.stem.2008.06.009

Brennan, J., Norris, D. P., & Robertson, E. J. (2002). Nodal activity in the node governs left-right asymmetry. *Genes Dev, 16*(18), 2339-2344. doi: 10.1101/gad.1016202

Brown, N. A., & Wolpert, L. (1990). The development of handedness in left/right asymmetry. *Development, 109*(1), 1-9.

Capdevila, J., Vogan, K. J., Tabin, C. J., & Izpisua Belmonte, J. C. (2000). Mechanisms of left-right determination in vertebrates. *Cell, 101*(1), 9-21. doi: 10.1016/S0092-8674(00)80619-4

Carlson, B. M. (2014). *Human embryology and developmental biology* (5th ed.). Philadelphia, PA: Elsevier/Saunders.

Cartegni, L., Chew, S. L., & Krainer, A. R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet, 3*(4), 285-298. doi: 10.1038/nrg775

Cecchetto, A., Rampazzo, A., Angelini, A., Bianco, L. D., Padalino, M., Stellin, G., & Daliento, L. (2010). From molecular mechanisms of cardiac development to genetic substrate of congenital heart diseases. *Future Cardiol, 6*(3), 373-393. doi: 10.2217/fca.10.10

Channabasappa, S. M., Mohan, H. S., & Sarma, J. (2013). A patient with situs inversus totalis presenting for emergency laparoscopic appendectomy: Consideration for safe anesthetic management. *Anesth Essays Res, 7*(1), 127-129. doi: 10.4103/0259-1162.114019

Chong, J. J., Forte, E., & Harvey, R. P. (2014). Developmental origins and lineage descendants of endogenous adult cardiac progenitor cells. *Stem Cell Res, 13*(3 Pt B), 592-614. doi: 10.1016/j.scr.2014.09.008

Cochard, L. R., & Netter, F. H. (2002). *Netter's atlas of human embryology* (1st ed.). Teterboro, N.J.: Icon Learning Systems.

Cohen, M. S., Anderson, R. H., Cohen, M. I., Atz, A. M., Fogel, M., Gruber, P. J., . . . Weinberg, P. M. (2007). Controversies, genetics, diagnostic assessment, and outcomes relating to the heterotaxy syndrome. *Cardiol Young, 17 Suppl 2*, 29-43. doi: 10.1017/S104795110700114X

Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C., & Kehrer-Sawatzki, H. (2013). Where genotype is not predictive of phenotype: towards an understanding of the molecular basis of reduced penetrance in human inherited disease. *Hum Genet, 132*(10), 1077-1130. doi: 10.1007/s00439-013-1331-2

Cowan, J. R., & Ware, S. M. (2015). Genetics and genetic testing in congenital heart disease. *Clin Perinatol, 42*(2), 373-393, ix. doi: 10.1016/j.clp.2015.02.009

De Luca, A., Sarkozy, A., Consoli, F., Ferese, R., Guida, V., Dentici, M. L., . . . Dallapiccola, B. (2010). Familial transposition of the great arteries caused by multiple mutations in laterality genes. *Heart, 96*(9), 673-677. doi: 10.1136/hrt.2009.181685

Dell'Era, P., Benzoni, P., Crescini, E., Valle, M., Xia, E., Consiglio, A., & Memo, M. (2015). Cardiac disease modeling using induced pluripotent stem cell-derived human cardiomyocytes. *World J Stem Cells, 7*(2), 329-342. doi: 10.4252/wjsc.v7.i2.329

Deng, H., Xia, H., & Deng, S. (2015). Genetic basis of human left-right asymmetry disorders. *Expert Rev Mol Med, 16*, e19. doi: 10.1017/erm.2014.22

Deng, X., Zhou, J., Li, F. F., Yan, P., Zhao, E. Y., Hao, L., . . . Liu, S. L. (2014). Characterization of nodal/TGF-lefty signaling pathway gene variants for possible roles in congenital heart diseases. *PLoS One, 9*(8), e104535. doi: 10.1371/journal.pone.0104535

Dunwoodie, S. L. (2007). Combinatorial signaling in the heart orchestrates cardiac induction, lineage specification and chamber formation. *Semin Cell Dev Biol, 18*(1), 54-66. doi: 10.1016/j.semcdb.2006.12.003

Dykes, I. (2014). Left Right Patterning, Evolution and Cardiac Development. *Journal of Cardiovascular Development and Disease, 1*(1), 52.

Ebert, A. D., Diecke, S., Chen, I. Y., & Wu, J. C. (2015). Reprogramming and transdifferentiation for cardiovascular development and regenerative medicine: where do we stand? *EMBO Mol Med, 7*(9), 1090-1103. doi: 10.15252/emmm.201504395

Fahed, A. C., Gelb, B. D., Seidman, J. G., & Seidman, C. E. (2013). Genetics of congenital heart disease: the glass half empty. *Circ Res, 112*(4), 707-720. doi: 10.1161/CIRCRESAHA.112.300853

Fahed, A. C., & Nemer, G. M. (2012). Genetic Causes of Syndromic and Non-Syndromic Congenital Heart Disease. In D. Cooper (Ed.), *Mutations in Human Genetic Disease*: InTech.

Field, S., Riley, K. L., Grimes, D. T., Hilton, H., Simon, M., Powles-Glover, N., . . . Norris, D. P. (2011). Pkd1l1 establishes left-right asymmetry and physically interacts with Pkd2. *Development, 138*(6), 1131-1142. doi: 10.1242/dev.058149

Fraser, H. B., & Xie, X. (2009). Common polymorphic transcript variation in human disease. *Genome Res, 19*(4), 567-575. doi: 10.1101/gr.083477.108

Furtado, M. B., Biben, C., Shiratori, H., Hamada, H., & Harvey, R. P. (2011). Characterization of Pitx2c expression in the mouse heart using a reporter transgene. *Dev Dyn, 240*(1), 195-203. doi: 10.1002/dvdy.22492

Gaio, U., Schweickert, A., Fischer, A., Garratt, A. N., Muller, T., Ozcelik, C., . . . Birchmeier, C. (1999). A role of the cryptic gene in the correct establishment of the left-right axis. *Curr Biol, 9*(22), 1339-1342.

