

Thèse de Doctorat

Marta SÁNCHEZ CASTRO

*Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université de Nantes
sous le label de L'Université Nantes Angers Le Mans*

École doctorale : Biologie-santé Nantes (ED 502)

Discipline : Sciences de la vie et de la santé

Spécialité : Génétique moléculaire

Unité de recherche : Institut du thorax, Inserm UMR 1087 /CNRS UMR6291, IRS

Soutenue le 2 octobre 2014

Génétique des malformations cardiaques congénitales

Apport de la technique de puces à ADN génomique
(Comparative Genomic Hybridisation ; aCGH)

JURY

Rapporteurs :	M. Patrice BOUVAGNET , Praticien Hospitalier, Maître de Conférences des Universités, Lyon, France M. Koenraad DEVRIENDT , Professeur des Universités, Leuven, Belgique
Examineurs :	M. Damien BONNET , Praticien Hospitalier, Professeur des Universités, Paris, France Mme. Véronique GOURNAY , Praticien Hospitalier, Professeur des Universités, Nantes, France
Directeur de Thèse :	M.Cédric LE CAIGNEC , Praticien Hospitalier, Professeur des Universités, CHU Nantes, France

*"They were maps that lived, maps that one could study, frown over, and add to; maps, in short,
that really meant something."*

— Gerald Durrell, *My Family and Other Animals*

Table des matières

Remerciements	6
Liste des publications	10
Liste des illustrations	11
Liste des abréviations	13
AVANT PROPOS	17
RÉSUMÉ.....	20
INTRODUCTION	21
I. Épidémiologie des malformations cardiaques congénitales.....	21
II. Étiologie des malformations cardiaques congénitales	27
II.1 Facteurs environnementaux et tératogènes	27
II.2 Anomalies chromosomiques.....	28
II.3 Formes monogéniques	30
➤ II.3.1 Formes syndromiques.....	30
➤ II.3.2 Formes non syndromiques.....	31
II.4 Formes multifactorielles	33
II.5 Risque de récurrence des MCC	33
III. Éléments d'embryologie cardiaque	35
III.1 Formation du tube cardiaque primitif.....	35
III.2 Formation des cavités cardiaques.....	39
➤ III.2.1 La boucle cardiaque (looping)	39
➤ III.2.2 La convergence	40
➤ III.2.3 Le wedging.....	41
III.3 Développement du conotruncus	42

➤ III.3.1 Rôle des cellules de la crête neurale cardiaque.....	43
➤ III.3.2 Septation du conotruncus	46
IV. Les malformations cardiaques congénitales de la voie efférente	46
IV.1 Incidence des malformations cardiaques congénitales de la voie efférente.....	48
IV.2 Coarctation de l'aorte	49
IV.3 Transposition des gros vaisseaux	51
IV.4 Tétralogie de Fallot	52
V. Génétique du développement cardiaque	54
V.1 Stades précoces de la différenciation cardiaque	54
V.2 Fusion des tubes cardiaques primitifs.....	55
V.3 Formation de la boucle cardiaque.....	56
V.4 Formation des cavités cardiaques	56
VI. Génétique des malformations cardiaques congénitales non syndromiques.....	60
VII. Le rôle de la variabilité du nombre de copies (CNVs) d'ADN génomique.....	61
VIII. Stratégies d'identification des CNVs.....	63
VIII.1 Hybridation In Situ en Fluorescence (FISH).....	63
VIII.2 Hybridation Génomique Comparative (CGH) array	64
OBJECTIFS.....	68
PROJET 1.....	69
PROJET 2.....	115
PROJET 3.....	138
DISCUSSION GÉNÉRALE ET CONCLUSION	161
ANNEXES.....	163
I. Annexe 1. Tableau complet de correspondance entre le code IPCCC et la CIM-10.....	163
II. Annexe 2. Facteurs de risque non génétiques et MCC (Patel and Burns 2013).....	168
II. Annexe 2. Table 1	168
II. Annexe 2. Table 2	170

II. Annexe 2. Table 3	172
II. Annexe 2. Table 4	172
II. Annexe 2. Table 5	173
RÉFÉRENCES BIBLIOGRAPHIQUES	178

Remerciements

J'adresse mes sincères remerciements tout d'abord au professeur Hervé Le Marec, directeur de l'unité, et au docteur Jean-Jacques Schott, directeur de l'équipe de génétique, pour m'avoir accueillie au sein de son groupe de recherche et pour son soutien tout au long de ma thèse.

Je voudrais aussi remercier les membres du jury : le docteur Patrice Bouvagnet, le professeur Koenraad Devriendt, le professeur Damien Bonnet et le professeur Véronique Gournay, d'avoir accepté de juger mon travail et d'avoir dédié de leur temps précieux à la lecture et à l'examen de ma thèse.

Je remercie les membres de mon comité de thèse, le docteur Richard Redon et le docteur Daniel Guerrier, d'avoir examiné mon travail chaque année, merci pour nos discussions, pour vos réflexions et pour les idées apportées qui m'ont aidée à améliorer et réorienter ma recherche quand cela a été nécessaire. Merci Richard pour ton accueil, ton soutien et ta disponibilité pendant ces années. Merci pour ton aide et pour tes conseils personnels et génétiques, et merci de me considérer comme un membre à part entière de ton équipe.

Je tiens à témoigner ma reconnaissance et ma gratitude au professeur Cédric Le Caignec pour m'avoir accueillie dans son équipe. Merci Cédric de m'avoir confié tes projets de recherche et pour ton soutien pendant mes années de thèse. Merci pour ta gentillesse, pour ta rigueur scientifique, pour ta grande patience surtout pendant l'écriture de la thèse et pour tes conseils. Merci encore pour nos discussions qui ont contribué à améliorer mes connaissances en génétique, et le plus important, à m'amuser avec la recherche.

Un grand merci à Hadja Eldjouzi qui était en charge de ce projet de recherche tout au début et qui a réalisé les tâches les plus essentielles et dures sans lesquelles ce projet n'aurait pas pu arriver au point où il en est aujourd'hui. Merci pour tes rires, tes gâteaux, ta complicité, ta patience, tes efforts pour essayer de comprendre mon mauvais français surtout au début de ma thèse et merci de m'avoir transmis tes connaissances de laboratoire. C'est toi qui m'as appris les bases des techniques de génétique

moléculaire. Merci aussi à Quentin Hauet de m'avoir aidée tout au début de ma thèse à acquérir une vision plus physiopathologique des maladies traitées dans ce travail, de ta gentillesse, ta bonne humeur et de ta disponibilité. Et merci aussi à Imen Fellah pour ton efficacité dans les tâches de recrutement des patients pour cette étude et de ta gentillesse.

Je remercie spécialement Olivier Pichon, que je considère un peu comme mon mentor dans le laboratoire. Merci de m'avoir appris les techniques de génétique moléculaire dont j'ai eu besoin, merci pour tes précieux conseils, pour nos discussions et pour ta grande culture en génétique qui a apporté à ce travail une grande expertise moléculaire. Merci pour ta complicité, ta grande disponibilité et ta patience inépuisable. Grand merci aussi à Annaïg et Damien de votre disponibilité, de votre aide, de votre sympathie et du cours de crochet qui m'a bien aidé à me détendre en fin de thèse. Merci aussi à M. Piloquet pour votre gentillesse et pour vos conseils sur la technique de Fiber FISH. Merci aussi à Sophie pour ta gentillesse et pour tous ces bons moments partagés à Bordeaux. Et grand merci à Kamran pour ta positivité, ta gentillesse et pour nos discussions dans le domaine professionnel et personnel qui m'ont aidée à clarifier mes idées vers la fin de ma thèse. Merci aussi de me rappeler ce qui est le plus important dans la vie.

J'aimerais remercier Simon, Estelle et Stéphanie Bonnaud de m'avoir aidée tout au long de ma thèse à réussir les différentes techniques moléculaires que j'ai eues à réaliser au laboratoire et de m'avoir donné plein de bonnes idées pour améliorer mes résultats. Merci de votre gentillesse et de votre efficacité. Merci Simon de nous faire plaisir avec ton délicieux gâteau fondant au chocolat plus que fondant.

J'adresse mes sincères remerciements à tous les bioinformaticiens : Edouard, Raluca, Audrey, Pierre, Solena et Eric pour votre patience, votre bonne humeur et de votre bonne volonté pour m'avoir aidé à faciliter l'analyse des données issues de mon travail. Un merci spécial à Eric Charpentier pour avoir réalisé toutes les analyses des puces à ADN et pour les améliorations apportées. Merci aussi pour ta disponibilité inépuisable et pour ta bonne humeur en tout moment.

Grand merci aussi au groupe essentiel des biostatisticiens : Christian, Mathilda, Floriane et Pierre-François. Merci de m'avoir aidée avec les questions statistiques, toujours complexes, dérivées de mon travail. Merci d'avoir essayé de trouver des solutions à mes questions, merci de votre gentillesse, de votre bonne humeur et de votre disponibilité.

Je remercie toutes les composantes de la plateforme de séquençage : la douceur de Françoise, le sourire permanent de Jade (et précédemment d'Aurore), l'expertise de Laetitia et l'amabilité d'Audrey, Marine, Catherine et Béatrice. Merci de m'avoir appris la technique de séquençage, de m'avoir fait confiance pour utiliser tous les appareils quand j'en avais besoin et de m'avoir aidée dès que je l'ai demandé.

Je ne peux pas oublier de remercier tous les thésards de l'UMR 1087 qui ont été mes compagnons de route pendant ces trois années de thèse. Je voudrais remercier spécialement Vincent, Xavier, Antoine, Jean-Baptiste et Romain avec lesquels j'ai aussi partagé le bureau. Merci pour nos discussions, pour votre patience, pour votre sens de l'humour et pour votre aide. Merci aussi d'avoir essayé d'améliorer mon français en me corrigeant à chaque fois que je commettais une erreur de langage, c'est-à-dire toujours. Et surtout merci pour « l'heure de ball », tous les jours vers 18h, absolument nécessaire pour se détendre et évacuer tout le stress accumulé pendant la journée.

Un grand merci à l'équipe de la « biocol » composée de Thierry et des sœurs Martine et Marie-France. Merci pour votre organisation, merci de veiller sur les échantillons, sur la conformité des consentements et sur le laboratoire en général. Merci aussi de votre sympathie, de votre bonne humeur et des histoires de vos voyages de rêve !

Merci aussi à Aurélie et Stéphanie Châtel pour votre gentillesse et de votre aide quand j'en ai eu besoin. Merci au groupe des secrétaires : Isabelle, Ophélie, Corinne, Marie-Pierre et Aurélie pour votre efficacité et votre aide pour simplifier toutes les questions administratives, toujours avec le sourire. Et merci à Vimla pour ta gentillesse, ta motivation et ta capacité d'organisation.

D'une manière générale je voudrais remercier tous les membres de l'UMR 1087 qui m'ont réservé un excellent accueil et qui d'une façon ou d'une autre ont contribué à faire de ces trois années de thèse un excellent souvenir. Je peux dire que je me suis sentie très à l'aise parmi vous.

Je remercie aussi de manière générale les équipes de cytogénétique, génétique moléculaire et de diagnostic préimplantatoire du CHU de Nantes. Merci pour votre accueil chaleureux, votre patience et votre aide à tout instant. C'était un luxe et un plaisir d'avoir partagé ces trois années de thèse avec vous.

J'adresse aussi un grand merci à tous mes amis qui n'ont pas cessé de me donner des encouragements pour arriver à la fin de ma thèse.

Je dois remercier mes parents (Blas et Lola), ma sœur (Ma Dolores), mon beau-frère (David) et mes deux neveux (Pau et Oriol) pour leur patience et leurs encouragements inconditionnels. Sans votre soutien je n'aurais pas été capable de finir cette thèse. Merci maman pour tous les déplacements à Nantes afin de garder Laia, qui m'ont permis de souffler de temps en temps.

Merci aussi à mes beaux-parents (Enric et Nuria) et ma belle-sœur (Neus) pour vos encouragements.

Finalement je voudrais remercier particulièrement ma petite famille qui a enduré la pire partie de cette thèse. Merci à mon compagnon Jordi pour tes sacrifices, ton encouragement, ta patience devant mes frustrations et pour ta détermination quand j'en ai manqué. Tu n'arriveras jamais à réaliser l'importance de ton soutien. Merci aussi à ma fille aînée Laia d'être « apparue » au milieu de ma thèse. Tes sourires et ton amour m'ont donné des forces pour continuer. Merci aussi à ma petite fille Lena pour m'avoir accompagnée dans les derniers mois de thèse. Tes petits coups de pied m'ont aussi encouragée à chaque instant.

Liste des publications

Rare Copy Number Variants in patients with Tetralogy of Fallot, Transposition of the Great Arteries, and Coarctation of the aorta

Marta Sanchez-Castro^{1,2}, et al.

Version préliminaire

Congenital heart defects in patients with deletions upstream of SOX9

Marta Sanchez-Castro^{1,2} Christopher T. Gordon,^{3,4} Florence Petit,⁵ Alex S. Nord,^{6,7} Patrick Callier,⁸ Joris Andrieux,⁵ Patrice Guérin,^{1,2} Olivier Pichon,⁹ Albert David,⁹ Véronique Abadie,¹⁰ Damien Bonnet,¹¹ Axel Visel,^{6,7} Len A. Pennacchio,^{6,7} Jeanne Amiel,^{3,4,12} Stanislas Lyonnet,^{3,4,12} Cédric Le Caignec^{1,2,9*}

Article accepté et publié dans le journal : Human mutation

Disruption of the SEMA3D gene in a patient with congenital heart defects

Marta Sanchez-Castro^{1,2}, Olivier Pichon³, Annaig Briand³, Damien Poulain³, Véronique Gournay⁴, Albert David³, Cedric Le Caignec^{1,2,3*}

Soumis dans le journal : Human Mutation

Liste des illustrations

I. Figure 1- Prévalence totale des MCC au cours du temps.	22
II. Figure 2. Prévalence totale de MCC par continent à la naissance.	23
III. Figure 3. Prévalence à la naissance des sous-types de MCC au cours du temps.	24
IV. Figure 4. Variabilité de l'incidence par million des différentes formes de MCC à la naissance.	26
V. Figure 5. Rapport de vraisemblance, en fonction du nombre de loci, appliqué aux tétralogies de Fallot.	35
VI. Figure 6. Vue dorsale et section transversale d'un embryon au stade présomitique.	36
VII. Figure 7. Transformation des amas cellulaires angioformateurs en deux tubes cardiaques ou endocardiques puis fusion sur la ligne médiane pour former le tube cardiaque primitif.	37
VIII. Figure 8. Tube cardiaque primitif rectiligne.	38
IX. Figure 9. Formation normale de la boucle cardiaque.	39
X. Figure 10. Les trois principales étapes du développement cardiaque : <i>early looping, convergence, wedging et cœur normal</i>	40
XI. Figure 11. Représentation schématique de la migration des cellules de la crête neurale cardiaque des rhombomères 6-8 à travers les arcs pharyngés 3, 4 et 6.44	
XII. Figure 12. Représentation schématique du cœur normal et des principales cardiopathies conotruncales.	48
XIII. Figure 13. Représentation schématique de la coarctation de l'aorte post- ductale.	50

XIV. Figure 14. Représentation schématique de la transposition des gros vaisseaux.	52
XV. Figure 15. Représentation schématique de la tétralogie de Fallot.	53
XVI. Figure 16. Représentation schématique des principaux stades du développement du coeur.	57
XVII. Figure 17. Principe de la technique de la CGH array.	66
XVIII. Figure 18. Le résultat de la CGH array.	67
XXI. Table 1. Incidence des MCC par million.	24
XXII. Table 2. Anomalies chromosomiques associées aux MCC.	29
XXIII. Table 3. Microdeletions et syndromes monogéniques associés à une MCC.	31
XXIV. Table 4. Gènes sélectionnés associés à des formes non syndromiques de MCC.	32
XXV. Table 5. Risque de récurrence des formes non syndromiques des MCC.	34
XXVI. Table 6. Mutations / délétions dans des facteurs de transcription associés au développement cardiaque.	58
XXVII. Table 7. Gènes impliqués dans des formes non-syndromiques de malformations cardiaques congénitales.	61

Liste des abréviations

A : oreillette

aCGH : array comparative genomic hybridisation (Hybridation Génomique Comparative array)

Ao : aorte

ARNm : ARN messenger

AS ou AoS : sténose aortique

ASD: communication inter-auriculaire

AVSD: canal atrio-ventriculaire

BAC : bacterial artificial chromosome (chromosome artificiel bactérien)

BAV : valve aortique bicuspide

BMPs : Protéine morphogénétique osseuse

cDNA: complementary DNA

CGH : comparative genomic hybridization (Hybridation Génomique Comparative)

CHD: congenital heart defects

ChIP: chromatin immunoprecipitation

CHM : Congenital heart malformation

CNV: copy number variant ou variation en nombre de copies

CoA ou Coarc: coarctation de l'aorte

d-TGA: dextro-transposition des gros vaisseaux

DDD : deciphering developmental disorders

DORV : ventricule droit à double issue

Ebstein's : anomalie d'Ebstein

FGF: facteurs de croissance des fibroblastes

FISH : hybridation in situ en fluorescence

CRGh37 : Genome Reference Consortium Human genome build 37

HCM : cardiomyopathie hypertrophique

HLH ou HLHS : hypoplasie du cœur gauche

HRH : hypoplasie du cœur droit

HRP : horseradish peroxidase

IAA : interruption de l'arc aortique

IAA(B) : interruption de l'arche aortique (type B)

ISCA : international Standards for Cytogenomic Arrays

L-TGA : levo-transposition des gros vaisseaux

LA : oreillette gauche

LR-PCR : long range PCR

LV : ventricule gauche

LVOTO : obstruction de la voie d'éjection ventriculaire gauche

M-MLV : Moloney Murine Leukemia Virus

MCC : malformations cardiaques congénitales

miARN : micro-ARNs

MVP : prolapsus valvulaire mitral

NCBI36 : National Center for Biotechnology Information build 36

NERICP : New England Regional Infant Cardiac Program

NMD : non-mediated decay

NRP : neuropilin

OFT : outflow tract

OT : voie afférente

PA : artère pulmonaire

PAPVR : retour veineux pulmonaire anormal partiel

PCR : polymerase chain reaction (réaction en chaîne par polymérase)

PDA : canal artériel patent

PPS : sténose pulmonaire périphérique

PS : sténose pulmonaire

PTA : persistent truncus arteriosus

Pul atresia IVS : atrésie pulmonaire à septum ventriculaire intact

PVS : sténose pulmonaire valvulaire

qPCR : quantitative PCR (PCR quantitative)

RA : oreillette droite

RT-PCR : reverse transcription polymerase chain reaction

RV : ventricule droit

Sema : semaphorin

Seq : séquençage

SHF : second champ cardiaque

SRF : serum response Factor (facteur de réponse sérique)

SV : ventricule unique

SV : sinus venosus

SVAS : sténose aortique supravalvulaire

TA : truncus arteriosus

TA : atrésie tricuspide

TAPVC : connection veineuse pulmonaire anormale totale

TAPVR : retour veineux pulmonaire anormal total

TFBS : transcription factor binding sites

TGA : transposition des gros vaisseaux

ToF : tétralogie of Fallot

3'UTR : three prime untranslated region

V : ventricule

VSD : communication inter-ventriculaire

WTCCC : Wellcome Trust Case Control Consortium

YAC : yeast artificial chromosome (chromosome artificiel de levure)

AVANT PROPOS

Le terme « maladies cardiovasculaires » englobe des anomalies très variées du cœur et des vaisseaux. Parmi celles-ci, les malformations cardiaques congénitales (MCC) sont les malformations les plus fréquentes et la principale cause de mortalité chez les enfants nés vivants. Plusieurs millions d'enfants dans le monde en sont affectés, avec une sévérité variable. Beaucoup reste à faire, mais l'amélioration de la prise en charge médicale et chirurgicale a permis une réduction significative de la mortalité et de la morbidité d'une majorité d'enfants, y compris pour les malformations les plus sévères.