Gelb, B. D., & Chung, W. K. (2014). Complex genetics and the etiology of human congenital heart disease. *Cold Spring Harb Perspect Med, 4*(7), a013953. doi: 10.1101/cshperspect.a013953

Gilbert, S. F. (2010). *Developmental biology* (9th ed.). Sunderland, Mass.: Sinauer Associates.

Goldmuntz, E., Bamford, R., Karkera, J. D., dela Cruz, J., Roessler, E., & Muenke, M. (2002). CFC1 mutations in patients with transposition of the great arteries and double-outlet right ventricle. *Am J Hum Genet, 70*(3), 776-780. doi: 10.1086/339079

Goodship, J. A., Hall, D., Topf, A., Mamasoula, C., Griffin, H., Rahman, T. J., . . . Keavney, B. (2012). A common variant in the PTPN11 gene contributes

to the risk of tetralogy of Fallot. *Circ Cardiovasc Genet, 5*(3), 287-292. doi: 10.1161/CIRCGENETICS.111.962035

Grüneberg, H. (1963). *The pathology of development; a study of inherited skeletal disorders in animals*. New York: J. Wiley.

Hamada, H., Meno, C., Watanabe, D., & Saijoh, Y. (2002). Establishment of vertebrate left-right asymmetry. *Nat Rev Genet, 3*(2), 103-113. doi: 10.1038/nrg732

Hamada, H., & Tam, P. P. (2014). Mechanisms of left-right asymmetry and patterning: driver, mediator and responder. *F1000Prime Rep, 6*, 110. doi: 10.12703/P6-110

Harvey, R. P. (2002). Patterning the vertebrate heart. *Nat Rev Genet, 3*(7), 544-556. doi: 10.1038/nrg843

Hashimoto, M., & Hamada, H. (2010). Translation of anterior-posterior polarity into left-right polarity in the mouse embryo. *Curr Opin Genet Dev, 20*(4), 433-437. doi: 10.1016/j.gde.2010.04.002

Hirokawa, N., Tanaka, Y., & Okada, Y. (2009). Left-right determination: involvement of molecular motor KIF3, cilia, and nodal flow. *Cold Spring Harb Perspect Biol, 1*(1), a000802. doi: 10.1101/cshperspect.a000802

Hirokawa, N., Tanaka, Y., & Okada, Y. (2012). Cilia, KIF3 molecular motor and nodal flow. *Curr Opin Cell Biol, 24*(1), 31-39. doi: 10.1016/j.ceb.2012.01.002

Hirokawa, N., Tanaka, Y., Okada, Y., & Takeda, S. (2006). Nodal flow and the generation of left-right asymmetry. *Cell, 125*(1), 33-45. doi: 10.1016/j.cell.2006.03.002

Huang, J. B., Liu, Y. L., Sun, P. W., Lv, X. D., Du, M., & Fan, X. M. (2010). Molecular mechanisms of congenital heart disease. *Cardiovasc Pathol, 19*(5), e183-193. doi: 10.1016/j.carpath.2009.06.008

Icardo, J. M., Garcia Rincon, J. M., & Ros, M. A. (2002). [Congenital heart disease, heterotaxia and laterality]. *Rev Esp Cardiol, 55*(9), 962-974.

Inacio, J. M., Marques, S., Nakamura, T., Shinohara, K., Meno, C., Hamada, H., & Belo, J. A. (2013). The dynamic right-to-left translocation of Cerl2 is involved in the regulation and termination of Nodal activity in the mouse node. *PLoS One, 8*(3), e60406. doi: 10.1371/journal.pone.0060406

Jouni, M., Si-Tayeb, K., Es-Salah-Lamoureux, Z., Latypova, X., Champon, B., Caillaud, A., . . . Gaborit, N. (2015). Toward Personalized Medicine: Using Cardiomyocytes Differentiated From Urine-Derived Pluripotent Stem Cells to Recapitulate Electrophysiological Characteristics of Type 2 Long QT Syndrome. *J Am Heart Assoc, 4*(9), e002159. doi: 10.1161/JAHA.115.002159

Kamura, K., Kobayashi, D., Uehara, Y., Koshida, S., Iijima, N., Kudo, A., . . . Takeda, H. (2011). Pkd1l1 complexes with Pkd2 on motile cilia and functions to establish the left-right axis. *Development, 138*(6), 1121-1129. doi: 10.1242/dev.058271

Karcher, C., Fischer, A., Schweickert, A., Bitzer, E., Horie, S., Witzgall, R., & Blum, M. (2005). Lack of a laterality phenotype in Pkd1 knock-out embryos correlates with absence of polycystin-1 in nodal cilia. *Differentiation, 73*(8), 425-432. doi: 10.1111/j.1432-0436.2005.00048.x

Kathiriya, I. S., & Srivastava, D. (2000). Left-right asymmetry and cardiac looping: implications for cardiac development and congenital heart disease. *Am J Med Genet, 97*(4), 271-279.

Kawasumi, A., Nakamura, T., Iwai, N., Yashiro, K., Saijoh, Y., Belo, J. A., . . . Hamada, H. (2011). Left-right asymmetry in the level of active Nodal protein produced in the node is translated into left-right asymmetry in the lateral plate of mouse embryos. *Dev Biol, 353*(2), 321-330. doi: 10.1016/j.ydbio.2011.03.009

Kim, C. (2014). Disease modeling and cell based therapy with iPSC: future therapeutic option with fast and safe application. *Blood Res, 49*(1), 7-14. doi: 10.5045/br.2014.49.1.7

Kim, S. J. (2011). Heterotaxy syndrome. *Korean Circ J, 41*(5), 227-232. doi: 10.4070/kcj.2011.41.5.227

Kirby, M. L. (2007). *Cardiac development*. Oxford ; New York: Oxford University Press.