Le déterminisme de ces malformations est complexe. Des facteurs environnementaux associés à une prédisposition génétique expliquent probablement la plupart des formes observées. Les facteurs génétiques étaient inconnus, il y a encore peu, mais au cours de la dernière décennie, des avancées majeures ont permis une meilleure compréhension des causes génétiques des MCC chez l'Homme et l'identification d'anomalies génétiques spécifiques de certains types de malformations. Ces nouveaux apports de la recherche sont importants à plusieurs niveaux. Au plan fondamental, ils représentent la première étape avant des études physiopathologiques qui permettront une meilleure compréhension du développement cardiaque normal et pathologique conduisant aux principales anomalies. Sur le plan clinique, ils sont essentiels aux soignants. Ils permettent de donner un meilleur conseil génétique aux parents d'enfants malades et, de plus en plus, aux patients eux-mêmes qui, opérés, souhaitent une descendance. On estime en effet qu'actuellement 85 % - 90 % de ces enfants atteignent l'âge adulte (Gurvitz et al. 2007).

Nous avons choisi de nous intéresser à trois types de malformations cardiaques congénitales sévères : la tétralogie de Fallot, la coarctation de l'aorte et la transposition des gros vaisseaux. Dans la première partie de cette thèse, nous donnerons quelques éléments d'épidémiologie des malformations cardiaques congénitales puis de développement cardiaque sur le plan embryologique puis génétique. Le contrôle génétique du développement cardiaque est d'une extrême complexité. À titre d'exemple, plus de 150 modèles murins avec mutations induites présentent un développement

cardiaque anormal. Ce chapitre ne prétend donc pas être exhaustif, mais centré sur les étapes clés à l'origine des cardiopathies affectant la région du conotruncus. Il est possible de classer ces malformations de nombreuses façons. Nous avons choisi de distinguer les malformations cardiaques congénitales isolées des formes dites syndromiques. Une forme syndromique est définie par l'association d'une malformation cardiaque à un ou plusieurs autres signes, entrant alors dans le cadre d'un syndrome.

Dans une seconde partie, nous détaillerons et discuterons les résultats obtenus dans le cadre de ce travail réalisé au sein de l'unité de génétique UMR 1087 à l'Institut du thorax de Nantes et du service de Génétique Médicale du centre hospitalo-universitaire de Nantes. Ce travail a consisté à recruter, grâce à un Projet Hospitalier de Recherche Clinique, des patients atteints de forme sporadique de malformations cardiaques congénitales afin de contribuer à l'identification des bases génétiques de ces malformations. Notre hypothèse centrale est que le risque de malformation cardiaque congénitale est modulé par des variations génétiques dérégulant certaines voies essentielles du développement embryonnaire du cœur. Afin de détecter certaines variations génétiques, nous avons mené une étude utilisant la technique de puces à ADN ou CGH array (*Comparative Genomic Hybridization array*) sur une série de patients présentant une tétralogie de Fallot (ToF), une coarctation de l'aorte (CoA) et/ou une transposition des gros vaisseaux (TGA).

Il est communément admis par les cliniciens et les généticiens que l'expressivité de ces pathologies est très variable, allant d'un phénotype normal à des formes cliniques extrêmement sévères. Par conséquent, des variations altérant la fonction d'un gène majeur sont parfois observées dans certaines familles, mais pas toujours suffisantes pour développer une maladie et d'autres facteurs restent à identifier. Ces facteurs pouvant être d'origine endogène (gène à effet mineur) ou exogène (environnement, mode de vie, alimentation, etc.). L'identification de ces modulateurs implique une évolution conceptuelle des modèles génétiques, une rigueur extrême de phénotypage afin de caractériser des variations d'expressivité et un développement de nouvelles technologies pouvant répondre à ce nouveau défi.

Mon travail de thèse s'inscrit dans ce contexte de recherche appliquée mêlant données cliniques, recherche de l'origine moléculaire à l'aide de technologies et de concepts en constante évolution et compréhension physiopathologique de ces variations

génétiques. L'objectif final est d'utiliser ces nouvelles connaissances pour améliorer la prise en charge des patients.

RÉSUMÉ

Génétique des malformations cardiaques congénitales : apport de la technique de puces à ADN génomique (Comparative Genomic Hybridisation ; aCGH)

Notre projet a consisté à analyser par aCGH 316 patients présentant une malformation cardiaque congénitale (MCC) à type de transposition des gros vaisseaux, tétralogie de Fallot ou coarctation de l'aorte avec pour objectif la détection de microdélétions et microduplications génomiques et ainsi l'identification de nouveaux gènes contribuant à ces MCC. L'identification de nouveaux gènes permet une meilleure compréhension des mécanismes moléculaires et embryologiques à l'origine de ces malformations et un meilleur conseil génétique.

Notre étude a conduit à l'identification de 21 microdélétions et 50 microduplications rares *de novo* ou héritées dont certaines comprennent des gènes candidats aux MCC. L'analyse bioinformatique des données a permis de montrer que de nombreuses microdélétions/microduplications comprennent des gènes possédant des sites de fixation de FOXC1. Ces microdélétions/microduplications pourraient donc altérer l'expression de certains gènes régulés par FOXC1 au cours du développement cardiaque. Par ailleurs, parmi les remaniements identifiés, nous avons détecté une délétion d'environ 1Mb en amont de SOX9 chez des patients avec malformation cardiaque et syndrome de Pierre Robin. Nous avons montré que cette délétion comprend plusieurs activateurs cardiaques putatifs. Ces résultats suggèrent que la dérégulation de SOX9 peut être responsable de MCC. Enfin, nous avons identifié une duplication partielle de la région 5' du gène SEMA3D donnant lieu à un ARNm tronqué de SEMA3D. Ces résultats suggèrent que la portion tronquée de SEMA3D pourrait perturber la migration des cellules de la crête neurale au cours du développement et ainsi contribuer aux MCC.

Mots-clés : malformations cardiaques congénitales, transposition des gros vaisseaux, coarctation de l'aorte, tétralogie de Fallot, CGH array, CNV, FOXC1, SOX9, SEMA3D.

INTRODUCTION

I. Épidémiologie des malformations cardiaques congénitales

Les malformations cardiaques congénitales (MCC) représentent un problème majeur de santé publique : il s'agit de la première cause de mortalité au cours de la première année de vie dans les pays développés (Gillum 1994; Petrini et al. 1997). La prévalence des cardiopathies congénitales est estimée à 0.5-1% des naissances vivantes (Hoffman and Kaplan 2002; Øyen et al. 2009). En France, cela correspond à 5000 nouveaux cas par an (Bajolle and Bonnet 2009). Les enfants atteints occupent entre 25 et 30% des lits dans la plupart des unités de soins intensifs pédiatriques.

Les progrès médicaux et chirurgicaux réalisés chez les enfants atteints de MCC ont contribué à augmenter significativement la prévalence des MCC. Qui plus est, les MCC sont ensuite associées à une forte morbidité à long terme chez les patients (De Galan-Roosen et al. 1998; Botto and Correa 2003). La moyenne d'âge des patients avec une MCC sévère est passée de 11 ans en 1985 à 17 ans en 2000 et la moyenne d'âge globale au moment du décès est également passée de 37 ans en 2002 à 57 ans en 2007 (Morris and Menashe 1991; Boneva et al. 2001). Dans les deux dernières décennies, la mortalité des patients atteints d'une MCC a diminué de façon variable selon le type de MCC. Par exemple, elle a diminué de 50% pour la tétralogie de Fallot et la communication inter-auriculaire ; de 55% pour le patent ductus arteriosus ; de 60% pour la sténose aortique congénitale ; et de 70% pour la transposition des gros vaisseaux et la coarctation de l'aorte (Pillutla et al. 2009). Parallèlement, on estime que la population adulte avec MCC augmente de ~5% par an (van der Bom et al. 2011). Par ailleurs, le risque de malformation cardiaque pour la descendance serait dix fois supérieur à la population générale (Burn et al. 1998).

Prévalence des malformations cardiaques congénitales

Les études de prévalence des MCC réalisées depuis 1930 montrent une augmentation de l'incidence des MCC au cours du temps (Figure 1) (van der Linde et al. 2011). Cette augmentation au fil du temps est représentée par une courbe en S : avec une première pente d'augmentation de 1930 à 1960, suivie par une stabilisation autour de 5.3 pour mille naissances vivantes de 1961 à 1975, une deuxième pente d'augmentation de la fin des années 70s à 1995 et finalement, une stabilisation autour de 9.1 pour mille naissances vivantes au cours des 15 dernières années.

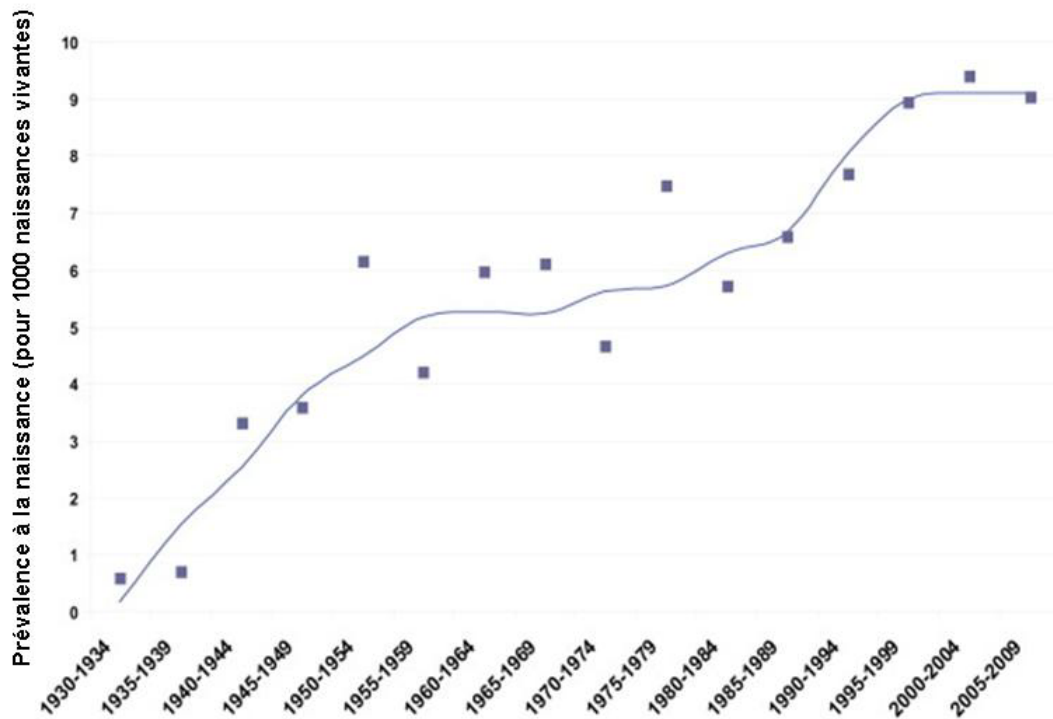


Figure 1- Prévalence totale des MCC au cours du temps. Prévalence à la naissance des MCC rapportée au cours du temps de 1930 à 2010. La ligne montre la tendance temporelle et les carrés représentent les valeurs de la prévalence de MCC calculées à la naissance pour chaque période de temps. Modifié d'après (van der Linde et al. 2011).

Des différences géographiques significatives ont également été observées entre les différents continents (Figure 2). La prévalence totale la plus élevée de MCC à la

naissance a été trouvée en Asie (9,3 pour mille naissances vivantes) et la plus faible en Afrique (1,9 pour mille naissances vivantes). La prévalence totale rapportée de MCC à la naissance en Asie est significativement plus élevée comparée à tous les autres continents ($p < 0,001$). L'Europe a la deuxième plus haute prévalence totale de MCC à la naissance (8,2 pour mille naissances vivantes). La plus faible incidence des MCC observée en Afrique comprenant des pays à faible revenu comparée à celle observée en Europe comprenant des pays aux revenus élevés pourrait être expliquée par un sous diagnostic dans les pays africains plutôt que par une plus faible prévalence réelle de MCC à la naissance.

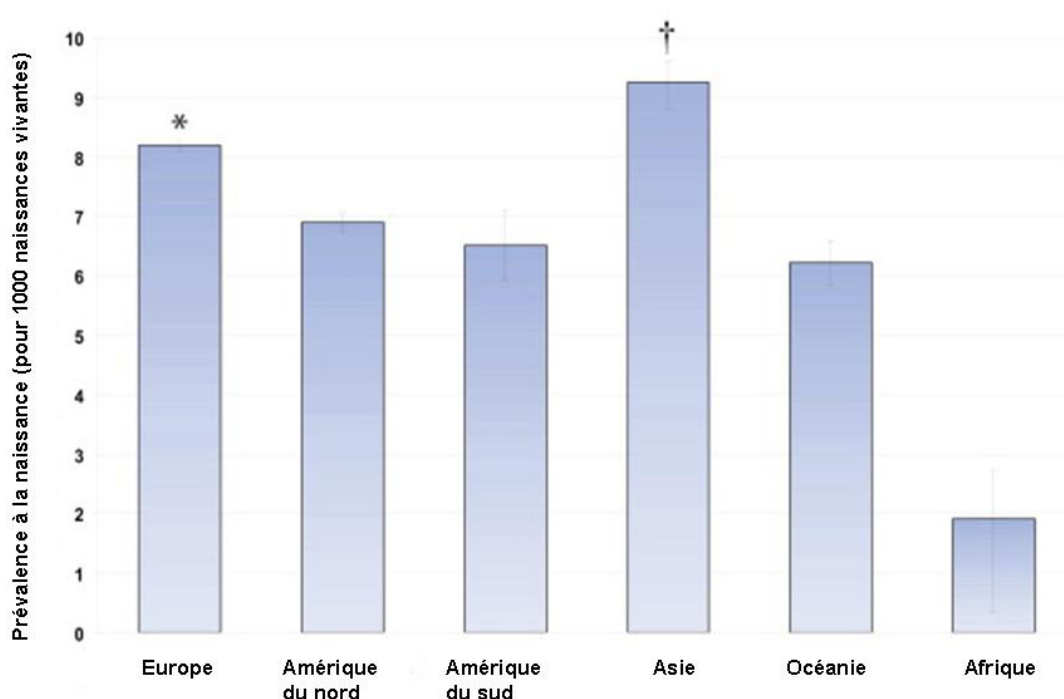


Figure 2. Prévalence totale de MCC par continent à la naissance. *La prévalence totale de MCC à la naissance en Europe est significativement plus élevée qu'en Amérique du nord ($p < 0,001$), Amérique du sud ($p < 0,001$), Océanie ($p < 0,001$) et Afrique ($p < 0,001$). †La prévalence totale de MCC à la naissance en Asie est significativement plus élevée qu'en Europe ($p < 0,001$), Amérique du nord ($p < 0,001$), Amérique du sud ($p < 0,001$), Océanie ($p < 0,001$) et Afrique ($p < 0,001$). Modifié d'après (van der Linde et al. 2011).

Par rapport à la prévalence à la naissance des 8 sous-types de MCC, les plus communes (Figure 3) (sténose aortique (AoS); communication inter-auriculaire (ASD); coarctation de l'aorte (CoA); patent ductus arteriosus (PDA); sténose pulmonaire (PS); transposition des gros vaisseaux (TGA); tétralogie de Fallot (TOF); communication inter-

ventriculaire membraneuse (VSD)) depuis 1945, on observe qu'au fil du temps il y a eu une légère augmentation des VSD, ASD et PDA. En revanche, la prévalence des PS, TOF, CoA, TGA et AoS est restée plus ou moins stable au cours de cette période. Ainsi, au sein du large spectre de malformations, la prévalence de chaque forme est très variable. La communication inter-ventriculaire membraneuse (VSD) est l'anomalie la plus courante avec une incidence à la naissance d'environ 3.5 pour mille naissances vivantes (Table 1) (Figure 4) (Hoffman and Kaplan 2002).

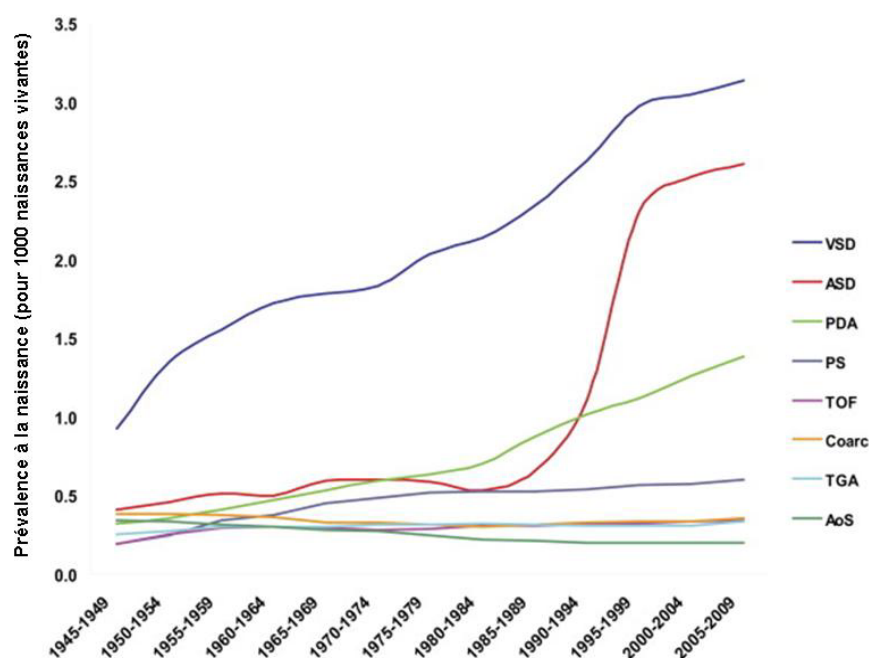


Figure 3. Prévalence à la naissance des sous-types de MCC au cours du temps. Prévalence à la naissance des 8 sous-types de MCC les plus fréquentes de 1945 à 2010. AoS = sténose aortique; ASD = communication inter-auriculaire; Coarc = coarctation de l'aorte; PDA = patent ductus arteriosus; PS = sténose pulmonaire; TGA = transposition des gros vaisseaux; TOF = tétralogie de Fallot; VSD = communication inter-ventriculaire membraneuse. Modifié d'après (van der Linde et al. 2011).

Table 1. Incidence des MCC par million.

Lesion	Number of Studies	Mean	SD	Lower Quartile	Median	Upper Quartile	NERICP 1975-1977
VSD	43	3,570	2,878	1,757	2,829	4,482	345
PDA	40	799	1,399	324	567	782	135
ASD	43	941	1,043	372	564	1,059	65
AVSD	40	348	165	242	340	396	110
PS	39	729	731	355	532	836	73
AS	37	401	543	161	256	388	41
Coarc	39	409	246	289	356	492	165
Tetralogy	41	421	188	291	356	577	196
d-TGA	41	315	115	231	303	388	218
HRH	32	222	199	105	160	224	—
Tricuspid atresia	11	79	52	24	92	118	56
Ebstein's anomaly	5	114	138	38	40	161	12
Pul Atresia	11	132	123	76	83	147	69
HLH	36	266	216	154	226	279	163
Truncus	30	107	71	61	94	136	30
DORV	16	157	103	82	127	245	32
SV	23	106	70	54	85	136	54
TAPVC	25	94	46	60	91	120	58
All cyanotic	37	1,391	590	1,078	1,270	1,533	888
All CHD*	43	9,596	7,484	6,020	7,669	10,567	2,033
BAV	10	13,556	13,049	5,336	9,244	13,817	—

BAV = valve aortique bicuspidée; MCC = malformation cardiaque congénitale; Coarc = coarctation de l'aorte; NERICP = New England Regional Infant Cardiac Program. Les autres abréviations sont définies dans la légende de la Figure 4. D'après (Hoffman and Kaplan 2002).