Kitamura, K., Miura, H., Miyagawa-Tomita, S., Yanazawa, M., Katoh-Fukui, Y., Suzuki, R., . . . Yokoyama, M. (1999). Mouse Pitx2 deficiency leads to anomalies of the ventral body wall, heart, extra- and periocular mesoderm and right pulmonary isomerism. *Development, 126*(24), 5749-5758.

Klaus, A., Muller, M., Schulz, H., Saga, Y., Martin, J. F., & Birchmeier, W. (2012). Wnt/beta-catenin and Bmp signals control distinct sets of transcription factors in cardiac progenitor cells. *Proc Natl Acad Sci U S A, 109*(27), 10921-10926. doi: 10.1073/pnas.1121236109

Koefoed, K., Veland, I. R., Pedersen, L. B., Larsen, L. A., & Christensen, S. T. (2014). Cilia and coordination of signaling networks during heart development. *Organogenesis, 10*(1), 108-125. doi: 10.4161/org.27483

Kosaki, K., Bassi, M. T., Kosaki, R., Lewin, M., Belmont, J., Schauer, G., & Casey, B. (1999). Characterization and mutation analysis of human LEFTY A and LEFTY B, homologues of murine genes implicated in left-right axis development. *Am J Hum Genet, 64*(3), 712-721. doi: 10.1086/302289

Kosaki, R., Gebbia, M., Kosaki, K., Lewin, M., Bowers, P., Towbin, J. A., & Casey, B. (1999). Left-right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. *Am J Med Genet, 82*(1), 70-76.

Lalani, S. R., & Belmont, J. W. (2014). Genetic basis of congenital cardiovascular malformations. *Eur J Med Genet, 57*(8), 402-413. doi: 10.1016/j.ejmg.2014.04.010

Levin, M. (2003). Motor protein control of ion flux is an early step in embryonic left-right asymmetry. *Bioessays, 25*(10), 1002-1010. doi: 10.1002/bies.10339

Levin, M. (2004). The embryonic origins of left-right asymmetry. *Crit Rev Oral Biol Med, 15*(4), 197-206.

Liu, C., Liu, W., Palie, J., Lu, M. F., Brown, N. A., & Martin, J. F. (2002). Pitx2c patterns anterior myocardium and aortic arch vessels and is required for local cell movement into atrioventricular cushions. *Development, 129*(21), 5081-5091.

Lowe, L. A., Yamada, S., & Kuehn, M. R. (2001). Genetic dissection of nodal function in patterning the mouse embryo. *Development, 128*(10), 1831-1843.

Ma, L., Selamet Tierney, E. S., Lee, T., Lanzano, P., & Chung, W. K. (2012). Mutations in ZIC3 and ACVR2B are a common cause of heterotaxy and

associated cardiovascular anomalies. *Cardiol Young, 22*(2), 194-201. doi: 10.1017/S1047951111001181

Marjoram, L., & Wright, C. (2011). Rapid differential transport of Nodal and Lefty on sulfated proteoglycan-rich extracellular matrix regulates left-right asymmetry in Xenopus. *Development, 138*(3), 475-485. doi: 10.1242/dev.056010

Marques, S., Borges, A. C., Silva, A. C., Freitas, S., Cordenonsi, M., & Belo, J. A. (2004). The activity of the Nodal antagonist Cerl-2 in the mouse node is required for correct L/R body axis. *Genes Dev, 18*(19), 2342-2347. doi: 10.1101/gad.306504

Martin, J., Amendt, B., & Brown, N. (2010). Pitx2 in Cardiac Left–Right Asymmetry and Human Disease. In R. P. Harvey & N. Rosenthal (Eds.), *Heart Development and Regeneration* (Vol. 1, pp. 307-322): Elsevier

May-Simera, H. L., & Kelley, M. W. (2012). Cilia, Wnt signaling, and the cytoskeleton. *Cilia, 1*(1), 7. doi: 10.1186/2046-2530-1-7

McGrath, J., Somlo, S., Makova, S., Tian, X., & Brueckner, M. (2003). Two populations of node monocilia initiate left-right asymmetry in the mouse. *Cell, 114*(1), 61-73.

Megarbane, A., Salem, N., Stephan, E., Ashoush, R., Lenoir, D., Delague, V., . . . Bouvagnet, P. (2000). X-linked transposition of the great arteries and incomplete penetrance among males with a nonsense mutation in ZIC3. *Eur J Hum Genet, 8*(9), 704-708. doi: 10.1038/sj.ejhg.5200526

Meilhac, S. M., Lescroart, F., Blanpain, C., & Buckingham, M. E. (2014). Cardiac cell lineages that form the heart. *Cold Spring Harb Perspect Med, 4*(9), a013888. doi: 10.1101/cshperspect.a013888

Meno, C., Shimono, A., Saijoh, Y., Yashiro, K., Mochida, K., Ohishi, S., . . . Hamada, H. (1998). lefty-1 is required for left-right determination as a regulator of lefty-2 and nodal. *Cell, 94*(3), 287-297.

Miquerol, L., & Kelly, R. G. (2013). Organogenesis of the vertebrate heart. *Wiley Interdiscip Rev Dev Biol, 2*(1), 17-29. doi: 10.1002/wdev.68

Mohapatra, B., Casey, B., Li, H., Ho-Dawson, T., Smith, L., Fernbach, S. D., . . . Ware, S. M. (2009). Identification and functional characterization of NODAL rare variants in heterotaxy and isolated cardiovascular malformations. *Hum Mol Genet, 18*(5), 861-871. doi: 10.1093/hmg/ddn411

Moore, K. L., Persaud, T. V. N., & Torchia, M. G. *The developing human : clinically oriented embryology* (10th edition. ed.).