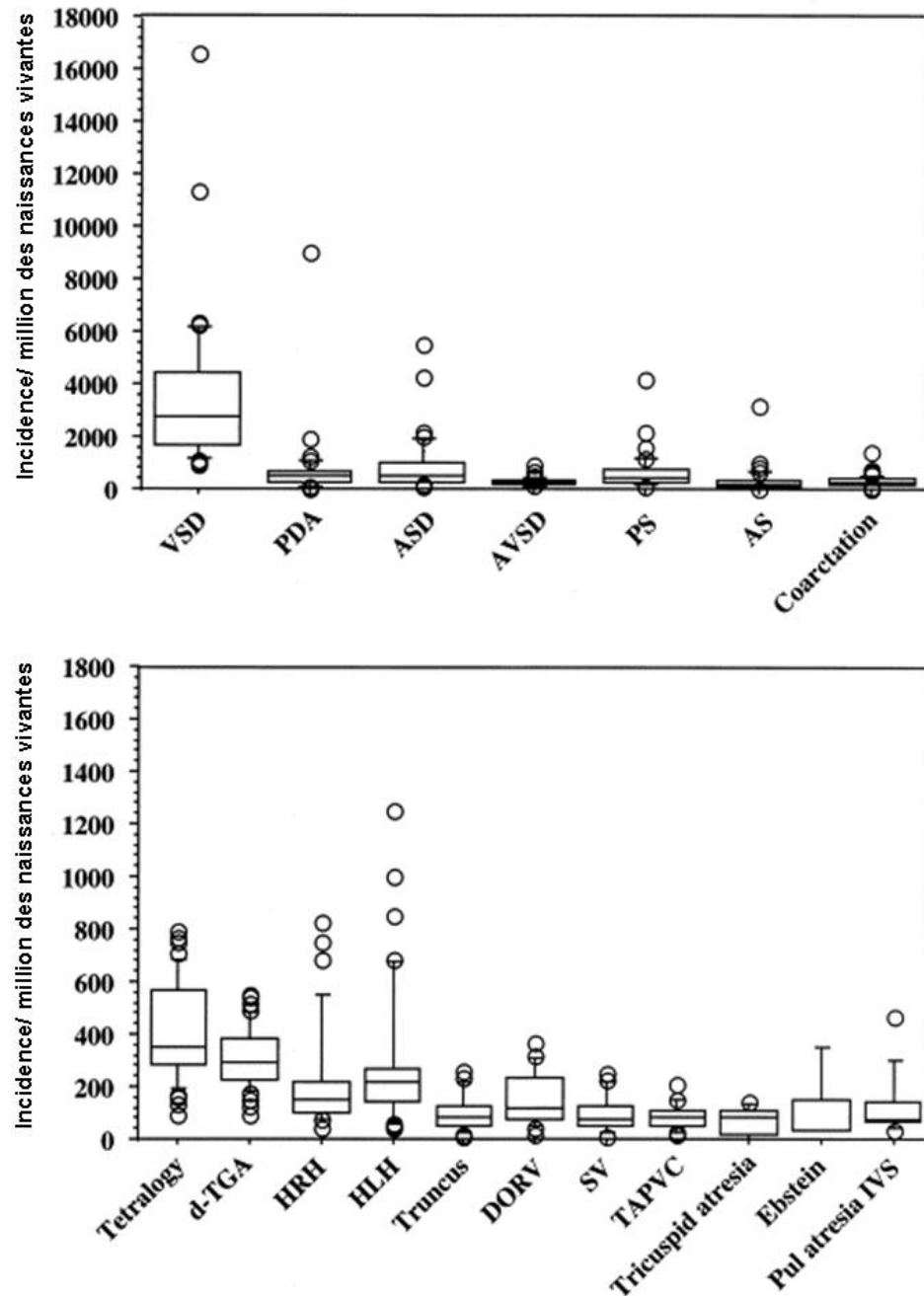


Figure 4. Variabilité de l'incidence par million des différentes formes de MCC à la naissance. Le rectangle montrant les quartiles au dessus et au dessous et la médiane est représentée par la ligne horizontale au milieu. Les deux lignes horizontales connectées avec le rectangle par lignes verticales minces indiquent les 90ème et 10ème percentiles, et les données isolées en dehors de ces limites sont représentées individuellement par des cercles. **(En dessus)** Les quatre principaux shunts gauche-droite (communication inter-ventriculaire membraneuse [VSD]; patent ductus arteriosus [PDA]; communication inter-auriculaire [ASD]; canal atrio-ventriculaire [AVSD]) et les trois principales lésions obstructives (sténose pulmonaire [PS]; sténose aortique [AS]; coarctation de l'aorte [Coarctation]). **(Au dessous)** Tétralogie de Fallot (Tetralogy); dextro-transposition des gros vaisseaux (d-TGA); hypoplasie du cœur droit (HRH), qui comprend atrésie tricuspide, maladie d'Ebstein, et atrésie pulmonaire à septum ventriculaire intact (pul atresia IVS); hypoplasie du cœur gauche (HLH); truncus arteriosus (Truncus); ventricule droit à double issue (DORV); ventricule unique (SV); retour veineux pulmonaire anormal total (TAPVC). Modifié d'après (Hoffman and Kaplan 2002).

II. Étiologie des malformations cardiaques congénitales

Malgré la prévalence élevée des malformations cardiaques congénitales, les étiologies de ces malformations sont peu connues : une cause est identifiée dans environ 20% des cas. Il peut s'agir d'une cause génétique, d'un tératogène ou d'un facteur environnemental. Cependant, aucune étiologie n'est mise en évidence dans la majorité des cas (80%). Une origine multifactorielle associant facteurs environnementaux et génétiques semble le mode de transmission habituel des malformations cardiaques congénitales (Nora 1968).

Les étiologies des malformations cardiaques peuvent être classées en quatre groupes :

- **les facteurs environnementaux et tératogènes**
- **les anomalies chromosomiques**
- **les formes monogéniques**
- **les formes multifactorielles** (mode de transmission proposé dans la plupart des cas).

II.1 Facteurs environnementaux et tératogènes

Une exposition à un facteur environnemental pendant la période périconceptionnelle est identifiée dans 2% des cas. Ces facteurs de risque sont classés en quatre grandes catégories (Patel and Burns 2013):

caractéristiques ou conditions parentales (Détailé dans la Table 1 de l'Annexe 2),

exposition maternelle à un médicament (Détailé dans la Table 2 de l'Annexe 2),

exposition parentale à une drogue (Détaillé dans la Table 3 de l'Annexe 2), et

exposition parentale environnementale (Détaillé dans la Table 4 de l'Annexe 2).

Certains des facteurs environnementaux décrits précédemment ont été associés à différents types de malformations cardiaques congénitales (Détaillé dans la Table 5 de l'Annexe 2).

Le diabète maternel est le risque environnemental le mieux documenté, associé à une augmentation du risque de malformations cardiaques congénitales cinq fois supérieur à la population générale (Wren et al. 2003). Le risque de malformation cardiaque lié au diabète est majeur lorsqu'il intervient avant la 7^{ème} semaine de grossesse. Par conséquent, les malformations cardiaques congénitales liées au diabète gestationnel interviennent pendant la formation du cœur primitif et incluent : anomalies de latéralité, anomalies de la boucle cardiaque, transposition des gros vaisseaux, canal atrio-ventriculaire, hypoplasie du cœur gauche, malformations de la voie efférente, cardiomyopathies et canal artériel patent (Jenkins et al. 2007). Les mécanismes par lesquels le diabète participe aux MCC sont peu connus. Un contrôle rigoureux de la glycémie maternelle préconceptionnelle et au cours de la grossesse permettrait une réduction du risque des MCC à la naissance (Cousins 1991).

Parmi les autres expositions environnementales qui peuvent augmenter le risque de MCC à la naissance, on note certaines maladies ou particularités maternelles (ex. phénylcétonurie, obésité, infection par rubéole, fièvre) pendant la grossesse, l'exposition maternelle à des médicaments ou drogues (ex. alcool) et l'exposition parentale à un tératogène (ex. herbicides).

II.2 Anomalies chromosomiques

Parmi les enfants nés vivants avec une malformation cardiaque, 8 à 10% présentent une anomalie chromosomique (Roos-Hesselink et al. 2005). Une anomalie chromosomique comprenant un ou plusieurs gènes impliqués dans le développement cardiaque peut entraîner une malformation cardiaque par différents mécanismes. Dans le cas de la

trisomie 21, c'est une copie supplémentaire d'un ou plusieurs gènes qui est responsable du phénotype anormal tandis que dans la microdélétion 22q11.2 (Lindsay 2001; Marino et al. 2001; Rauch et al. 2004; Théveniau-Ruissy et al. 2008), c'est l'haploinsuffisance du gène *TBX1* qui est en cause (Table 2). Cette perte ou gain de matériel chromosomique peut produire un effet sur des gènes sensibles au dosage et, en conséquence, être la cause de certaines malformations cardiaques congénitales. Concernant la fréquence des MCC chez les patients présentant un syndrome de Klinefelter, les estimations divergent. Ainsi, ce risque serait de 50% selon certaines publications (Pierpont et al. 2007; Mazen et al. 2010; Blue et al. 2012) (Table 2) tandis que d'autres rapportent rarement de MCC chez ces patients (Rosenthal 1972; Adatia et al. 1987; Visootsak et al. 2001).

Table 2. Anomalies chromosomiques associées aux MCC.

Syndrome	Chromosomal anomaly	Associated cardiac lesions	Proportion of patients with CHD
Down	Trisomy 21	AVSD, ASD, VSD, TOF	40%–50%
Edwards	Trisomy 18	VSD, ASD, DORV, TOF, CoA, HLHS	90%–100%
Patau	Trisomy 13	ASD, VSD, DORV, HLHS, L-TGA, AVSD, TAPVR, dextrocardia, PDA	80%
Turner	Monosomy X	CoA, AS, HLHS, PAPVR	25%–35%
Klinefelter	47, XXY	ASD, PDA, MVP	50%
Cat eye	Tetrasomy 22p	TAPVR, PAPVR	50%
Pallister–Killian	Tetrasomy 12p	VSD, CoA, PDA, ASD, AS	25%
Velocardiofacial	(Del 22q11.2)	IAA(B), TA, TOF, aortic arch anomalies	75%–85%
Williams	(Del 7q11.23)	SVAS ± PVS, PS, PPS	50%–80%

ASD = communication inter-auriculaire. AS = sténose aortique. AVSD = canal atrio-ventriculaire. CoA = coarctation de l'aorte. DORV = ventricule droit à double issue. HLHS = hypoplasie du cœur gauche. IAA(B) = interruption de l'arche aortique (type B). L-TGA = levo-transposition des gros vaisseaux. MVP = prolapsus valvulaire mitral. PAPVR = retour veineux pulmonaire anormal. PDA = patent ductus arteriosus. PPS = sténose pulmonaire périphérique. PS = sténose pulmonaire. PVS = sténose pulmonaire valvulaire. SVAS = sténose aortique supra-valvulaire. TA = truncus arteriosus. TAPVR = retour veineux pulmonaire anormal total. TOF = tétralogie de Fallot. VSD = communication inter-ventriculaire. D'après (Blue et al. 2012).

Les réarrangements chromosomiques tels que les translocations réciproques ou les inversions peuvent aussi être la cause de malformations en interrompant un gène à l'un des points de cassure des chromosomes affectés ou en dérégulant un gène lorsque le point de cassure est localisé entre le gène et des éléments régulateurs de ce gène.

Les variations en nombre de copies (CNV) d'une région spécifique de l'ADN d'un individu représentent un autre type d'anomalie chromosomique. Certains CNVs ont été associés à des maladies comme l'autisme et la schizophrénie, et également à des MCC ces dernières années. Ainsi, une étude récente a montré qu'une microduplication de la région 1q21.1 incluant le gène *GJA5* était identifiée chez environ 1% des patients présentant une TOF isolée (Soemedi et al. 2012a) faisant considérer *GJA5* comme un gène majeur des malformations cardiaques congénitales chez l'Homme.

II.3 Formes monogéniques

Les facteurs génétiques représentent un élément important dans le déterminisme de ces pathologies. Par exemple, le risque de malformation cardiaque chez un enfant issu de parents apparentés est considérablement plus élevé que celui observé pour des parents sans lien de parenté, suggérant une influence génétique très importante dans l'apparition des MCC (Naderi 1979).

➤ II.3.1 Formes syndromiques

Une transmission monogénique explique certaines formes syndromiques de MCC. Environ 3%-5% des MCC sont attribuées à des syndromes mendéliens dont les syndromes d'Alagille, Holt-Oram ou Noonan (van der Bom et al. 2011). La fréquence des anomalies cardiaques dans ces syndromes est importante (Table 3). De la même manière que pour les réarrangements chromosomiques, dans les syndromes mendéliens, pour une même mutation, le phénotype cardiaque peut être très variable.

Table 3. Microdeletions et syndromes monogéniques associés à une MCC.

Syndrome	Gene (region)	Associated cardiac lesions	Proportion of patients with CHD
Alagille	<i>JAG1, NOTCH1</i> (del 20p12)	PPS, TOF, ASD, PS	85%–95%
Noonan	<i>PTPN11, SOS1, KRAS, RAF1</i>	PVS, ASD, CoA, HCM	80%–90%
Holt–Oram	<i>TBX5</i>	ASD, VSD, AVSD, TOF	80%
Char	<i>TFAP2B</i>	PDA	60%
Ellis–van Creveld	<i>EVC, EVC2</i>	Primum ASD, common atrium, AVSD	60%
Smith–Lemli–Opitz	<i>DHCR7</i>	AVSD, primum ASD, VSD, PAPVR	45%
CHARGE	<i>CHD7, SEMA3E</i>	ASD, VSD, valve defects	50%–80%
Kabuki	<i>MLL2</i> in some cases	CoA, ASD, VSD	40%
Heterotaxy*	<i>ZIC3</i>	Dextrocardia, L-TGA, AVSD, TAPVR	90%–100%

ASD = communication inter-auriculaire. AVSD = canal atrio-ventriculaire. CoA = coarctation de l'aorte. HCM = cardiomyopathie hypertrophique. L-TGA = transposition corrigée des gros vaisseaux. PAPVR = retour veineux pulmonaire anormal. PDA = patent ductus arteriosus. PPS = sténose pulmonaire périphérique. PS = sténose pulmonaire. PVS = sténose pulmonaire valvulaire. TAPVR = retour veineux pulmonaire anormal total. TOF = tétralogie de Fallot. VSD = communication inter-ventriculaire.

* Syndrome d'hétérotaxie qui inclue des manifestations cardiaques et non cardiaques (ex. asplénie, polysplénie). L'hétérotaxie est aussi un terme général pour décrire différentes malformations cardiaques, incluant l'isomérisme des oreillettes droite et gauche. D'après (Blue et al. 2012).

➤ II.3.2 Formes non syndromiques

Le premier gène associé à des formes non syndromiques de MCC est *NKX2.5* (Schott et al. 1998). Plus récemment, un nombre relativement important de familles liées à des formes monogéniques non syndromiques de MCC ont été rapportées. A ce jour, environ 40 gènes différents ont été reliés à des formes non syndromiques de MCC (Wessels and Willems 2010). La majorité des gènes rapportés codent des facteurs de transcription mais on observe également d'autres types de protéines par exemple des protéines structurales musculaires (Table 4). Dans ces formes monogéniques non syndromiques, une transmission autosomique dominante à pénétrance incomplète est habituellement observée.

Table 4. Gènes sélectionnés associés à des formes non syndromiques de MCC.

Gene	Function	Associated cardiac lesions
<i>NKX2-5</i>	Transcription factor	ASD–AV block, TOF, HLHS, TGA, DORV, Ebstein anomaly, VSD
<i>NKX2-6</i>	Transcription factor	TA
<i>GATA4</i>	Transcription factor	ASD ± PS, TOF, VSD, DORV
<i>GATA6</i>	Transcription factor	TA, TOF, AVSD
<i>TBX1</i>	Transcription factor	IAA, aortic arch anomalies, VSD
<i>TBX5</i>	Transcription factor	ASD, VSD, AVSD, conduction abnormalities
<i>TBX20</i>	Transcription factor	ASD, VSD, valve defects, LVOTO
<i>CITED2</i>	Transcription factor	ASD, VSD, TOF, TGA
<i>ZIC3</i>	Transcription factor	Heterotaxy, ASD, AVSD, TGA, VSD, TAPVR, PS
<i>ZFPM2</i>	Transcription factor	TOF
<i>FOXH1</i>	Transcription factor	TOF, VSD
<i>HAND1</i>	Transcription factor	HLHS (somatic mutation)
<i>TFAP2B</i>	Transcription factor	PDA
<i>NOTCH1</i>	Membrane ligand–receptor	AS, BAV
<i>NODAL</i>	Membrane ligand–receptor	Heterotaxy, TGA
<i>JAG1</i>	Membrane ligand–receptor	PS, TOF
<i>CFC1</i>	Membrane ligand–receptor	Heterotaxy, TGA, DORV, TOF
<i>MYH6</i>	Sarcomeric protein	ASD
<i>MYH7</i>	Sarcomeric protein	ASD, Ebstein anomaly
<i>MYH11</i>	Sarcomeric protein	PDA
<i>ACTC1</i>	Sarcomeric protein	ASD, VSD
<i>GJA1</i>	Gap junction protein	HLHS (somatic mutation)
<i>GJA5</i>	Gap junction protein	TOF
<i>CRELD1</i>	Matricellular protein	AVSD, dextrocardia
<i>ELN</i>	Structural protein	SVAS
<i>VEGFA</i>	Mitogen	TOF

ASD = communication inter-auriculaire. AS = sténose aortique. AV = auriculo-ventriculaire. AVSD = canal atrio-ventriculaire. BAV = valve aortique bicuspidée. DORV = ventricule droit à double issue. HLHS = hypoplasie du cœur gauche. IAA = interruption de l'arc aortique. LVOTO = obstruction de la voie d'éjection ventriculaire gauche. PDA = patent ductus arteriosus. PS = sténose pulmonaire. SVAS = sténose aortique supra-avalvulaire. TA = truncus arteriosus. TAPVR = total anomalous pulmonary venous return. TGA = transposition des gros vaisseaux. TOF = tétralogie de Fallot. VSD = communication inter-ventriculaire. D'après (Blue et al. 2012).

II.4 Formes multifactorielles

Actuellement environ 20% des MCC sont attribuées à des causes connues décrites précédemment, mais aucune étiologie n'est retrouvée chez plus de 80% des patients avec une malformation cardiaque congénitale isolée. L'hypothèse proposée est que les MCC pour lesquelles l'étiologie est inconnue suivent un modèle de transmission multifactorielle, dans lequel sont impliqués des facteurs génétiques et environnementaux participant au développement de la maladie (Nora 1968; Driel et al. 2008; van et al. 2008; Verkleij-Hagoort et al. 2008). Le modèle suppose des variations dans différents gènes, chacune contribuant à la malformation (Burn et al. 1998). Ces gènes porteurs de variations interagissent entre eux ainsi qu'avec des facteurs environnementaux pour augmenter la probabilité d'apparition d'une MCC. La majorité des cas sporadiques de malformations cardiaques congénitales serait d'origine multifactorielle.

II.5 Risque de récurrence des MCC

Diverses études ont rapporté un risque élevé de récurrence des MCC compatibles avec le modèle multifactoriel (Gill et al. 2003; Loffredo et al. 2004; Calcagni et al. 2007; Øyen et al. 2009). Le risque de récurrence des MCC lorsqu'un enfant de la famille est atteint est estimé à environ 1-6% (Table 5) tandis que ce risque augmente à environ 1-20% (Table 5) lorsqu'un des parents est atteint d'une MCC. Dans le cas où deux frères ou sœurs sont atteints, le risque de récurrence peut augmenter jusqu'à 10%. Le risque de récurrence varie selon le type de MCC (Table 5). Les lésions obstructives, incluant l'hypoplasie du cœur gauche, la sténose valvulaire aortique et la coarctation de l'aorte ont un risque de récurrence généralement plus élevé dans la fratrie de parents non atteints et dans la descendance de parents atteints comparé aux autres types de MCC (Lewin et al. 2004).

Table 5. Risque de récurrence des formes non syndromiques des MCC.

Cardiac lesion	RR in siblings with unaffected parents		RR in children of affected parents	
	1 child affected	≥ 2 children affected	Mother affected	Father affected
VSD	3%	10%	9%–10%	2%–3%
ASD	2%–3%	8%	6%	1%–2%
TOF	2%–3%	8%	2%–5%	1%–2%
CoA	2%	6%	4%	2%–3%
AS	2%	6%	12%–20%	5%
PS	2%	6%	6%–7%	2%
HLHS	3%	10%	nr	nr
AVSD	3%–4%	nr	10%–14%	1%
PA	1%	3%	nr	nr
TA	1%	3%	nr	nr
TGA	1%–2%	5%	nr	nr
L-TGA	5%–6%	nr	nr	nr
Ebstein anomaly	1%	3%	6%	nr
Heterotaxy	5%–6%	nr	nr	nr
Overall	1%–6%	3%–10%	2%–20%	1%–5%

ASD=communication interauriculaire. AS =sténose aortique. AVSD= canal atrio-ventriculaire. CoA =coarctation de l'aorte. HLHS= hypoplasie du cœur gauche. L-TGA = transposition corrigée des gros vaisseaux. nr=non rapporté. PA= atrésie pulmonaire. PS= sténose pulmonaire. TA =truncus arteriosus. TGA= transposition des gros vaisseaux. TOF=tétralogie de Fallot. VSD=communication inter-ventriculaire. D'après (Blue et al. 2012).

Dans le cas des tétralogies de Fallot, l'analyse de ces données suggère une transmission polygénique faisant intervenir un petit nombre de gènes. Un modèle à trois gènes à effet multiplicatif semble le plus probable (Figure 5). Par ailleurs, le risque de récurrence des MCC est plus élevé chez les enfants dont les mères sont atteintes que chez les enfants dont les pères sont atteints d'une malformation cardiaque congénitale (Burn et al. 1998) (Table 5). La raison de ces différences n'est pas établie à ce jour.

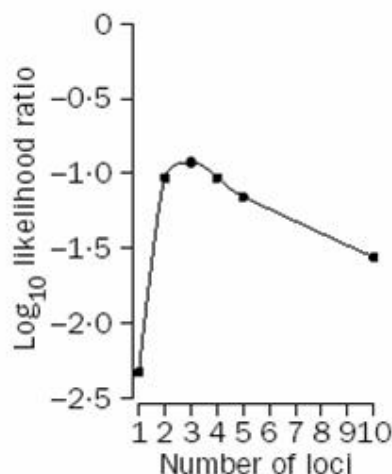


Figure 5. Rapport de vraisemblance, en fonction du nombre de loci, appliqué aux tétralogies de Fallot. *Le modèle à trois gènes est le plus probable. D'après (Burn et al. 1998).*

Le taux de concordance (c'est-à-dire le même type de malformation) des MCC chez les patients d'une même famille peut varier significativement selon le type de MCC. La concordance exacte globale des MCC rapportée est de 37% (Gill et al. 2003).

Estimer les parts respectives de la prédisposition génétique et des facteurs environnementaux est difficile. Récemment, de nombreux travaux contribuent à la compréhension des facteurs génétiques à l'origine des malformations cardiaques congénitales.

III. Éléments d'embryologie cardiaque

III.1 Formation du tube cardiaque primitif

La formation du tube cardiaque primitif et de la boucle cardiaque s'effectue du 15^{ème} jour au 21^{ème} jour de vie intra-utérine chez l'homme.

L'ébauche cardiaque n'apparaît qu'au début de la troisième semaine de vie de l'embryon. Pendant les deux premières semaines de vie intra-utérine, celui-ci est un

disque plat formé seulement par deux couches cellulaires (épiblaste et hypoblaste) et ne possède ni cœur ni système circulatoire.

La ligne primitive apparaît dans la partie médiane et caudale du disque embryonnaire, établissant l'axe longitudinal de l'embryon. Vers le 16^{ème} jour, les cellules épiblastiques migrent vers la ligne primitive et s'invaginent (gastrulation), conduisant à la différenciation de l'embryon en trois couches cellulaires : ectoderme, mésoderme et endoderme.

Le cœur dérive du mésoderme antérieur ou mésoblaste situé entre l'endoderme, feuillet ventral (entoblaste) et l'ectoderme, feuillet dorsal (ectoblaste). Les cellules précardiaques du mésoderme se groupent en amas cellulaires angioformateurs. Ces amas, qui trouvent leur origine de chaque côté de la ligne primitive, migrent vers la partie céphalique de l'embryon et se rejoignent à l'avant de la plaque neurale pour former un plexus vasculaire en forme de fer à cheval : la plaque cardiogénique (Figure 6).

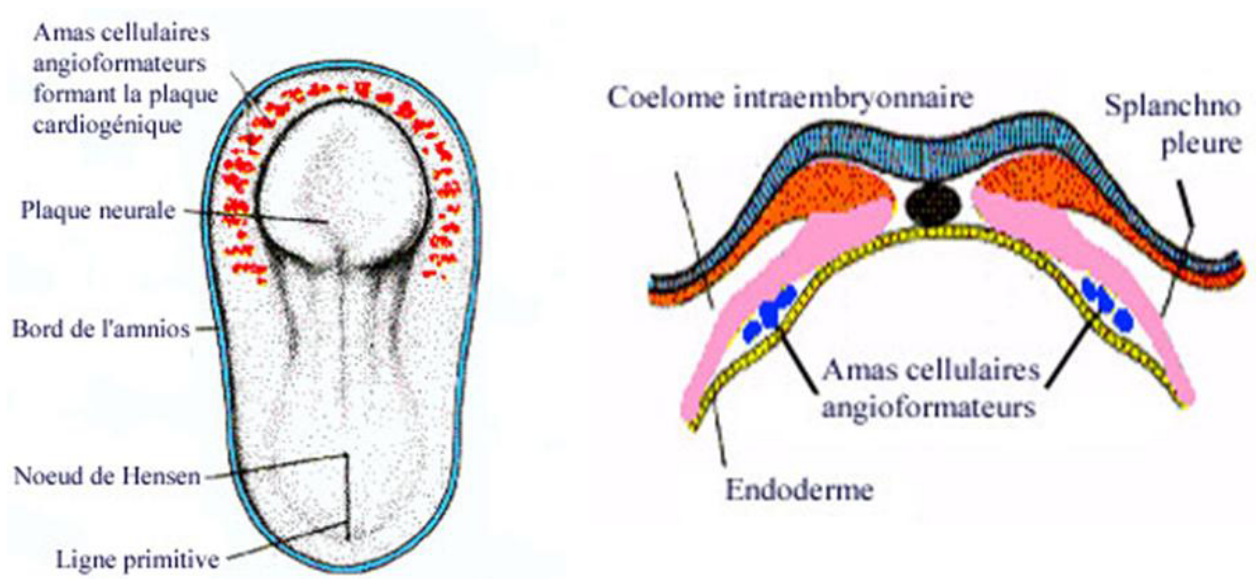


Figure 6. Vue dorsale et section transversale d'un embryon au stade présomitique. Embryon d'environ 18 jours après résection de l'amnios. Modifié d'après www.med.unc.edu/embryo_images/.

Secondairement, la plaque cardiogénique est clivée en deux couches, ventrale et dorsale, par la cavité coelomique. La couche dorsale ou somatopleure formera le péricarde. La plus grande partie de la couche ventrale ou splanchnopleure donnera

naissance au myocarde. Le coelome intraembryonnaire formera, outre la cavité péricardique, les cavités pleurale et péritonéale. C'est à ce stade que se produit, en raison de la croissance rapide du tube neural (futur cerveau) et de la flexion céphalocaudale de l'embryon, un changement d'orientation de la plaque cardiogénique, qui se trouve en arrière et en position ventrale par rapport à l'ébauche du système nerveux central et à la membrane buccopharyngée. L'ébauche cardiaque, sous forme de deux plaques latérales, va s'invaginer pour acquérir une structure tubulaire (les tubes cardiaques primitifs) et être repoussée en position cervicale, puis thoracique. Dans le même temps, en raison de la croissance rapide des somites, le disque embryonnaire se replie aussi transversalement. Ce phénomène d'enroulement latéral de l'embryon permet le rapprochement des deux tubes cardiaques latéraux, qui vont ensuite fusionner sur la ligne médiane pour former le tube cardiaque primitif (Figure 7). Le processus de fusion entre les deux tubes cardiaques primitifs, situés à l'origine en dehors de la cavité péricardique, s'accompagne de l'invagination de cette dernière. Au point où se produit cette invagination, le tube cardiaque reste attaché temporairement à la paroi dorsale de la cavité péricardique par un feuillet mésoblastique, le mésocarde dorsal.

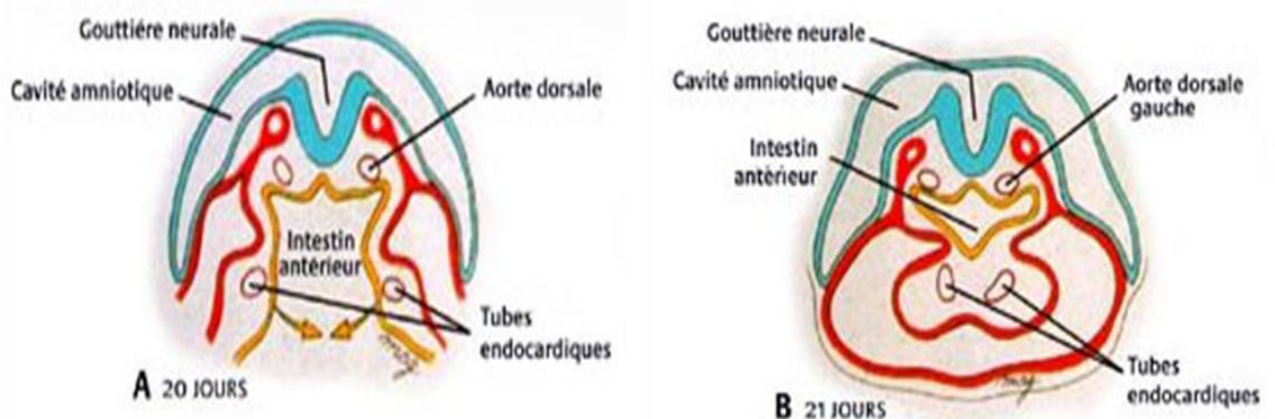


Figure 7. Transformation des amas cellulaires angioformateurs en deux tubes cardiaques ou endocardiques puis fusion sur la ligne médiane pour former le tube cardiaque primitif.

D'après www.uvp5.univ-paris5.fr/campus-gyneco-obst/.

À l'origine, le tube cardiaque n'est constitué que d'un endothélium. Dès le 22^{ème} jour, une épaisse masse de mésoblaste vient se disposer autour du tube cardiaque et se différencie en deux couches : le myocarde et la gelée cardiaque. La gelée cardiaque est

un épais manteau acellulaire sécrété par le myocarde en voie de développement contenant des glycosaminoglycanes et des protéines. Elle semble remplir un rôle mécanique essentiel lors des premiers moments de l'activité cardiaque. En effet, l'existence d'une zone amorphe, peu compressible, organisée en manchon glissé entre les premières cellules contractiles et le recouvrement interne de la lumière du tube cardiaque produit un effet démultiplicateur de la contraction des premières cellules myocardiques. La formation de tube cardiaque primitif est normalement achevée au 22^{ème} jour de la vie intra-utérine. C'est à ce stade que les premières contractions cardiaques vont apparaître. Dès cet instant, le cœur va devoir se développer en même temps qu'assurer son rôle de pompe. À ce stade, le tube cardiaque primitif est encore rectiligne (Figure 8).

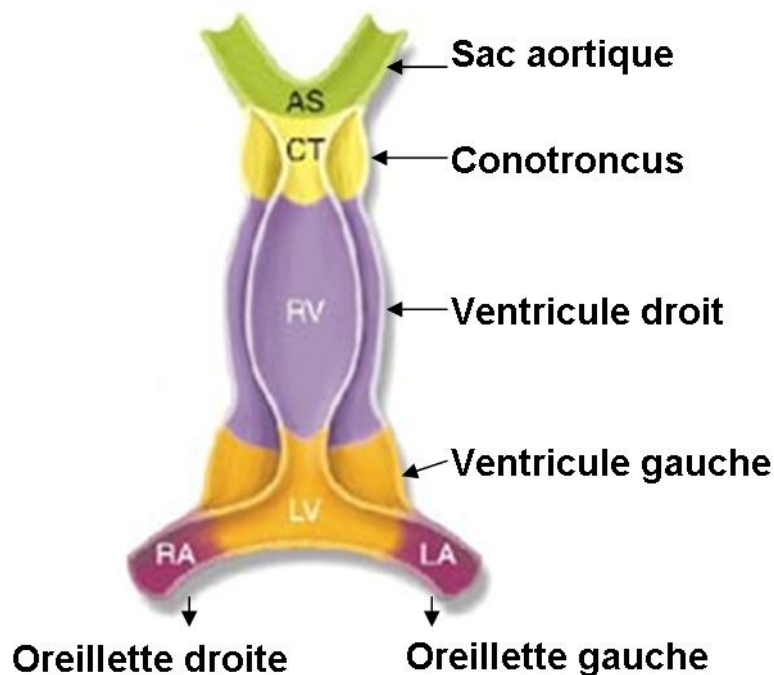


Figure 8. Tube cardiaque primitif rectiligne. Modifié d'après (Srivastava and Olson 2000).

III.2 Formation des cavités cardiaques

➤ III.2.1 La boucle cardiaque (looping)

La formation de la boucle cardiaque est la première manifestation anatomique de la latéralisation droite gauche de l'embryon (Männer 2000). Le tube cardiaque primitif rectiligne va s'allonger et s'incurver pour former une boucle à convexité droite le 23ème-24ème jour de vie intra-utérine (D-loop), initialement en forme de C, puis, après la rupture du mésocarde dorsal, de forme sigmoïde en S (Figure 9). Cette étape est cruciale pour la morphologie du cœur embryonnaire, apportant les futures cavités cardiaques dans sa position spatiale relative. En cas de renversement complet de la boucle, le cœur apparaît en *situs inversus totalis* ou en miroir. Le renversement peut être incomplet et aléatoire, donnant lieu à différentes anomalies, parfois intégrées à un syndrome avec hétérotaxie.

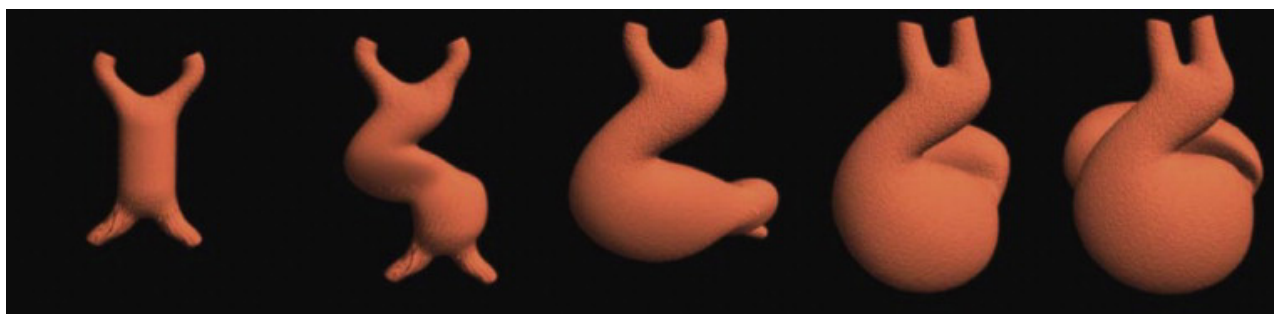


Figure 9. Formation normale de la boucle cardiaque. Au 23ème jour de vie intra utérine, le tube cardiaque primitif rectiligne s'allonge et s'incurve pour former une boucle à convexité droite (D-loop). D'après (Schleich et al. 2013).

➤ III.2.2 La convergence

Le deuxième stade de la boucle, ou complétion (*late looping*) implique la convergence de deux types de voies en parallèle : d'admission et d'éjection. La convergence amène les deux voies ensemble craniocaudalement pour que l'alignement du conotruncus avec le septum auriculaire, ventriculaire et atrio-ventriculaire se fasse de façon correcte (Yelbuz et al. 2002) (Figure 10). Lorsque la boucle est complétée, le cœur comporte cinq segments individualisables : oreillette gauche primitive, oreillette droite primitive, ventricule gauche primitif, ventricule droit primitif et conotruncus, reliés entre eux par des jonctions : canal auriculo-ventriculaire, foramen inter-ventriculaire primitif, sillon bulbo-ventriculaire et jonction conotroncale.

Les malformations cardiaques résultantes d'une erreur lors de la convergence sont souvent sévères car elles concernent les deux ventricules et les valves auriculo-ventriculaires. Elles peuvent être causées par un défaut de croissance ventriculaire ou par une absence ou anomalie du développement de la jonction auriculo-ventriculaire droite. Ces anomalies résultent d'un malalignement entre le septum atrial et ventriculaire donnant lieu à diverses malformations cardiaques congénitales, dont les ventricules à double issues, les atrésies tricuspides et les hypoplasies ventriculaires.

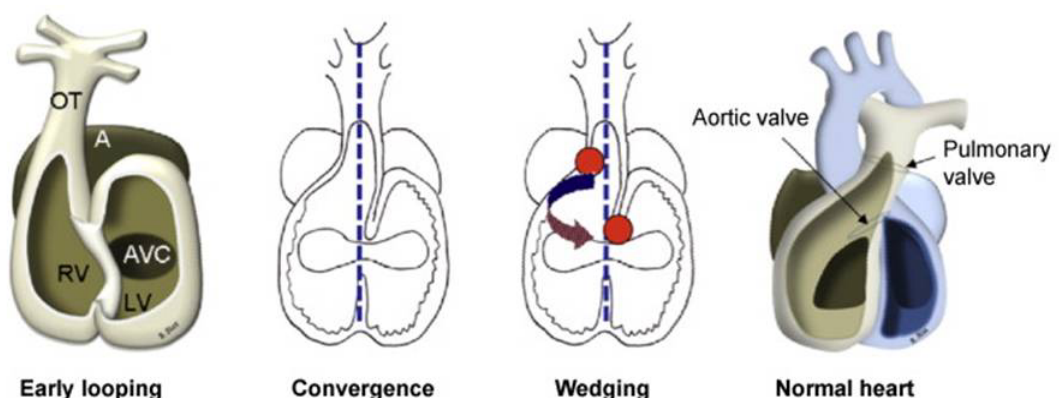


Figure 10. Les trois principales étapes du développement cardiaque : *early looping*, *convergence*, *wedging* et cœur normal. Les cercles rouges indiquent la position initiale et finale de la future valve aortique. A : oreillette ; AVC : canal auriculo-ventriculaire ; LV : ventricule gauche ; OT : voie efferente ; RV : ventricule droit. D'après (Schleich et al. 2013).