Mozaffarian, D., Benjamin, E. J., Go, A. S., Arnett, D. K., Blaha, M. J., Cushman, M., . . . Stroke Statistics, S. (2015). Heart disease and stroke statistics--2015 update: a report from the American Heart Association. *Circulation, 131*(4), e29-322. doi: 10.1161/CIR.0000000000000152

Nadeau, J. H. (2001). Modifier genes in mice and humans. *Nat Rev Genet, 2*(3), 165-174. doi: 10.1038/35056009

Nakamura, T., & Hamada, H. (2012). Left-right patterning: conserved and divergent mechanisms. *Development, 139*(18), 3257-3262. doi: 10.1242/dev.061606

Nakamura, T., Mine, N., Nakaguchi, E., Mochizuki, A., Yamamoto, M., Yashiro, K., . . . Hamada, H. (2006). Generation of robust left-right asymmetry in the mouse embryo requires a self-enhancement and lateral-inhibition system. *Dev Cell, 11*(4), 495-504. doi: 10.1016/j.devcel.2006.08.002

Nakamura, T., Saito, D., Kawasumi, A., Shinohara, K., Asai, Y., Takaoka, K., . . . Hamada, H. (2012). Fluid flow and interlinked feedback loops establish left-right asymmetric decay of Cerl2 mRNA. *Nat Commun, 3*, 1322. doi: 10.1038/ncomms2319

Nonaka, S., Tanaka, Y., Okada, Y., Takeda, S., Harada, A., Kanai, Y., . . . Hirokawa, N. (1998). Randomization of left-right asymmetry due to loss of nodal cilia generating leftward flow of extraembryonic fluid in mice lacking KIF3B motor protein. *Cell, 95*(6), 829-837.

Nonaka, S., Yoshiba, S., Watanabe, D., Ikeuchi, S., Goto, T., Marshall, W. F., & Hamada, H. (2005). De novo formation of left-right asymmetry by posterior tilt of nodal cilia. *PLoS Biol, 3*(8), e268. doi: 10.1371/journal.pbio.0030268

Norris, D. P. (2012). Cilia, calcium and the basis of left-right asymmetry. *BMC Biol, 10*, 102. doi: 10.1186/1741-7007-10-102

Norris, D. P., & Grimes, D. T. (2012). Developmental biology. Cilia discern left from right. *Science, 338*(6104), 206-207. doi: 10.1126/science.1230401

Norris, D. P., & Robertson, E. J. (1999). Asymmetric and node-specific nodal expression patterns are controlled by two distinct cis-acting regulatory elements. *Genes Dev, 13*(12), 1575-1588.

Nowotschin, S., Liao, J., Gage, P. J., Epstein, J. A., Campione, M., & Morrow, B. E. (2006). Tbx1 affects asymmetric cardiac morphogenesis by regulating Pitx2 in the secondary heart field. *Development, 133*(8), 1565-1573. doi: 10.1242/dev.02309

Okada, Y., Nonaka, S., Tanaka, Y., Saijoh, Y., Hamada, H., & Hirokawa, N. (1999). Abnormal nodal flow precedes situs inversus in iv and inv mice. *Mol Cell, 4*(4), 459-468.

Okada, Y., Takeda, S., Tanaka, Y., Izpisua Belmonte, J. C., & Hirokawa, N. (2005). Mechanism of nodal flow: a conserved symmetry breaking event in left-right axis determination. *Cell, 121*(4), 633-644. doi: 10.1016/j.cell.2005.04.008

Oki, S., Hashimoto, R., Okui, Y., Shen, M. M., Mekada, E., Otani, H., . . . Hamada, H. (2007). Sulfated glycosaminoglycans are necessary for Nodal signal transmission from the node to the left lateral plate in the mouse embryo. *Development, 134*(21), 3893-3904. doi: 10.1242/dev.009464

Oki, S., Kitajima, K., Marques, S., Belo, J. A., Yokoyama, T., Hamada, H., & Meno, C. (2009). Reversal of left-right asymmetry induced by aberrant Nodal signaling in the node of mouse embryos. *Development, 136*(23), 3917-3925. doi: 10.1242/dev.039305

Pediatric Cardiac Genomics, C., Gelb, B., Brueckner, M., Chung, W., Goldmuntz, E., Kaltman, J., . . . Rosenberg, E. (2013). The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. *Circ Res, 112*(4), 698-706. doi: 10.1161/CIRCRESAHA.111.300297

Peeters, H., & Devriendt, K. (2006). Human laterality disorders. *Eur J Med Genet, 49*(5), 349-362. doi: 10.1016/j.ejmg.2005.12.003

Praetorius, H. A., & Spring, K. R. (2001). Bending the MDCK cell primary cilium increases intracellular calcium. *J Membr Biol, 184*(1), 71-79.

Ramsdell, A. F. (2005). Left-right asymmetry and congenital cardiac defects: getting to the heart of the matter in vertebrate left-right axis determination. *Dev Biol, 288*(1), 1-20. doi: 10.1016/j.ydbio.2005.07.038

Rana, M. S., Christoffels, V. M., & Moorman, A. F. (2013). A molecular and genetic outline of cardiac morphogenesis. *Acta Physiol (Oxf), 207*(4), 588-615. doi: 10.1111/apha.12061

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi: 10.1038/gim.2015.30

Richardson, D. N., Rogers, M. F., Labadorf, A., Ben-Hur, A., Guo, H., Paterson, A. H., & Reddy, A. S. (2011). Comparative analysis of serine/arginine-rich proteins across 27 eukaryotes: insights into sub-family classification and extent of alternative splicing. *PLoS One, 6*(9), e24542. doi: 10.1371/journal.pone.0024542

Roessler, E., Ouspenskaia, M. V., Karkera, J. D., Velez, J. I., Kantipong, A., Lacbawan, F., . . . Muenke, M. (2008). Reduced NODAL signaling strength via mutation of several pathway members including FOXH1 is linked to human heart defects and holoprosencephaly. *Am J Hum Genet, 83*(1), 18-29. doi: 10.1016/j.ajhg.2008.05.012

Roessler, E., Pei, W., Ouspenskaia, M. V., Karkera, J. D., Velez, J. I., Banerjee-Basu, S., . . . Muenke, M. (2009). Cumulative ligand activity of NODAL mutations and modifiers are linked to human heart defects and holoprosencephaly. *Mol Genet Metab, 98*(1-2), 225-234. doi: 10.1016/j.ymgme.2009.05.005

Rosenthal, N., & Harvey, R. P. (2010). *Heart Development and Regeneration*: Academic Press -Elsevier.