➤ III.2.3 Le wedging

Le second champ cardiaque (second heart field : SHF) facilite l'élongation du conotruncus pour l'addition de cellules myocardiques, en réponse à la migration des cellules de la crête neurale vers les voies efférentes (Waldo et al. 2005b). Cette élongation du conotruncus est nécessaire au bon alignement (convergence) et au wedging.

Pendant le wedging, la paroi myocardique du conotruncus subit une rotation anti-horaire, vu du côté ventriculaire, de manière à ce que la valve aortique passe derrière le tronc pulmonaire, vers le bas et la gauche pour s'installer entre les deux valves auriculo-ventriculaires (Yelbuz et al. 2002; Bajolle et al. 2006) (Figure 10).

Un malalignement entre le conotruncus et les ventricules résulte inévitablement en une fêlure de la fusion du septum d'éjection avec le septum ventriculaire primitif et en une communication inter-ventriculaire (VSD). La Tétralogie de Fallot peut être le résultat d'un arrêt de la rotation du conotruncus à la base des gros vaisseaux (Lomonico et al. 1988). On peut considérer alors que la Tétralogie de Fallot est une erreur dans la dernière étape du looping cardiaque-wedging qui conduit à un malalignement entre le conotruncus et les ventricules (Anderson and Jacobs 2008; Van Praagh 2009).

Une convergence et wedging incomplet ou anormal sont la cause majeure de plusieurs malformations cardiaques congénitales, en particulier du ventricule droit à double issue (Bartram et al. 2001). Les anomalies pendant la convergence et le wedging produisent un malalignement des segments d'admission et d'éjection.

La suite du développement du cœur consiste en un remaniement de ces cavités et à l'apparition, entre elles, de septa et de valves.

III.3 Développement du conotruncus

Le conus est la région du cœur, en forme de cône, qui fait suite à la partie distale du ventricule primitif et la relie au troncus, la région entre les deux étant l'emplacement des futures valves semi-lunaires. Son devenir est de donner naissance aux voies d'éjection des deux ventricules. Jusqu'au wedging, le conotruncus reste entièrement au-dessus du ventricule droit en développement.

Le conus a en charge la répartition des entrées et des sorties ventriculaires et réalise ces deux missions par deux mouvements simultanés de migration et de rotation pendant le phénomène de wedging. La migration du cône est un phénomène complexe, qui le porte du flanc droit du cœur successivement vers l'avant, puis vers la gauche, puis vers l'arrière. Il s'insinue entre les deux orifices auriculo-ventriculaires issus du canal atrio-ventriculaire primitif et participe ainsi à la répartition des entrées ventriculaires, avec un effacement du sillon auriculo-conal. Il donne aussi le septum conal ou infundibulaire, qui participe à la formation du septum ventriculaire dans sa portion la plus antérieure et haute. La rotation du cône répartit les sorties ventriculaires. Initialement, le conus est surplombé par le troncus, dont la partie droite, rattachée aux III° et IV° arcs, donnera l'aorte. La partie gauche, rattachée au VI° arc, donnera l'artère pulmonaire. Ainsi, les deux vaisseaux primitifs sont en position de transposition avec une aorte antéro-droite et partent tous deux du futur ventricule droit. Le cœur embryonnaire est un ventricule droit à double issue jusqu'au 40^{ème} jour. La mise en position des deux vaisseaux se fait par une rotation inversée des deux ostia du cône. L'ostium proximal tourne de 90° en sens horaire et l'ostium distal de 150° en sens antihoraire. Les torsions débutent à la 5^{ème} semaine et se terminent à la 6^{ème}. La rotation de l'ostium proximal est due à un glissement de la gelée cardiaque sur le myocarde, et celle de l'ostium distal à une croissance différentielle de la paroi conale.

Le troncus suit passivement les mouvements du cône. Il se tord sur lui-même de 150°, portant l'artère pulmonaire en avant, à droite et en haut. L'aorte vient à gauche et en arrière, et s'abaisse du fait de la croissance différentielle de la paroi du cône. Elle

s'approche ainsi du futur ventricule gauche, profitant de l'effacement du sillon auriculo-conal. L'appariement et la fusion de crêtes conotruncales dont le nombre est encore discuté permet le cloisonnement normal en spirale du conus, la connexion ventriculo-artérielle et la disposition normale des vaisseaux (artère pulmonaire antérieure et droite s'enroulant au flanc gauche de l'aorte).

Deux structures contribuent à l'origine du développement du conotruncus : les cellules de la crête neurale cardiaque et le second champ cardiaque..

➤ III.3.1 Rôle des cellules de la crête neurale cardiaque

La crête neurale est une structure qui apparaît très tôt dans le développement de l'embryon et qui n'a qu'une existence temporaire. Les cellules dont elle est composée migrent à travers l'organisme et vont se différencier pour former les cellules de divers organes. Seule une région très limitée de la portion crâniale de la crête neurale contient des cellules à destinée cardiaque. Les cellules de la voie d'éjection cardiaque ont donc deux origines. Une fraction des cellules dérive de l'endocarde, l'autre a migré à partir de la crête neurale, par l'intermédiaire des arcs branchiaux III, IV et VI (Figure 11). Les cellules migrant vers la voie d'éjection colonisent les septa truncal, aortopulmonaire et conal et les valves semi-lunaires (Hutson and Kirby 2007).

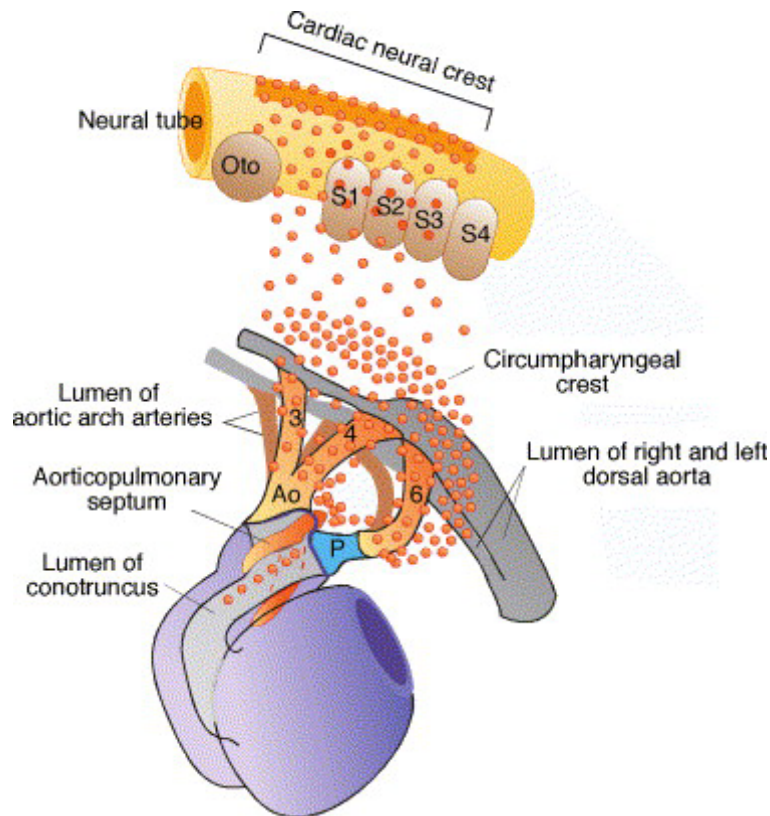


Figure 11. Représentation schématique de la migration des cellules de la crête neurale cardiaque des rhombomères 6-8 à travers les arcs pharyngés 3, 4 et 6. Les arcs aortiques 3, 4 et 6 traversent ces arcs pharyngés et se connectent au sac aortique. Modifié d'après (Hutson and Kirby 2007).

Les cellules de la crête neurale cardiaque contribuent à la morphogenèse cardiaque, mais également à la formation du thymus, de la thyroïde et des glandes parathyroïdes. Le rôle des cellules de la crête neurale cardiaque dans le développement du pôle artériel du cœur a été démontré au début des années 70s à partir des expériences d'ablation de la crête neurale, réalisées sur l'embryon de poulet lesquelles ont permis d'en préciser le rôle en étudiant les malformations observées. L'ablation de la crête neurale cardiaque avant que ne se produise la migration cellulaire aboutit à un large éventail de malformations cardiaques, selon l'étendue de l'ablation. Ces malformations peuvent être classées en deux catégories :

- un vaisseau unique par absence de septation de la voie d'éjection (tronc artériel commun ou troncus arteriosus)
- un malalignement au niveau du septum conal (ventricule droit à double issue).

L'ablation totale aboutit presque toujours à un tronc artériel commun. Une ablation partielle conduit à la présence de diverses anomalies cardiaques, comme la Tetralogie de Fallot (ToF) avec ou sans atrésie pulmonaire, ventricule droit à double issue (DORV) et malalignement au niveau du septum conal, toujours associé à des anomalies d'arc aortique (Le Lièvre and Le Douarin 1975). La position du vaisseau unique par rapport aux ventricules sous jacents est variable, mais il se trouve le plus souvent au-dessus du ventricule droit. La transposition des gros vaisseaux est exceptionnelle après ablation de la crête neurale. À ces malformations s'associent de façon variable des anomalies des arcs aortiques : interruption de l'arc (ou crosse) aortique, persistance de vaisseaux embryonnaires. Ces malformations cardiaques prises dans son ensemble sont connues sous le nom de malformations conotroncales.

Le conotruncus a alors deux origines cellulaires: les cellules de la crête neurale cardiaque pour la portion distale des gros vaisseaux ; et le second champ cardiaque pour la portion proximale des gros vaisseaux (cellules musculaires lisses) et le conus (cardiomyocytes). Un dysfonctionnement des cellules de la crête neurale cardiaque ou du second champ cardiaque se traduirait par un défaut de *wedging* et par une malformation cardiaque conotroncale (Waldo et al. 2005a). Le rôle majeur des cellules de la crête neurale est d'envoyer des signaux aux cellules du SHF du conotruncus afin d'orienter leur différenciation en cellules myocardiques (Waldo et al. 2005a). Ces cellules permettent la croissance et l'élongation du conotruncus, nécessaire au bon déroulement du phénomène de *wedging*. Les expériences d'ablation du second champ cardiaque réalisées chez des embryons de poulet ont permis d'observer des défauts d'alignement du conotruncus, de l'élongation et du *wedging*, conduisant à des malformations conotroncales (Ward et al. 2005). Si l'ablation est réalisée plus tard pendant le développement, le second champ cardiaque n'intègre pas de cellules musculaires lisses à la portion proximale des gros vaisseaux, ce qui peut être responsable d'anomalies de position des orifices coronaires (Ward et al. 2005).

➤ III.3.2 Septation du conotruncus

L'aorte et l'artère pulmonaire viennent toutes deux du sac aortique (Anderson et al. 2012). Dans la partie dorsale du sac aortique, une protubérance mésenchymateuse nommée épine artérielle va vers la partie distale des bourgeons endocardiques du conus créer le septum aortico-pulmonaire primitif. La fusion entre les parties distales des bourgeons endocardiques donne naissance à l'aorte à droite et à l'artère pulmonaire à gauche, qui se connecte avec les arcs aortiques 4 et 6, respectivement.

Le foramen aortico-pulmonaire embryonnaire est fermé par la fusion des bourgeons endocardiques distaux avec l'épine artérielle. Un défaut de fermeture du foramen donne lieu à une fenêtration aortico-pulmonaire.

Les malformations cardiaques résultant du développement anormal du conotruncus incluent : les malformations conotruncales obtenues à partir d'un wedging anormal (Tétralogie de Fallot et ses variantes ; truncus arteriosus ; ventricule droit à double issue, interruption de l'arc aortique et malalignement du septum conal) ; et la transposition des gros vaisseaux (TGA), qui implique le conotruncus mais est considérée comme une anomalie de latéralité. La TGA affecte la latéralité du wedging et a en commun des gènes impliqués dans les syndromes avec hétérotaxie (Ramsdell 2005).

IV. Les malformations cardiaques congénitales de la voie efférente

Les malformations cardiaques congénitales de la voie efférente constituent un groupe hétérogène de malformations tant sur le plan anatomique que physiologique puisqu'il peut s'agir soit de cardiopathies obstructives du cœur droit à bas débit pulmonaire (tétralogie de Fallot), soit d'obstacle gauche (interruption de la crosse aortique) ou de discordance ventriculo-artérielle (transposition des gros vaisseaux).

Au cours du développement cardiaque, une mauvaise rotation du septum conal avec bascule antérieure joue un rôle central dans l'apparition de certaines de ces cardiopathies. Par exemple, les quatre éléments de la tétralogie de Fallot ne sont que la conséquence de cette bascule antérieure : communication inter-ventriculaire par malalignement entre le septum conal et le septum musculaire rendant impossible la formation du septum membraneux, sténose infundibulaire par rétrécissement de la voie sous pulmonaire, chevauchement de la communication inter-ventriculaire par l'aorte déplacée antérieurement et hypertrophie ventriculaire droite par égalité de pression systolique entre les deux ventricules. La perturbation de la formation d'une structure embryonnaire provenant, soit de la migration, soit de la différenciation d'un groupe de cellules, dont les cellules de la crête neurale cardiaque, peut conduire à des phénotypes variés allant d'une forme mineure à une forme sévère. Appliquée aux cardiopathies conotruncales, la bascule antérieure du septum conal, quand elle est modérée, donne une communication inter-ventriculaire par malalignement. Plus marquée, elle conduit à une tétralogie de Fallot. Extrême, c'est l'atrésie pulmonaire avec communication inter-ventriculaire. La bascule postérieure du septum conal sous l'aorte, miroir de la bascule antérieure, conduit aux syndromes de coarctation et à l'interruption de la crosse aortique. À côté de ces mécanismes d'asymétrie de septation de la région conotruncale, une malrotation lors de la septation du conotruncus donnera une malposition des gros vaisseaux, une absence totale de septation un tronc artériel commun et une absence de résorption du conus sous aortique un ventricule droit à double issue (Figure 12).

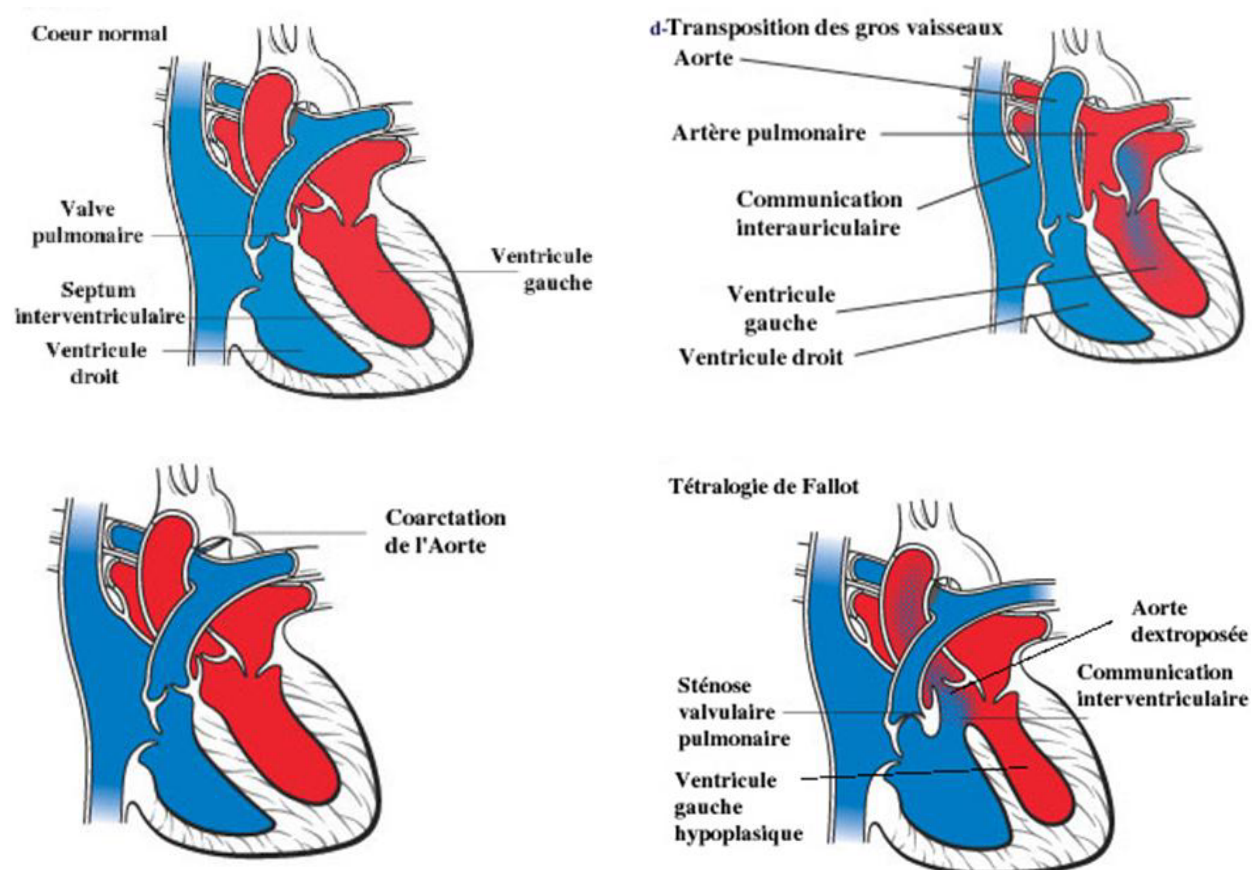


Figure 12. Représentation schématique du cœur normal et des principales cardiopathies conotroncales. Modifié d'après www.americanheart.org/.

IV.1 Incidence des malformations cardiaques congénitales de la voie efférente

Les malformations cardiaques affectant la voie efférente constituent 30% des malformations cardiaques congénitales (Thom et al. 2006). Ces malformations sont la cause d'une morbidité et mortalité significatives chez les enfants et les adultes.

IV.2 Coarctation de l'aorte

La coarctation de l'aorte a été décrite dès 1760 par l'anatomiste prussien Johann Friedrich Meckel (JARCHO 1961). Elle est la 5ème malformation cardiaque congénitale la plus fréquente. L'incidence de cette malformation est estimée à environ 1 pour 2.500 naissances vivantes. Elle représente 5-8% des malformations cardiaques congénitales totales (Kenny and Hijazi 2011), avec une prédominance plus forte chez les hommes que chez les femmes (2 :1). Cette malformation cardiaque peut être isolée ou associée à d'autres lésions, la plus habituelle est la valve aortique bicuspide et la communication inter-ventriculaire (VSD). La coarctation de l'aorte peut être associée à une transposition des gros vaisseaux. Par contre, il est très rare de la trouver chez des patients présentant une obstruction conotroncale sévère du ventricule droit comme la tétralogie de Fallot.

La coarctation de l'aorte est définie comme un rétrécissement du segment aortique proche du ligamentum arteriosum adjacent à l'artère subclavière gauche (Warnes et al. 2008). Cette coarctation peut être discrète ou étendue, simple ou complexe, impliquant l'arc aortique ou isthme, et peut avoir des vaisseaux collatéraux (Figure 13).

Autrefois, la coarctation de l'aorte était classée en pré-ductale ou post-ductale, selon que le segment rétréci était localisé en position proximale ou distale par rapport au ligamentum ou ductus arteriosus, respectivement. Actuellement, ces différentes formes de coarctations sont regroupées sous le terme de coarctation juxta-ductales.

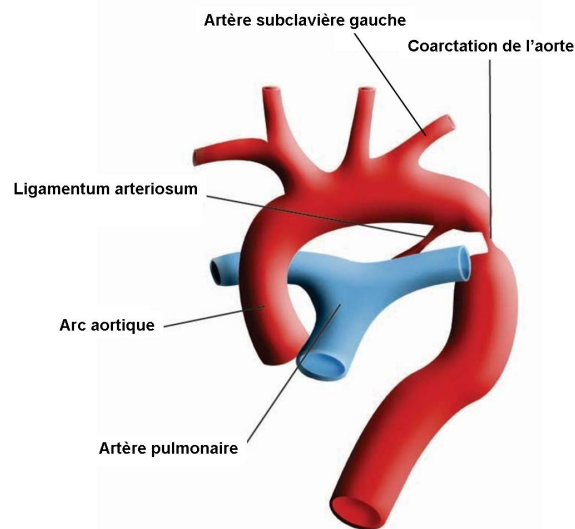


Figure 13. Représentation schématique de la coarctation de l'aorte post-ductale. Modifié d'après <https://gmep.org/media/13021>.

Le mécanisme exact expliquant la coarctation de l'aorte n'est pas clairement connu. Les deux hypothèses habituellement avancées sont la théorie hémodynamique d'une part et la théorie mécanique ou du tissu ductal ectopique d'autre part. La théorie hémodynamique fait jouer un rôle à l'hypoperfusion de l'isthme aortique qui résulte de certaines cardiopathies telles qu'une communication interventriculaire ou une sténose sous-aortique. Elle pourrait expliquer l'absence de développement isthmique pendant la vie foetale et l'hypoplasie tubaire fréquemment constatée dans la coarctation néonatale. La clôture spontanée du ductus arteriosum complète le développement de cette obstruction (Rudolph et al. 1972; Talner and Berman 1975). Dans la théorie mécanique, du tissu ductal ectopique migrant au niveau de la paroi aortique pourrait être responsable de la striction de cette paroi en regard de l'émergence du canal artériel. Cette théorie s'appuie sur la découverte, sur certaines pièces anatomiques de véritables anneaux de striction de tissu ductal dans la paroi aortique. Cette théorie n'explique pas les différents degrés d'hypoplasie de l'isthme et arc aortique associés à la coarctation de l'aorte.

Dans des séries historiques, les coarctations de l'aorte qui ne sont pas traitées donnent lieu à une mortalité de plus de 80% vers l'âge de 50 ans en raison de complications telles que la rupture aortique, l'insuffisance cardiaque et l'hémorragie intracrânienne (Brown et al. 2013).

IV.3 Transposition des gros vaisseaux

La première description de cette malformation est attribuée à Matthew Baillie (1761-1823) à partir de l'examen anatomique du cœur d'un nourrisson âgé de deux mois (en 1793) (Prichard 1979). La transposition des gros vaisseaux (transposition of the great arteries ; TGA) représente 5 à 7% des malformations cardiaques congénitales (Samánek et al. 1989), correspondant à une incidence d'environ 20-30 pour 100.000 naissances vivantes. Il existe une prédominance chez les hommes avec un ratio homme/femme qui varie de 1,5 :1 à 3,2 :1 (Samánek 1994; Bianca and Ettore 2001). Environ 10% des transpositions des gros vaisseaux sont associées à une malformation extra cardiaque (Güçer et al. 2005).

Dans 50% des cas, la transposition des gros vaisseaux est isolée et se présente sous la forme d'une transposition simple. A l'inverse, les transpositions complexes associent une transposition à une autre anomalie cardiaque telle qu'une communication inter-ventriculaire (VSD), une anomalie d'arc aortique, etc.

La transposition des gros vaisseaux est caractérisée par une concordance auriculo-ventriculaire et une discordance ventriculo-artérielle (dextro-transposition des gros vaisseaux ou transposition des gros vaisseaux complète). Ainsi, l'oreillette droite est connectée au ventricule droit lui-même connecté à l'aorte ; et l'oreillette gauche est connectée au ventricule gauche duquel émerge l'artère pulmonaire. En raison de cette discordance, l'aorte transporte du sang pauvre en oxygène du ventricule droit vers les organes périphériques tandis que l'artère pulmonaire transporte du sang riche en oxygène vers les poumons. Pour permettre une survie du nouveau né, il doit y avoir une communication entre les deux circulations, pulmonaire et systémique, à travers le ductus arteriosus ou le foramen oval (Figure 14) lesquels doivent rester ouverts à la naissance permettant ainsi au sang riche en oxygène d'arriver au reste du corps (Ashworth et al. 2006).

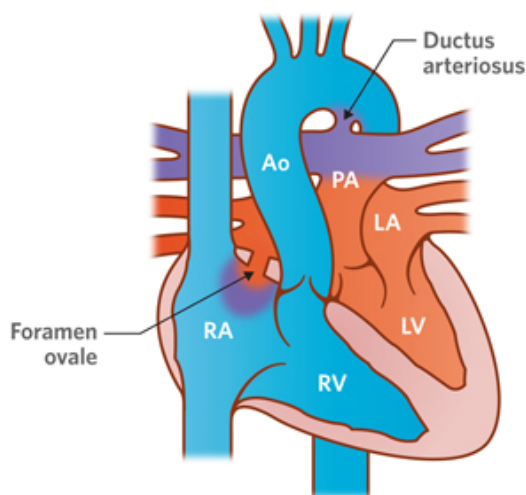


Figure 14. Représentation schématique de la transposition des gros vaisseaux. Ao : aorte, PA : artère pulmonaire, RA : oreillette droite, RV : ventricule droit, LA : oreillette gauche, LV : ventricule gauche. D'après http://www.rch.org.au/cardiology/heart_defects/Transposition_of_the_Great_Arteries/.

IV.4 Tétralogie de Fallot

La première description de la maladie a été faite en 1671 par Niels Stenson (Apitz et al. 2009) mais son nom a été associé à Étienne-Louis Arthur Fallot qui publie sa propre description en 1888. Le terme « tétralogie de Fallot » n'est cependant utilisé qu'à partir de 1924 (Apitz et al. 2009). La tétralogie de Fallot représenterait environ 10 % des malformations cardiaques congénitales (Huehnergath et al. 2008), correspondant à une incidence de 4 pour 10 000 naissances vivantes, ce qui en fait la cardiopathie congénitale cyanogène la plus fréquente chez l'enfant. Il n'existe pas de prédominance de sexe (Apitz et al. 2009).

La tétralogie de Fallot, comme son nom l'indique, comprend quatre éléments (Figure 15):

1. La sténose pulmonaire ou rétrécissement de la voie d'éjection du ventricule droit.

2. La présence d'une communication entre le cœur droit et le cœur gauche à travers le septum inter-ventriculaire (communication inter-ventriculaire, CIV).

3. L'hypertrophie du ventricule droit

4. La dextroposition de l'aorte, c'est-à-dire son déplacement partiel au-dessus du ventricule droit.

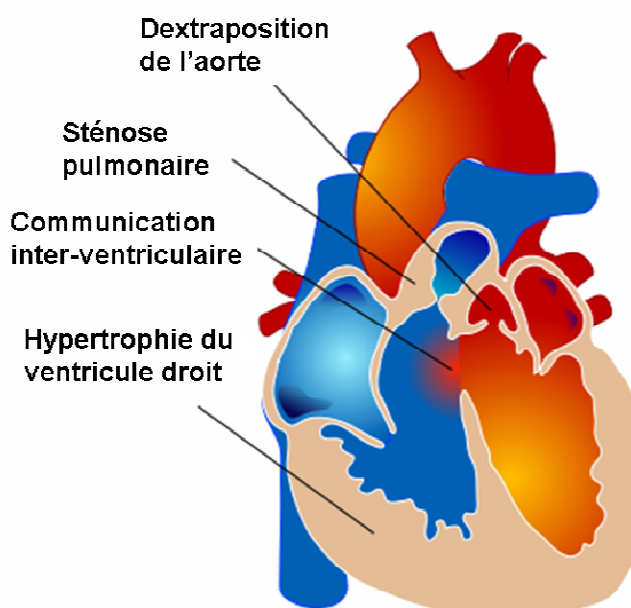


Figure 15. Représentation schématique de la tétralogie de Fallot. Modifié d'après http://fr.wikipedia.org/wiki/T%C3%A9tralogie_de_Fallot#mediaviewer/Fichier:Tetralogy_of_Fallot.svg

L'association d'un rétrécissement de l'artère pulmonaire et d'une communication inter-ventriculaire provoque une augmentation de la pression dans le ventricule droit, qui devient égale ou supérieure à celle du ventricule gauche. L'obstruction sévère vers la voie d'éjection du ventricule droit produit un shunt droite gauche avec un faible débit sanguin pulmonaire et une cyanose sévère, nécessitant une intervention chirurgicale à cœur ouvert, soit peu après la naissance ou plus tard dans l'enfance (Sommer et al. 2008). Le sang éjecté dans l'aorte par le ventricule gauche contient en partie du sang oxygéné et en partie du sang pauvre en oxygène provenant du cœur droit. Ce sang

appauvri en oxygène explique la cyanose et la couleur bleue typique des nouveaux nés atteints de tétralogie de Fallot.

V. Génétique du développement cardiaque

V.1 Stades précoces de la différenciation cardiaque

La formation du cœur chez les vertébrés supérieurs est un processus complexe qui commence aux stades précoces de l'embryogenèse, avant la fin de la gastrulation, avec la participation des cellules du mésoderme splanchnique antérieur à l'intérieur duquel les cellules précardiaques vont se différencier en cellules cardiaques en réponse à des signaux d'induction venant de l'endoderme. Ces cellules vont migrer vers la partie céphalique de l'embryon pour former la plaque cardiogénique, en forme de croissant. La découverte du second champ cardiaque a été une avancée majeure dans la compréhension du développement cardiaque (Buckingham et al. 2005). Le SHF comprend le croissant cardiaque et génère une grande portion du cœur, incluant la voie afférente, le ventricule droit et la plus grande partie des oreillettes. Ces données contribuent à expliquer pourquoi des mutations associées à certaines malformations cardiaques peuvent, en affectant des lignées cellulaires spécifiques uniquement au sein du SHF, entraîner des défauts dans des structures cardiaques spécifiques.

Certains des signaux provenant du mésoderme sous jacent participent à la différenciation en cellules cardiaques. Ces signaux comprennent des protéines morphogéniques (BMPs), des facteurs de croissance des fibroblastes (FGFs) et des protéines Wnt. Une attention toute particulière a été portée aux facteurs de transcription de la famille NK. Une mutation homozygote du gène *tinman* de la Drosophile, un facteur de transcription de la famille NK, est responsable d'une absence complète de développement cardiaque (Harvey 1996). Des études postérieures ont montré que des mutations du gène *NKX2.5*, homologue du gène *tinman* chez l'homme, sont la cause de certaines malformations cardiaques congénitales non syndromiques (Schott et al. 1998; Elliott et al. 2003; McElhinney et al. 2003; Gioli-Pereira et al. 2010). Le gène *GATA4* est également essentiel aux premiers stades de la différenciation cardiaque. Un facteur de

transcription GATA est défini par la présence d'un motif en doigt de zinc capable de reconnaître une séquence d'ADN « GATA » (Ko and Engel 1993). Les gènes de la drosophile *pannier* et *tinman*, homologues des gènes des familles GATA et NK chez l'Homme, sont capables, en synergie, d'induire une différenciation des cellules non-myocardiques en cardiomyocytes (Gajewski et al. 2001).

Parmi les nombreux autres gènes qui font actuellement l'objet d'études, deux sont également importants aux premiers stades de la formation cardiaque : SRF (*Serum Response Factor*) et le facteur de transcription myocardin. L'expression de SRF n'est pas spécifiquement cardiaque, mais sa présence est nécessaire à la formation du mésoderme permettant une différenciation cardiaque (Arsenian et al. 1998; Niu et al. 2005). De nombreux gènes d'expression cardiaque possèdent des sites de fixation potentiels à SRF. L'action de SRF est augmentée de plus de 1000 fois par la protéine myocardin, dont l'expression est essentiellement cardiaque (Wang et al. 2001; Chen et al. 2008). L'expression d'un mutant dominant négatif de myocardin dans des embryons de *Xenope* interfère avec la différenciation des cellules myocardiques laissant penser que myocardin joue un rôle important dans le développement cardiaque.

V.2 Fusion des tubes cardiaques primitifs

Le facteur de transcription GATA4 est important pour la fusion des tubes cardiaques primitifs (Kuo et al. 1997; Molkentin et al. 1997). Les gènes *GATA1*, 2 et 3 sont importants dans le développement du système hématopoïétique (Orkin 1992; Weiss and Orkin 1995), tandis que *GATA4*, 5 et 6 sont exprimés dans le cœur (Laverriere et al. 1994; Charron and Nemer 1999). L'inactivation de *GATA4* chez la souris est responsable d'une mort embryonnaire précoce. Les tubes cardiaques primitifs sont présents, mais leur fusion ne se fait pas (Kuo et al. 1997; Molkentin et al. 1997).

Des sites de fixation potentiels de GATA sont retrouvés dans les régions promotrices de nombreux gènes dont l'expression est cardiaque. Ainsi *GATA4*, en synergie avec *NKX2.5* et SRF sont capables d'activer des gènes dont l'expression est cardiaque (Durocher et al. 1997; Sepulveda et al. 1998). Des facteurs associés à

GATA4, dont FOG2 (*Friend of GATA*) et NFAT3 (*nuclear factor of activated T cells 3*) ont été identifiés comme partenaires potentiels d'hétérodimérisation de GATA4 (Molkentin et al. 1998).

V.3 Formation de la boucle cardiaque

Une légère perturbation à cette étape critique du développement cardiaque peut être responsable d'une malformation cardiaque congénitale sévère et explique la probable implication de nombreux gènes. Par ailleurs, l'hétérogénéité phénotypique observée au sein de familles avec mutation de *NKX2.5* illustre le rôle essentiel des gènes modificateurs, des facteurs environnementaux, des polymorphismes génétiques ou d'événements stochastiques dans la sévérité et le type de malformation. L'inactivation de *Nkx2.5* chez la souris perturbe la formation de la boucle cardiaque donnant lieu à une voie efférente et un ventricule droit d'une taille plus petite que la normale (Prall et al. 2007). *Nkx2.5* est orthologue au gène *tinman* de la mouche *Drosophila melanogaster* (Bodmer 1993). Chez la mouche, *tinman* est nécessaire à la formation de la structure qui correspond au cœur chez les mammifères (Harvey 1996). De nombreux gènes de la famille NKX ont des profils d'expression communs chez les mammifères, mais *Nkx2.5* semble jouer un rôle unique dans le développement cardiaque. Ce gène est exprimé dès les premiers stades de la différenciation du cardiomyocyte. Le tube cardiaque primitif se forme normalement chez les embryons de souris mutés, mais ceux-ci meurent en milieu de gestation en raison d'une mauvaise formation de la boucle cardiaque (Lyons et al. 1995).

V.4 Formation des cavités cardiaques

Pendant la formation du cœur, différents gènes ont été associés à une étape de développement déterminée (Figure 16). A mesure que le développement avance, le tube cardiaque primitif est soumis à une série d'événements complexes, incluant la spécification des cavités cardiaques, la septation et la trabéculatation, afin de former le cœur définitif. Les quatre cavités cardiaques sont morphologiquement différentes à des

stades très précoces du développement. Ces différences ne sont pas seulement liées à des facteurs hémodynamiques mais également à des expressions géniques particulières. De nombreux facteurs de transcription ont un profil d'expression spécifique d'une cavité cardiaque et participent au développement de chacune d'elles.

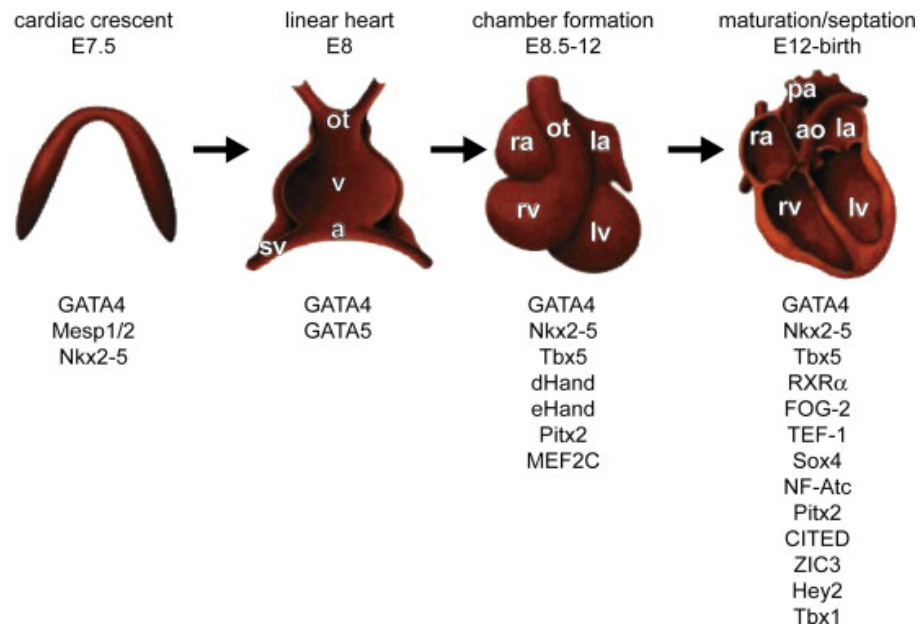


Figure 16. Représentation schématique des principaux stades du développement du cœur. Les jours de développement indiqués correspondent à des embryons de souris. A l'exception de GATA-5 qui a été étudié chez le Zebrafish, le rôle des Facteurs de Transcription a été principalement étudié chez la Souris. A : oreillette, ao : aorte, la : oreillette gauche, lv : ventricule gauche, ot : voie afférente, pa : artère pulmonaire / troncus, ra : oreillette droite, rv : ventricule droit, sv : sinus venosus, v : ventricule. D'après (Nemer 2008).

Bien que les mécanismes sous-jacents à la morphogenèse cardiaque restent encore en grand partie inconnus, la production de souris avec des gènes spécifiques inactivés a permis l'identification de protéines impliquées dans la maturation des cavités et dans la morphogenèse cardiaque (Moens et al. 1993; Chen et al. 1994; Sucov et al. 1994; Lyons et al. 1995; Conway et al. 1997; Lin et al. 1997; Molkentin et al. 1997; Srivastava et al. 1997; de la Pompa et al. 1998; Pereira et al. 1999; Reiter et al. 1999; Galvin et al. 2000; Tevosian et al. 2000; Zhong et al. 2000; Bruneau et al. 2001a; Merscher et al. 2001; Bruneau et al. 2001b).