Saijoh, Y., Oki, S., Tanaka, C., Nakamura, T., Adachi, H., Yan, Y. T., . . . Hamada, H. (2005). Two nodal-responsive enhancers control left-right asymmetric expression of Nodal. *Dev Dyn, 232*(4), 1031-1036. doi: 10.1002/dvdy.20192

Saijoh, Y., Viotti, M., & Hadjantonakis, A. K. (2014). Follow your gut: relaying information from the site of left-right symmetry breaking in the mouse. *Genesis, 52*(6), 503-514. doi: 10.1002/dvg.22783

Sauer, S., & Klar, A. J. (2012). Left-right symmetry breaking in mice by left-right dynein may occur via a biased chromatid segregation mechanism, without directly involving the Nodal gene. *Front Oncol, 2*, 166. doi: 10.3389/fonc.2012.00166

Saund, R. S., Kanai-Azuma, M., Kanai, Y., Kim, I., Lucero, M. T., & Saijoh, Y. (2012). Gut endoderm is involved in the transfer of left-right asymmetry from the node to the lateral plate mesoderm in the mouse embryo. *Development, 139*(13), 2426-2435. doi: 10.1242/dev.079921

Savona, V. C., & Grech, V. (1999). Concepts in cardiology - a historical perspective. *Images Paediatr Cardiol, 1*(1), 22-31.

Schleich, J. M., Abdulla, T., Summers, R., & Houyel, L. (2013). An overview of cardiac morphogenesis. *Arch Cardiovasc Dis, 106*(11), 612-623. doi: 10.1016/j.acvd.2013.07.001

Schoenwolf, G. C., Bleyl, S. B., Brauer, P. R., & Francis-West, P. H. *Larsen's human embryology* (Fifth edition. ed.).

Shiraishi, I., & Ichikawa, H. (2012). Human heterotaxy syndrome - from molecular genetics to clinical features, management, and prognosis. *Circ J, 76*(9), 2066-2075.

Shiratori, H., & Hamada, H. (2006). The left-right axis in the mouse: from origin to morphology. *Development, 133*(11), 2095-2104. doi: 10.1242/dev.02384

Shiratori, H., & Hamada, H. (2014). TGFbeta signaling in establishing left-right asymmetry. *Semin Cell Dev Biol, 32*, 80-84. doi: 10.1016/j.semcdb.2014.03.029

Shiratori, H., Sakuma, R., Watanabe, M., Hashiguchi, H., Mochida, K., Sakai, Y., . . . Hamada, H. (2001). Two-step regulation of left-right asymmetric expression of Pitx2: initiation by nodal signaling and maintenance by Nkx2. *Mol Cell, 7*(1), 137-149.

Srivastava, D. (2001). Genetic assembly of the heart: implications for congenital heart disease. *Annu Rev Physiol, 63*, 451-469. doi: 10.1146/annurev.physiol.63.1.451

Srivastava, D. (2006). Making or breaking the heart: from lineage determination to morphogenesis. *Cell, 126*(6), 1037-1048. doi: 10.1016/j.cell.2006.09.003

Srivastava, D., & Olson, E. N. (2000). A genetic blueprint for cardiac development. *Nature, 407*(6801), 221-226. doi: 10.1038/35025190

Sylva, M., van den Hoff, M. J., & Moorman, A. F. (2014). Development of the human heart. *Am J Med Genet A, 164A*(6), 1347-1371. doi: 10.1002/ajmg.a.35896

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell, 131*(5), 861-872. doi: 10.1016/j.cell.2007.11.019

Takahashi, K., & Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell, 126*(4), 663-676. doi: 10.1016/j.cell.2006.07.024

Takeda, S., Yonekawa, Y., Tanaka, Y., Okada, Y., Nonaka, S., & Hirokawa, N. (1999). Left-right asymmetry and kinesin superfamily protein KIF3A: new insights in determination of laterality and mesoderm induction by kif3A-/- mice analysis. *J Cell Biol, 145*(4), 825-836.

Tanaka, Y., Okada, Y., & Hirokawa, N. (2005). FGF-induced vesicular release of Sonic hedgehog and retinoic acid in leftward nodal flow is critical for left-right determination. *Nature, 435*(7039), 172-177. doi: 10.1038/nature03494

Vandenberg, L. N., Lemire, J. M., & Levin, M. (2013). It's never too early to get it Right: A conserved role for the cytoskeleton in left-right asymmetry. *Commun Integr Biol, 6*(6), e27155. doi: 10.4161/cib.27155

Vandenberg, L. N., & Levin, M. (2012). Polarity proteins are required for left-right axis orientation and twin-twin instruction. *Genesis, 50*(3), 219-234. doi: 10.1002/dvg.20825

Vandenberg, L. N., & Levin, M. (2013). A unified model for left-right asymmetry? Comparison and synthesis of molecular models of embryonic laterality. *Dev Biol, 379*(1), 1-15. doi: 10.1016/j.ydbio.2013.03.021

Viotti, M., Niu, L., Shi, S. H., & Hadjantonakis, A. K. (2012). Role of the gut endoderm in relaying left-right patterning in mice. *PLoS Biol, 10*(3), e1001276. doi: 10.1371/journal.pbio.1001276

Waardenberg, A. J., Ramialison, M., Bouveret, R., & Harvey, R. P. (2014). Genetic networks governing heart development. *Cold Spring Harb Perspect Med, 4*(11), a013839. doi: 10.1101/cshperspect.a013839

Wang, B., Wang, J., Liu, S., Han, X., Xie, X., Tao, Y., . . . Ma, X. (2011). CFC1 mutations in Chinese children with congenital heart disease. *Int J Cardiol, 146*(1), 86-88. doi: 10.1016/j.ijcard.2009.07.034

Wang, B., Yan, J., Mi, R., Zhou, S., Xie, X., Wang, J., & Ma, X. (2010). Forkhead box H1 (FOXH1) sequence variants in ventricular septal defect. *Int J Cardiol, 145*(1), 83-85. doi: 10.1016/j.ijcard.2009.05.030

Wang, B., Yan, J., Peng, Z., Wang, J., Liu, S., Xie, X., & Ma, X. (2011). Teratocarcinoma-derived growth factor 1 (TDGF1) sequence variants in patients with congenital heart defect. *Int J Cardiol, 146*(2), 225-227. doi: 10.1016/j.ijcard.2009.08.046

Wang, J., Xin, Y. F., Xu, W. J., Liu, Z. M., Qiu, X. B., Qu, X. K., . . . Yang, Y. Q. (2013). Prevalence and spectrum of PITX2c mutations associated with congenital heart disease. *DNA Cell Biol, 32*(12), 708-716. doi: 10.1089/dna.2013.2185

Ward, A. J., & Cooper, T. A. (2010). The pathobiology of splicing. *J Pathol, 220*(2), 152-163. doi: 10.1002/path.2649

Wessels, M. W., & Willems, P. J. (2010). Genetic factors in non-syndromic congenital heart malformations. *Clin Genet, 78*(2), 103-123. doi: 10.1111/j.1399-0004.2010.01435.x

Whiley, P. J., Guidugli, L., Walker, L. C., Healey, S., Thompson, B. A., Lakhani, S. R., . . . Spurdle, A. B. (2011). Splicing and multifactorial analysis of intronic BRCA1 and BRCA2 sequence variants identifies clinically significant splicing aberrations up to 12 nucleotides from the intron/exon boundary. *Hum Mutat, 32*(6), 678-687. doi: 10.1002/humu.21495

Wu, S. M., Chien, K. R., & Mummery, C. (2008). Origins and fates of cardiovascular progenitor cells. *Cell, 132*(4), 537-543. doi: 10.1016/j.cell.2008.02.002

Xin, M., Olson, E. N., & Bassel-Duby, R. (2013). Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. *Nat Rev Mol Cell Biol, 14*(8), 529-541. doi: 10.1038/nrm3619

Yoshiba, S., & Hamada, H. (2014). Roles of cilia, fluid flow, and Ca2+ signaling in breaking of left-right symmetry. *Trends Genet, 30*(1), 10-17. doi: 10.1016/j.tig.2013.09.001

Yoshioka, H., Meno, C., Koshiba, K., Sugihara, M., Itoh, H., Ishimaru, Y., . . . Noji, S. (1998). Pitx2, a bicoid-type homeobox gene, is involved in a lefty-signaling pathway in determination of left-right asymmetry. *Cell, 94*(3), 299-305.

Yuan, F., Zhao, L., Wang, J., Zhang, W., Li, X., Qiu, X. B., . . . Yang, Y. Q. (2013). PITX2c loss-of-function mutations responsible for congenital atrial septal defects. *Int J Med Sci, 10*(10), 1422-1429. doi: 10.7150/ijms.6809

Yuan, S., Zhao, L., Brueckner, M., & Sun, Z. (2015). Intraciliary calcium oscillations initiate vertebrate left-right asymmetry. *Curr Biol, 25*(5), 556-567. doi: 10.1016/j.cub.2014.12.051

Zhou, T., Benda, C., Dunzinger, S., Huang, Y., Ho, J. C., Yang, J., . . . Esteban, M. A. (2012). Generation of human induced pluripotent stem cells from urine samples. *Nat Protoc, 7*(12), 2080-2089. doi: 10.1038/nprot.2012.115

Chapter 6

Supplementary material

Supplementary Material 1

Declaração de consentimento informado

Para a realização de estudos genéticos

Eu, _____, portador do Bilhete de Identidade nº _____ do Arquivo de Identificação de _____, autorizo que se faça a colheita de células bucais por esfregaço da parte interna da bochecha (técnica não invasiva nem dolorosa), ao meu/minha filho(a) _____, nascido(a) em __ / ___/ _____, portador(a) de Cardiopatia Congénita Complexa.

O material biológico recolhido destina-se à realização de estudos genéticos sobre a base molecular da Cardiopatia Congénita e Defeitos da Lateralização, no âmbito do projeto de investigação "*Estudo Molecular dos Genes Homólogos Humanos Envolvidos na Definição da Assimetria do Coração* ", em curso no Centro de Biomedicina Molecular e Estrutural, Universidade do Algarve.

Declaro ter sido informado(a) que o material biológico assim recolhido será utilizado apenas para os fins atrás definidos.

Declaro ter sido informado(a) que é assegurada, em qualquer situação, a confidencialidade dos dados pessoais e genéticos resultantes deste (s) estudo (s) e que fica salvaguardada a possibilidade de eu aceder à informação existente sempre que o desejar.

Declaro autorizar / não autorizar* a recolha de imagens e a utilização dos resultados para fins de investigação, no âmbito deste conjunto de patologias, bem como a sua apresentação em reuniões científicas, desde que fique garantida a confidencialidade.