Les facteurs de transcription dHand (Hand2) et eHand (Hand1) possédant un motif helix-loop-helix sont nettement plus exprimés dans les ventricules droit et gauche, respectivement, au moment du développement embryonnaire de la souris (Srivastava et

al. 1995). Des embryons dont le gène *dHand* est inactivé présentent un ventricule droit peu développé, alors que le tissu est correctement différencié (Srivastava et al. 1997). L'inactivation du gène *Hand2* chez la souris est associée à une létalité embryonnaire au moment du looping et le segment du tube cardiaque destiné à former le ventricule droit est absent (Srivastava et al. 1997). Le facteur de transcription Mef2 est essentiel à une bonne croissance des deux ventricules mais pas des oreillettes (Lin et al. 1997). L'inactivation du gène *Tbx5* chez la souris conduit à une hypoplasie des oreillettes et du ventricule gauche avec un développement normal du ventricule droit et des gros vaisseaux (Bruneau et al. 2001b). Le développement des ventricules est également sous le contrôle du gène à homéoboîte *Irx4* dont l'expression est principalement ventriculaire (Bao et al. 1999). Des patients avec hypoplasie du ventricule droit et/ou gauche pourraient présenter des mutations au sein de ces gènes. Enfin, *Tbx5* contrôle l'expression de gènes dans le canal atrioventriculaire, mais est contrebalancé par une autre protéine de la famille T-Box, *Tbx2* (Habets et al. 2002). Les mécanismes par lesquels ces facteurs de transcription affectent différentes structures du cœur ne sont pas clairement identifiés, mais impliquent probablement l'activation ou la répression de gènes cibles importants dans le développement spécifique des cavités cardiaques.

Les défauts de compartimentation et de communication entre les quatre cavités cardiaques provoquent la majorité des malformations cardiaques congénitales (Table 6).

Table 6. Mutations / délétions dans des facteurs de transcription associés au développement cardiaque chez la souris.

TF	Phenotype
GATA-4	Cardia bifida
GATA-5	Cardia bifida
Nkx2.5	Septation and differentiation defects
MEF2C	Right ventricle hypoplasia, septal defects
Tbx5	Atrial hypoplasia, septal defects
dHand	Right ventricle hypoplasia
TEF-1	Myocardial thinning
N-myc	Myocardial thinning
COUP-TFII	Atrial hypoplasia
RXR/RAR	Septal defects
Fog2	Outflow tract defects
Tbx1	Septation defects
Pax-3	Septation defects
NFATc	Defective valve development
Hrt2	Aortic coarctation
Smad6	Valve thickening
Irx4	Postnatal cardiomyopathy

Modifiée d'après (Nemer 2008).

Le rôle de certains facteurs de transcription pendant le développement cardiaque a été bien décrit, mais on en connaît peu sur le rôle des facteurs qui modifient la structure de la chromatine dont le degré de compaction peut restreindre ou à l'inverse faciliter l'activation génique. BAF60C (SMARCD3), une sous-unité du complexe BAF de remodelage de la chromatine, interagit avec divers facteurs de transcription. La perte de fonction de BAF60C a été associée à des malformations sévères pendant la morphogenèse cardiaque entravant l'activation de certains gènes à expression cardiaque (Lickert et al. 2004). Une réduction partielle des niveaux d'expression de BAF60C est liée à des malformations conotruncales (Lickert et al. 2004), ce qui suggère qu'un bon dosage de complexes régulateurs de la chromatine est essentiel pour le développement normal du cœur. Tandis que des mutations au sein de gènes participant aux complexes BAF altèrent la structure de la chromatine, d'autres complexes régulateurs de la chromatine modifient les histones qui sont aussi très importantes pour la formation du cœur. L'histone SMYD1 (BOP) est un régulateur essentiel de la croissance et de la différenciation des cavités cardiaques (Gottlieb et al. 2002). Un travail récent montre que les mutations de gènes modificateurs d'histones joueraient un rôle dans 10% des malformations cardiaques congénitales sévères (Zaidi et al. 2013).

Au cours des dernières années, de nombreux micro-ARNs (miARN) non codants ont été identifiés. Ces miARN s'accrochent à la région 3' UTR de l'ARN messager (ARNm) transcrit à partir d'un gène donné ce qui provoque une dégradation de cet ARNm et ainsi une répression de la traduction du gène donné, (Ambros 2004). Une partie de ces miARNs ont une fonction dans le cœur (van Rooij and Olson 2007). Le miARN qui est potentiellement le plus pertinent pour les malformations cardiaques congénitales est le miR-1, étant donné son importance au cours du développement embryonnaire cardiaque (Zhao et al. 2005; Zhao et al. 2007). miR-1 est exprimé dans le cœur en développement et les études de surexpression génique suggèrent que les gènes régulés par miR-1 sont impliqués dans la prolifération des cardiomyocytes et le programme de développement régulé par des facteurs de transcription cardiaques (Zhao et al. 2005). Il a été démontré que miR-1 régule le facteur de transcription cardiaque HAND2, lequel est impliqué dans la croissance du cœur embryonnaire. Ainsi, une dérégulation de miR-1 ou d'autres miARN importants dans le développement cardiaque pourrait contribuer à certaines malformations cardiaques congénitales chez l'homme.

VI. Génétique des malformations cardiaques congénitales non syndromiques

Des études génétiques chez des modèles animaux et chez l'Homme ont permis d'identifier un grand nombre de gènes associés à des malformations cardiaques congénitales héritées ou sporadiques. La connaissance des bases moléculaires du développement normal du cœur est indispensable pour améliorer la compréhension de la pathophysiologie des MCC. La majorité de ces gènes codent des facteurs de transcription (Table 4) qui régulent des événements spécifiques au cours du développement du cœur, telles que la septation ventriculaire ou la morphogenèse de la voie efférente.

La découverte de ces facteurs de transcription aux interactions complexes (Temsah and Nemer 2005) peut expliquer comment un même type de malformation cardiaque congénitale peut être liée à différents gènes (Bruneau et al. 2001b; Garg et al. 2003; Nemer et al. 2006). Par exemple, une mutation au sein de *TBX5*, *GATA4* ou *NKX2.5* peut conduire à des défauts similaires du septum cardiaque

Les dernières avancées dans la compréhension des mécanismes moléculaires de la formation du cœur ont permis l'identification de divers loci liés à des malformations cardiaques congénitales. Deux découvertes intéressantes issues de ces études génétiques sont :

1. Une malformation cardiaque congénitale est souvent liée à plus d'un locus (Table 7). .
2. L'expressivité variable du phénotype est souvent observée même dans les familles où les malformations cardiaques congénitales sont liées à une mutation génétique spécifique.

La majorité des MCC dont la transmission est monogénique sont associées à des anomalies extra-cardiaques et constituent les formes syndromiques décrites précédemment. A l'inverse, la plupart des malformations cardiaques congénitales non-

syndromiques sont sporadiques. Le fait que les formes familiales non syndromiques soient rarissimes limite l'identification de nouveaux gènes humains impliqués dans ces maladies par des méthodes classiques de génétique. Cependant, ces rares familles ont conduit à l'identification de gènes impliqués dans ces MCC.

La majorité des gènes impliqués dans des formes non-syndromiques de ces maladies sont des régulateurs transcriptionnels de la morphogenèse cardiaque. (Table 7).

Table 7. Gènes impliqués dans des formes non-syndromiques de malformations cardiaques congénitales.

CHM	BAV/AS	VSD	ASD	PS	TOF	AVSD	TGA	HLHS	DORV	Heterotaxy
Incidence ^a	14 / 0.8	4	1.0	0.7	0.4	0.3	0.2	0.2	0.16	0.1
Genes	<i>NOTCH1</i>	<i>NKX2.5^b</i> <i>GATA4</i> <i>TBX20</i> <i>TBX1</i> <i>ACTC1</i> <i>MYBPC3</i> <i>TBX5</i> <i>CITED2</i>	<i>NKX2.5^b</i> <i>GATA4</i> <i>TBX20</i> <i>MYH6</i> <i>ACTC1</i> <i>MYBPC3</i> <i>MYH7</i> <i>MYBPC3</i> <i>TBX5^b</i> <i>CITED2</i> <i>TLL1</i>	<i>GATA4</i> <i>MYOCD</i> <i>JAG1</i> <i>GATA6</i>	<i>NKX2.5</i> <i>GATA4</i> <i>TBX1</i> <i>FOG2</i> <i>CFC1</i> <i>NOTCH1</i> <i>GDF1</i> <i>HEY2^c</i> <i>TDGF1</i> <i>JAG1</i> <i>FOXH1</i>	<i>GATA4</i> <i>CRELD1</i> <i>CFC1</i> <i>GDF1</i> <i>NKX2.5^c</i> <i>TBX5^c</i> <i>HEY2^c</i> <i>PTPN11</i> <i>ALK2</i>	<i>NKX2.5</i> <i>THRAP2</i> <i>ZIC3</i> <i>CFC1</i> <i>GDF1</i> <i>THRAP2</i> <i>FOXH1</i>	<i>NKX2.5</i> <i>NOTCH1</i> <i>HAND1^c</i> <i>GJA1^c</i>	<i>NKX2.5</i> <i>GDF1</i> <i>CFC1</i>	<i>ZIC3</i> <i>GDF1</i> <i>CFC1</i> <i>LEFTY2</i> <i>ACVR2B</i> <i>NODAL</i> <i>NKX2.5</i> <i>CRELD1</i> <i>FOXH1</i>

AS : sténose de la valve aortique, ASD : communication inter-auriculaire, AVSD : canal atrio-ventriculaire, BAV : valve aortique bicuspidale, CHM : malformation cardiaque congénitale, DORV : ventricule droit à double issue, HLHS : hypoplasie du cœur gauche, PS : sténose pulmonaire, TGA : transposition des gros vaisseaux, TOF : tétralogie de Fallot, VSD : communication inter-ventriculaire. a : par 1000 naissances vivantes, b : mutations de la ligne germinale et somatiques, c : mutations somatiques. D'après (Wessels and Willems 2010).

VII. Le rôle de la variabilité du nombre de copies (CNVs) d'ADN génomique

La variabilité du nombre de copies (CNV) d'ADN génomique désigne une forme particulière de variation génétique pour laquelle le nombre de copies d'un même gène ou d'un segment chromosomique est variable. Le nombre de copies peut être variable d'une personne à une autre au sein de la population générale, il s'agit d'un polymorphisme. Cette forme de polymorphisme, découverte récemment, constitue une part importante de

la variabilité individuelle au sein des espèces. Ainsi, plus de 12 % du génome humain serait concerné par le polymorphisme lié aux CNV (Redon et al. 2006). A l'inverse, un CNV peut être retrouvé chez un patient et ne pas être présent chez les individus de la population générale et expliquer la pathologie présentée par le patient. La variabilité du nombre de copies d'un gène est souvent secondaire à une recombinaison méiotique inégal entraînant des événements de duplication et délétion d'une région pouvant impliquer un ou plusieurs gènes.

Le nombre de copies d'un gène peut être soumis à une sélection. En effet, le nombre de copies modifie la quantité de protéines produites dans une cellule. L'augmentation ou la diminution du nombre de copie peut donc être sélectionnée au sein d'une espèce ou d'une population en fonction de contraintes environnementales. Ainsi, on observe une augmentation du nombre moyen de copies de *AMY1* (gène codant la protéine AMY1, l'amylase, une enzyme salivaire responsable de la dégradation de l'amidon) dans les populations humaines ayant un régime plus riche en amidon. Il semblerait donc que cette augmentation ait été le fruit de la sélection (Perry et al. 2007). Il est alors raisonnable de penser que les variations en nombre de copies d'autres gènes sont ou ont été soumis à de fortes pressions de sélection naturelle.

CNVs et les malformations cardiaques congénitales

Des CNVs de grande taille ont été associées à des malformations cardiaques congénitales, depuis longtemps pour certains (Chelly et al. 1986; Scambler et al. 1991; Sparkes et al. 2005). Un CNV particulièrement bien étudié est la microdélétion 22q11.2 comprenant le gène *TBX1* et responsable du syndrome velocardiofacial ou de DiGeorge. Un autre exemple de CNV récurrent responsable d'un syndrome associé à une malformation cardiaque congénitale est la microdélétion 7q11.23 qui comprend le gène *ELN* (syndrome de Williams-Beuren).

Un grand nombre de CNVs rares ont également été associés à des formes non syndromiques de malformations cardiaques congénitales, dont certaines microdélétions 1q21.1 (Soemedi et al. 2012b) comprenant le gène *GJA5*. Les phénotypes cardiaques associés à ces CNVs sont très variables, ce qui est également observé pour des

mutations ponctuelles au sein de gènes responsables de malformations cardiaques congénitales.

Outre l'identification de CNV rares et récurrents, une incidence globale significative des CNVs rares et non récurrents a été montrée dans les cas de malformations cardiaques congénitales (Goldmuntz et al. 2011; Soemedi et al. 2012b). De multiples CNVs rares contribuent à l'hétérotaxie (Fakhro et al., 2011), la tétralogie de Fallot (Greenway et al. 2009), et la canal atrio-ventriculaire (Priest et al. 2012). Plusieurs études ont montré que les CNVs rares expliquent 5 à 10% des différentes malformations cardiaques congénitales non syndromiques (Breckpot et al., 2011, 2010; Erdogan et al., 2008; Greenway et al., 2009; Hitz et al., 2012; Lalani et al., 2013; Silversides et al., 2012; Soemedi et al., 2012b).

Environ 10% de CNVs *de novo* ont été observés dans des cas sporadiques de tétralogie de Fallot (Greenway et al. 2009). L'incidence globale est estimée à environ 5% de CNVs *de novo* si l'on prend en compte différents types de malformations cardiaques congénitales (Yuan et al. 2013), ce qui correspond au double de ce qui est observé dans une étude portant sur l'autisme (une incidence de CNVs *de novo* a été observée dans environ 2% des cas) (Sanders et al. 2011).

VIII. Stratégies d'identification des CNVs

VIII.1 Hybridation In Situ en Fluorescence (FISH)

L'étude d'anomalies chromosomiques permet l'identification de gènes responsables de maladies ou de syndromes génétiques. Les translocations réciproques en apparence équilibrée peuvent interrompre un gène. Ce gène n'est alors plus fonctionnel et explique un phénotype anormal. Des méthodes de clonage positionnel basées sur l'utilisation de sondes d'ADN génomique marquées – cosmide, BAC (bacterial artificial chromosome), ou YAC (yeast artificial chromosome) – permettent la mise en évidence du point de cassure, lieu d'interruption du gène.

D'autres remaniements chromosomiques consistent en des délétions ou des duplications chromosomiques. Les délétions et les duplications peuvent emporter un ou plusieurs gènes. Ces anomalies permettent une première localisation chromosomique d'un gène responsable d'une maladie.

Les anomalies chromosomiques responsables d'un syndrome polymalformatif intégrant une malformation cardiaque sont très variées. Plusieurs études réalisées grâce aux données de la littérature ont tenté d'identifier les anomalies chromosomiques récurrentes associées à une malformation cardiaque (Johnson et al. 1997; van Karnebeek and Hennekam 1999). Ainsi, une cardiopathie est souvent associée à la microdélétion 22q11.2. L'observation d'enfants porteurs d'une délétion dans la région chromosomique 20p12 a conduit à l'identification du gène JAG1. De la même façon, le syndrome de « San Luis valley » décrit des enfants polymalformés avec délétion 8p terminale et duplication 8q terminale, résultat d'une inversion péricentrique parentale (Gelb et al. 1991). Par la suite, d'autres enfants présentant un tableau proche incluant une malformation cardiaque volontiers conotruncale, lié à une délétion 8p terminale, ont permis de suspecter qu'un gène important dans le développement cardiaque siégeait dans cette région chromosomique (Reddy 1999; Giglio et al. 2000). Plus récemment, des mutations du gène GATA4 ont été identifiées chez des patients de plusieurs familles confirmant cette suspicion (Garg et al. 2003). De nombreux gènes ayant un rôle dans le développement cardiaque restent cependant à identifier. Par exemple, la délétion 3p terminale est souvent associée à une malformation cardiaque, mais le ou les gènes impliqués ne sont pas encore identifiés.

VIII.2 Hybridation Génomique Comparative (CGH) array

L'identification de nouveaux gènes par FISH reste laborieuse. Elle impose de disposer de cellules techniques en cytogénétique et de tester un grand nombre de sondes marquées. Le temps et l'argent nécessaire en limitent l'intérêt. En revanche, la technique des puces à ADN (CGH array) constitue une meilleure alternative. Ces dernières années, des améliorations techniques décisives ont été réalisées. Le principe de la technique est le suivant (Figure 19) : sur une lame (puce), est déposé un grand nombre de sondes ou fragments d'ADN (clones) balisant l'ensemble du génome (par

exemple, à raison d'une sonde tous les 100 kilobases) ou permettant de tester des régions préférentiellement délétées ou dupliquées dans les syndromes polymalformatifs. Parallèlement, les ADN du patient à tester et d'un sujet contrôle sont marqués de deux fluorochromes de couleurs différentes. Une cohybridation de ces deux ADN est ensuite réalisée sur la puce. Cette technique permet de tester plusieurs centaines ou milliers de régions en un seul temps. Ensuite, un logiciel détermine le rapport de fluorescence pour chaque sonde, le gain ou la perte de matériel chromosomique se traduisant par une déviation de l'intensité de fluorescence de l'ADN du patient (Figure 20). Cette approche par puce à ADN permet un diagnostic rapide par la mise en évidence d'un remaniement de petite taille impliquant un ou plusieurs des multiples clones fixés sur la puce. L'étude d'enfants ou de fœtus avec plusieurs malformations dont une anomalie cardiaque peut contribuer à identifier des régions chromosomiques au sein desquelles siègent des gènes importants dans le développement du cœur.

Cette approche a été validée par plusieurs études préliminaires. L'une d'elles a consisté à rechercher par CGH array un remaniement subtélomérique préalablement identifié par la technique de FISH chez 20 patients avec retard mental associé ou non à des malformations (Veltman et al. 2002). La technique a permis de retrouver les anomalies connues mais également d'identifier 3 remaniements subtélomériques supplémentaires non détectés par FISH. Un autre travail a permis de caractériser la délétion 1p36. Vingt-cinq patients avec délétion 1p36 connue ont été testés grâce à une puce sur laquelle 97 clones de la région 1p36 ont été déposés (Yu et al. 2003). Les remaniements connus ont été confirmés ainsi que des remaniements plus complexes passés inaperçus, confirmant ainsi l'intérêt de cette approche.