Assinatura: _____; Data: __ / __ / _____

Local da colheita: _____; Rúbrica do técnico: _____

* riscar o que não interessa Identificação da(s) amostra(s): CBME HCF ☐☐☐☐

Supplementary Material 2

Estudo Molecular dos Genes Homólogos Humanos, Envolvidos na Definição da Assimetria do Coração

Identificação da amostra:

nº processo origem: CBME

Data Nascimento: / / ; Sexo M ☐ / F ☐

Instituição de origem:

Médico de Contacto: Telefone:

Amostra biológica - Data da colheita: / /

Familiares:

CBME ; grau de parentesco: _____ Data de Nascimento: / /

Observações:_____

CBME ; grau de parentesco: _____ Data de Nascimento: / /

Observações:_____

Diagnóstico clínico

Caso único: S/N História familiar: S/N

Heterotaxia		Isomerismo Esquerdo/ Direito	
Malformação cardíaca			
Transposição/mal posição dos grandes vasos		Atrésia da artéria pulmonary	
Ventrículo esquerdo de dupla entrada		Defeitos do septo atrioventricular DISAV*	
Ventrículo direito de dupla saída		Coração hipoplástico	
Hipertrofia do ventrículo direito		Outra:	
Malformação gastrointestinal			
Estômago rodado para o lado direito		Fígado na linha média	
Anomalias do tracto biliar		Mal rotação intestinal	
Poli ou asplenia		Outra: _____	

Observações:

Assinatura da Declaração de Consentimento Informado: S/N

Data: / / Rubrica:

Supplementary Material 3

DNA isolation from buccal swabs samples, BIOLINE protocol

1. Place the swab in a 1.5ml microcentrifuge tube. Add 400µl Lysis Buffer D and 25µl Proteinase K. Mix by vortexing vigorously for 5 seconds. Incubate at 50ºC for 10-15 minutes.

2. Remove the swab from the tube and scrape it against the side of the tube to recover as much of the liquid as possible. Discard the swab.

3. Add 200µl Binding Buffer D to the lysed sample and mix by vortexing for 15 seconds.

4. Transfer the sample to Spin Column D placed in a 2ml Collection Tube. Centrifuge at 10,000 x g (12,000rpm) for 2 minutes. Discard the Collection Tube and place the Spin Column in a new Collection Tube.

5. Add 700µl Wash Buffer D. Centrifuge at 10,000 x g (12,000rpm) for 1 minute. Discard the filtrate and reuse the Collection Tube.

6. Repeat step 6.

7. Centrifuge at maximum speed for 2 minutes to remove all traces of ethanol. Discard the Collection Tube.

8. Place Spin Column D in a 1.5ml Elution Tube. Add 200µl Elution Buffer or H2O nuclease free directly to the membrane of the column. Incubate at room temperature for 1 minute. Centrifuge at 6000 x g (8000rpm) for 1 minute to elute the DNA.

Supplementary Material 4

Ethanol precipitation protocol

Ethanol and sodium acetate precipitation protocol

1. Add 1/10 volume of sodium acetate, pH 5.2, (final concentration of 3 M) and mix well.
2. Add 2 to 2.5 volumes of cold 100% ethanol. Mix well.
3. Place on ice or at -20ºC for >20 minutes.
4. Spin at maximum speed and 4ºC in a microfuge for 20 minutes. Carefully decant supernatant.
5. Add 250 µl of cold 70% ethanol. Mix.
6. Spin at maximum speed and 4ºC in a microfuge for 5 minutes. Carefully decant supernatant.
7. Air dry or briefly vacuum dry pellet and resuspend the pellet in the appropriate volume of TE or nuclease-free water.

Supplementary Material 5

INFORMAÇÃO AO DADOR DE AMOSTRAS BIOLÓGICAS

Título do projecto de investigação

" Controlo da proliferação de cardiomiócitos na doença e em medicina regenerativa"

Objetivo do Estudo

Do projeto de investigação, do qual o Professor Doutor José Belo é o responsável, "Estudo Molecular dos Genes Homólogos Humanos Envolvidos na Definição da Assimetria do Coração", protocolado anteriormente entre o CBME (Centro de Biomedicina Molecular e Estrutural, da Universidade do Algarve) e o Hospital de Faro (Centro Hospitalar Algarve, EPE), resultou a identificação de variantes genómicas em doentes. No seguimento deste projeto e com o objetivo da realização de um estudo mais aprofundado do mecanismo da doença, será criada uma linha de células estaminais pluripotentes induzidas (*iPS cells*) a partir de células somáticas isoladas dos doentes. Para tal será necessário a recolha de produto biológico (urina) de cada doente.

O material biológico recolhido destina-se apenas à realização de estudos de investigação biomédica no âmbito da continuação dos estudos "Controlo da proliferação de cardiomiócitos na doença e em medicina regenerativa", agora em curso no Centro de Estudos de Doenças Crónicas, Faculdade de Ciências Médicas, Universidade Nova de Lisboa.

Procedimentos

No caso de concordar continuar a participar neste projeto, ser-lhe-á apenas pedido uma amostra de urina (técnica não invasiva nem dolorosa). Estas colheitas serão efetuadas sem alterar os procedimentos médicos habituais e sem interferir com a consulta dos doentes. A amostra será preservada em condições apropriadas e transportada para o *Laboratório de Células Estaminais e Desenvolvimento*, localizado no CEDOC, de modo a ser processada. A partir de Células epiteliais isoladas da urina, estas serão reprogramadas por métodos convencionais de modo a criar linhas de células estaminais pluripotentes induzidas (*iPS cells*). Estas células poderão então ser utilizadas para derivação de células diferenciadas de várias linhagens somáticas nomeadamente cardíacas de modo a podermos estudar o papel desta mutação na diferenciação e função de células cardíacas. As informações clínicas relacionadas com a amostra serão introduzidas numa base de dados, passando a sua identificação pessoal a estar codificada e só acessível aos investigadores que realizam o estudo. A decisão de doar amostras é totalmente voluntária, pelo que o dador tem o direito de recusar. O doente é livre para interromper a sua colaboração no estudo, em qualquer momento, sem que com isso seja

prejudicado no seu acompanhamento clínico, na instituição. Serão cumpridas todas as normas éticas aceites internacionalmente para o uso de materiais biológicos para fins de investigação. Não se antevê a utilização comercial ou benefícios financeiros que resultem diretamente das células obtidas do dador.

Identificação das amostras e Confidencialidade

Após a colheita, as amostras serão identificadas por um código de forma a preservar a privacidade.

Os dados serão tratados confidencialmente, de acordo com a Lei (LEI 12/2005, ARTIGO Nº 19 DIÁRIO DA REPÚBLICA – I SÉRIE A, Nº 18 DE 26 DE JANEIRO DE 2005), com os regulamentos e de acordo com as normas éticas aprovadas pela Comissão de Ética do Hospital de Faro (Centro Hospitalar Algarve). Os dados resultantes dos estudos realizados serão apenas alvo de publicação científica em revistas da especialidade de forma anónima. Toda a investigação efetuada estará sob a supervisão da Comissão de Ética do Centro Hospital Algarve que assegurará o cumprimento dos regulamentos mencionados.

Possíveis benefícios para os participantes

Esta é uma doação altruísta, não havendo por isso qualquer compensação para o dador. Não se garante que este estudo envolva quaisquer benefícios directos para o participante. No entanto, a sua participação proporcionará a aquisição de conhecimentos que poderão vir a beneficiar terceiros no futuro. Fica salvaguardada a possibilidade dos participantes acederem à informação existente sempre que o desejarem.

Riscos físicos previsíveis

Tendo em conta a natureza do procedimento utilizado para a recolha da amostra biológica, os riscos e o desconforto associados serão mínimos ou inexistentes.

Se tiver qualquer dúvida, em qualquer momento, mesmo após a colheita, sobre este estudo poderá contactar o Responsável pelo estudo:

José António Belo, Faculdade de Ciências Médicas, Universidade Nova de Lisboa.
Edifício CEDOC II, Rua Câmara Pestana n.º 6, 6- A e 6- B. 1150-082 Lisboa. Portugal
Email: jose.belo@fcm.unl.pt Tel: +351 21 8803102 Fax: +351 21 8803010

DECLARAÇÃO DE CONSENTIMENTO INFORMADO

Eu, _____, portador do Bilhete de Identidade nº/CC nº _____ do Arquivo de Identificação de _____, autorizo que se faça a colheita de amostra biológica, urina, ao meu/minha filho(a) _____, nascido(a) em __ / ___/ _____, portador(a) de Cardiopatia Congénita Complexa.

O material biológico recolhido destina-se à realização de estudos de investigação biomédica sobre a base molecular da Cardiopatia Congénita e Defeitos da Lateralização, no âmbito do projeto de investigação "*Controlo da proliferação de cardiomiócitos na doença e em medicina regenerativa* ", em curso no Centro de Estudos de Doenças Crónicas, NOVA Medical School | Faculdade de Ciências Médicas, Universidade Nova de Lisboa.

Declaro ter sido informado(a) que o material biológico assim recolhido será utilizado apenas para os fins atrás definidos.

Declaro ter sido informado(a) que é assegurada, em qualquer situação, a confidencialidade dos dados pessoais e genéticos resultantes deste (s) estudo (s) e que fica salvaguardada a possibilidade de eu aceder à informação existente sempre que o desejar.

Declaro autorizar / não autorizar* a recolha de imagens e a utilização dos resultados para fins de investigação, no âmbito deste conjunto de patologias, bem como a sua apresentação em reuniões científicas, desde que fique garantida a confidencialidade.

Assinatura: _____; Data: __ / __ / _____

Local da colheita: _____;Rúbrica do técnico: _____

* riscar o que não interessa

Identificação da(s) amostra(s): CEDOC HCF ☐☐☐☐

(formato HCF 0000 – numeração sequencial)

Supplementary Material 6

Table S1

Table S1. DAND5 c.455 G>A allele frequency according to ExAC genome project

Gene	Variant	Variant rs number
DAND5	c.455 G>A / p.R152H	45513495
Population	**Allele frequency (%)**	**Number of homozygous**
European (Finnish)	2,4	2
European (Non-Finnish)	1,5	14
All populations	1	16

Supplementary Material 7

Table S2

Table S2. Genotypes of the 38 patients cohort for the variants found in this study.

Proband	*DAND5*	*NODAL*	*CFC1*	*PITX2C*
HStM001	---	C193+12T A494G	---	---
HStM004	---	---	---	A205+71G
HStM006	---	---	---	---
HsTC001	---	C193+12T A494G	G523A	G205+88C
HStC002	---	A494G	---	---
HsTC003	---	A494G	---	---
HsTC004	---	---	---	G205+88C
HsTC005	---	C193+12T A494G	---	A205+71G
HsTC006	---	T193+12T G494G	---	G205+71G
HsTC007	A455G	C193+12T G494G	---	G205+88C
HsTC008	---	C193+12T A494G	---	A205+71G
HsTC009	---	A494G	---	C205+88C
HsTC010	---	A494G	---	---
HsTC011	---	A494G	---	---
HCF001	---	C193+12T A494G	---	G205+88C
HCF002	---	C193+12T G494G	---	G205+88C
HCF003	---	C193+12T A494G	---	A205+71G
HCF004	---	A494G	---	C205+88C
HCF005	---	A494G	---	---
HCF006	---	---	---	A205+71G
HCF007	---	C193+12T A494G	---	G205+71G

HCF008	---	A494G	---	---
HCF009	---	A494G	---	---
HCF010	---	G494G	---	C205+88C
HCF011	---	T193+12T G494G	---	C205+88C
HCF012	---	C193+12T A494G	---	G205+88C
HCF013	---	C193+12T A494G	---	G205+88C
HCF014	---	A494G	---	---
HCF015	---	A494G	---	---
HCF017	---	C193+12T A494G	---	G205+88C
HCF019	---	T193+12T G494G	---	G205+88C
HCF021	---	C193+12T A494G	---	A205+71G G205+88C
HCF023	---	C193+12T A494G	---	---
HCF025	A455G	C193+12T A494G	---	---
HCF027	---	C193+12T A494G	---	G205+88C
HCF029	---	C193+12T A494G	---	G205+88C
HCF030	---	G494G	---	
HCF032	---	---	---	G205+88C