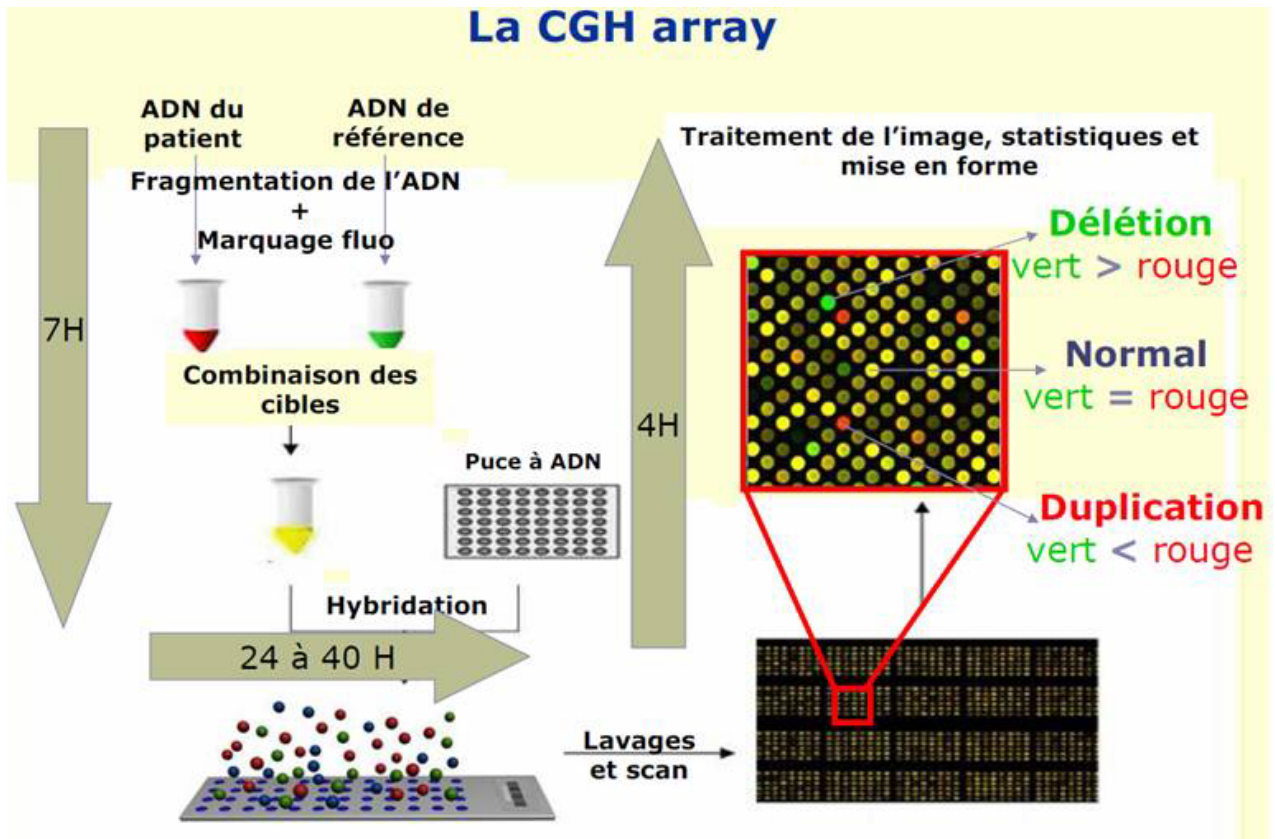


Figure 17. Principe de la technique de la CGH array. D'après <http://www.techniciens-cytogenetique.com/>

La technique de CGH array est utilisée actuellement en recherche et en diagnostic dans les domaines du cancer, des malformations congénitales, de la déficience intellectuelle, etc.

Certaines considérations et limitations techniques de la CGH array doivent être prises en compte. Cette technique ne détecte que les modifications chromosomiques déséquilibrées. Les anomalies chromosomiques structurales équilibrées telles que les translocations réciproques ou les inversions ne sont pas détectées. L'apport des puces d'hybridation génomique comparative est le niveau de résolution bien supérieur à celui obtenu avec les techniques de cytogénétiques classiques, jusqu'à 1000 fois plus résolutive qu'un caryotype haute résolution.

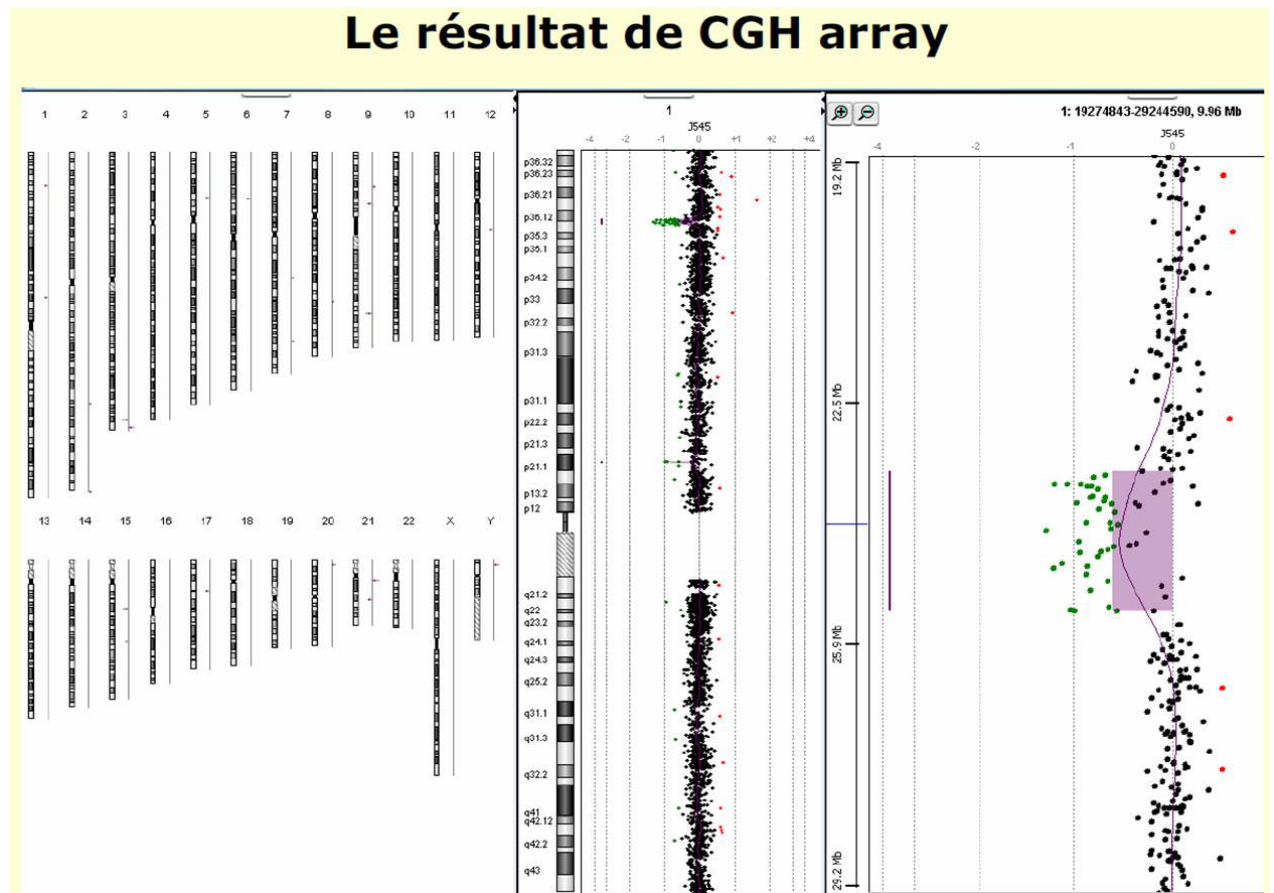


Figure 18. Le résultat de la CGH array. On peut observer une déviation de l'intensité de fluorescence de l'ADN du patient correspondant à un gain de matériel chromosomique. La taille et la localisation de cette duplication est obtenue avec une grande précision. Chaque point correspond à une sonde différente. D'après <http://www.techniciens-cytogenetique.com/>

OBJECTIFS

L'étude de patients présentant une malformation cardiaque congénitale sporadique constitue une alternative aux études familiales classiques d'identification de nouveaux gènes impliqués dans ces pathologies. L'approche familiale est limitée par le fait que les familles au sein desquelles ségrègent une malformation cardiaque sont extrêmement rares et également par la faible pénétrance de ces malformations au sein de ces familles. A l'inverse, l'étude de patients présentant une cardiopathie isolée et sporadique est une alternative intéressante permettant l'identification de nouveaux gènes impliqués dans ces pathologies. L'émergence de nouvelles technologies de génétique moléculaire permet d'envisager de nouvelles études portant sur ces cas sporadiques.

L'objectif principal de cette thèse sur la génétique des malformations cardiaques congénitales de la voie efférente est l'identification de nouveaux gènes à partir de patients présentant une forme non syndromique et sporadique de la maladie. Ces études permettent une meilleure compréhension des mécanismes moléculaires et embryologiques à l'origine de ces malformations, qui restent peu connus à ce jour. L'identification de nouveaux gènes permettra également une meilleure compréhension du développement cardiaque normal. D'une manière plus générale, l'objectif est d'améliorer le conseil génétique destiné aux couples dont l'un des membres est atteints d'une malformation cardiaque congénitale ou ayant eu un enfant atteint.

Dans l'avenir, l'identification de nouveaux gènes impliqués dans des formes non syndromiques des malformations cardiaques congénitales contribuera à améliorer le diagnostic génétique de ces pathologies.

PROJET 1

Rare Copy Number Variants in patients with Tetralogy of Fallot, Transposition of the Great Arteries, and Coarctation of the aorta

Marta Sanchez-Castro^{1,2}, et al.

Version préliminaire

¹INSERM, UMR1087, l'institut du thorax, Nantes, France

²Université de Nantes, Nantes, France

³CHU Nantes, Service de Cardiologie Pédiatrique, Nantes, France

⁴Centre de Pédiatrie Clocheville, Service de Pédiatrie A, Tours, France

⁵CHU Angers, Service de pédiatrie, Angers, France

⁶Centre de Référence Malformations Cardiaques Congénitales Complexes-M3C, Hôpital Necker-Enfants Malades, AP-HP, Paris, France

⁷CHU Nantes, Service de Génétique Médicale, Nantes, France

*Correspondence to: Cédric Le Caignec, MD, PhD, Service de Génétique Médicale, CHU, 9, quai Moncousu 44093 Nantes, France, Tél.: +33 2 40 08 42 84, Fax: +33 2 40 08 39 43, E-mail: cedric.lecaignec@chu-nantes.fr

Abstract

Congenital heart defects (CHD) are the leading malformations among all newborns, and one of the leading causes of morbidity and mortality in developed countries. Although genetic variation contributes to CHD, the genetic basis for CHD remains unknown in the majority of patients. We analysed by array comparative genomic hybridization (aCGH) 316 children with nonsyndromic CHD - 76 coarctation of the aorta (CoAo), 159 transposition of the great arteries (TGA) and 81 tetralogy of Fallot (ToF) - and their unaffected parents. We identified by aCGH and confirmed by qPCR 71 rare *de novo* or inherited copy number variants (CNV) (50 duplications and 21 deletions) in the affected children. No *de novo* CNV was identified in patients with TGA in contrast to CoAo and ToF ($p < 0.01$). Transcription Factor Binding Sites (TFBS) approach showed that 77% of the rare CNV contained at least one gene with FOXC1 binding sites and thus might have altered genes' expression regulated by FOXC1. FOXC1 belongs to the forkhead transcription factors family, which play a major role in the regulation of cardiovascular development and in cardiac outflow tract morphogenesis in particular. These data suggest that FOXC1 is a strong candidate for CHD.

Key Words

Congenital heart defects, transposition of the great arteries, coarctation of the aorta, tetralogy of fallot, outflow tract defect, CGH array, CNV, FOXC1

Introduction

Congenital heart defects (CHD) are the most common congenital malformations with an incidence of 0.5-1% of live births (Øyen et al., 2009), being the first cause of mortality during the first year of newborns' life in developed countries (Gillum, 1994; Petrini et al., 1997). Despite therapeutic advances, CHD are associated with a high proportion of long term patients morbidity (Botto and Correa, 2003; De Galan-Roosen et al., 1998). Among CHD, a large subset of CHD involves the outflow tract. This heterogeneous group of malformations represents 20-30% of CHD diagnosed in

newborns (Kodo et al., 2009). Transposition of the great arteries (TGA) accounts for 5-7% of all CHD (Digilio et al., 2001), and is one of the most common cyanotic disorder, (after the Tetralogy of Fallot (ToF)) and one of the most frequent CHD diagnosed in the neonatal period (after the Atrial Septal Defect (ASD) and the Ventricular Septal Defect (VSD)), with a prevalence of 0.2 per 1000 live births. The most common form of TGA is the dextro-looped type, which consists of a discordant ventriculo-arterial connection so that the aorta incorrectly arises anterior and right-sided from the right ventricle, whereas the pulmonary artery incorrectly arises posterior and left-sided from the left ventricle. In contrast to the normal heart in which both outflow tracts and great vessels show a dextral (right-handed) spiralisation in TGA the great vessels present a parallel course lacking normal spiralisation. Coarctation of the aorta (CoA) is an outflow tract defect whereby the aorta narrows in the area where the ductus arteriosus inserts. This is a relatively common defect that accounts for around 7% of all CHD (Rosenthal, 2005). Tetralogy of Fallot (ToF) is defined by a combination of malpositioned aorta that overrides both ventricles, ventricular septal defect, pulmonary stenosis obstructing blood flow into the lungs and right ventricular hypertrophy. ToF is the most common cyanotic congenital cardiac disease in humans with an occurrence of one per 3000 live births and accounts for 10% of all CHD (Ferencz et al., 1985).

Although most of the patients are successfully operated, the risk of cardiac malformation in their offspring is 10-fold higher than in the general population (Burn et al., 1998). Despite the high incidence of CHD, the etiology of these malformations remains largely unknown. A multifactorial origin associating environmental and genetic factors seems to be the usual mode of inheritance (Blue et al., 2012; Gillum, 1994; Nora, 1968; Petrini et al., 1997). The identification of new genes involved in non-syndromic forms of CHD will help to better understand the molecular mechanisms leading to these malformations, and to improve the genetic counseling and disease prevention of couples having an affected child as well as adult patients willing to reproduce. Currently, about 20% of CHD can be attributed to known causes such as chromosomal abnormalities, single gene disorders or teratogens. However, no aetiology is identified in about 80% of the patients (Blue et al., 2012). Genomic rearrangements are known to play a major role in CHD. For example, microdeletions and microduplications at 22q11.2 (Christiansen et al., 2004; Greenway et al., 2009), 9q31.1

(van Bon et al., 2008) and 1q21.1 (Tyreman et al., 2009) are frequently associated with outflow tract defects in humans.

A number of animal studies showed that CHD originates from the early steps of cardiogenesis during embryo development (Bruneau, 2008). Abnormal development of cardiac neural crest cells, an ectoderm-derived cell lineage, contributes significantly to the cardiac outflow tract defects. During early embryogenesis, cardiac neural crest cells arise from the dorsal neural tube and migrate ventrally as mesenchymal cells to populate the cardiac outflow tract, where they coalesce to form the aorticopulmonary septum, which divides the single truncus arteriosus into the aorta and pulmonary artery, resulting in the establishment of separate systemic and pulmonary circulations. Several mutated genes acting in different stages of embryo development have been associated to non syndromic forms of CHD in humans (Bruneau, 2008; Wessels and Willems, 2010).

The study of sporadic patients presenting with an isolated non-syndromic CHD is an alternative to the classic familial studies for the identification of new genes implicated in these pathologies. The array comparative genomic hybridization (aCGH) is a method able to detect genomic copy number variations (CNVs) (i.e. deletions and duplications) that can be used in sporadic patients and therefore able to pinpoint candidate genes. Few studies have evidenced rare CNVs in patients with CHD using the aCGH method (van Bon et al., 2008; Breckpot et al., 2010; Derwińska et al., 2012; Fakhro et al., 2011; Goldmuntz et al., 2011; Greenway et al., 2009; Hartman et al., 2011; Hitz et al., 2012; Lalani et al., 2013; Lu et al., 2008; Mefford et al., 2008; Serra-Juhé et al., 2012; Silversides et al., 2012; Soemedi et al., 2012; Thienpont et al., 2007; Warburton et al., 2014).

We studied by aCGH 316 children with nonsyndromic CHD and their normal parents. Our data showed a high contribution of rare inherited and also *de novo* CNVs to human CHD.

Methods

Patients

Informed consent for genetic analyses was obtained from all individuals participating in the study. The protocol was approved by the ethics committee of the University Hospital Center of Nantes (BRD 09/3A). Children from 516 families presenting with a sporadic and non syndromic outflow tract (OFT) defect were referred to the university hospital of Nantes, the hospital Necker Enfants Malades of Paris, and the university hospital of Tours between July 2008 and august 2011. Patients with extra-cardiac features, such as learning disability, brain or renal anomalies, or carrying a microdeletion/microduplication syndrome, or a monogenic syndrome were excluded from the study. The child for whom one of the parents was known to present a CHD was excluded from the study.

The phenotype of the patients was classified as following: transposition of the great arteries (TGA) in 291 cases, coarctation of the aorta (CoA) in 85 cases and tetralogy of Fallot (ToF) in 92 cases. We pooled other OFT defects in a single group: pulmonary atresia with intact ventricular septum (16 cases), truncus arteriosus (4 cases), interrupted aortic arch (3 cases), double outlet right ventricle (2 cases), aneurismal dilatation of the ascending aorta (1 case) and pulmonary stenosis (1case). Twenty-one cases were unclassified.

DNA extraction

DNA from all probands and their parents was extracted from whole peripheral blood using NucleoSpin® Blood XL (Macherey Nagel), illustra™ DNA Extraction Kit BACC2 (GE Healthcare) or UltraPure™ Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) (Life Technologies™) according to manufacturer's instructions.

Array comparative genomic hybridization (aCGH)

316 family trios (76 CoAo, 159 TGA and 81 ToF) were selected for aCGH. aCGH experiments were performed using 2*400K Agilent custom-designed arrays (024825_D_F_20090731). Microarrays contained 300,000 probes covering >95% of exonic regions. In addition, 100,000 probes covered with a high density 297 candidate genes known to play a role in heart development in humans or in different animal models (Supplementary table 1). To reduce the risk for detecting unsolicited findings (i.e. the detection of genomic imbalances unrelated to CHD), more than 800 genes responsible for dominant or recessive genetic diseases were excluded from the design of the array. Digestion, labelling, and hybridization were performed following the protocols provided by Agilent. The DNA of the child was hybridized twice: once with the DNA of the father and once with the DNA of the mother (Figure 1).

CNV detection workflow: bioinformatics and statistical analysis

The arrays were analyzed with the Agilent scanner and the Feature Extraction software (v.9.1.3). Graphical overview was obtained using the customized SigFrame software. All genomic coordinates are based on February 2009 assembly of the reference genome (GRCh37/hg19).

Database of Genomic Variants Archive (DGVa)

All raw data and analyzed data were submitted to DGVa public database <http://www.ebi.ac.uk/dgva/>.

Copy number variant (CNV) detection and validation

CNVs were selected according to the following criteria: 1) CNVs >2 kb (except CNV containing at least one complete exon). 2) CNVs absent from the CNV consensus

reference set version 2.1 (42 million probes study (Conrad et al., 2010), WTCCC study <http://www.wtccc.org.uk/>, 1000 genomes project <http://www.1000genomes.org/> and DDD controls project <http://www.ddduk.org/>) integrating a number of high quality copy number based studies, 3) or uncommon (< 4 times in the Database of Genomic Variants (Iafrate et al., 2004)). 4) Only CNVs containing at least one exon were included. 5) CNVs corresponding to syndromic regions were discarded.

QPCR validation

The CNVs that fulfilled the selection criteria were subjected to validation by qPCR. At least one couple of primers was designed inside of each rare CNV regions (Supplementary table 2). qPCR reactions were carried out in a 10 µL reaction volume containing 5 ng of genomic DNA, 0.5 µM each primer and 5 µL of SYBR®Premix Ex Taq™ (Takara Bio Inc., Shiga, Japan). qPCR conditions were: Initial denaturation at 95°C for 4min, followed by 40 cycles at 95 °C for 15 sec, 60°C for 10sec and 72 °C for 10sec. qPCR reactions were carried out in a LightCycler®480 System (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany). The plate used was a LightCycler®480 white Multiwell Plate 96 (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany). Amplification products were analyzed using LightCycler®480 software version 1.5.0 (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany). $\Delta\Delta C_t$ method was used as described by (Livak and Schmittgen, 2001) to quantify the number of DNA copies. $\Delta\Delta C_t$ were calculated with the *ALB* and *TNNI3K* genes for normalization.

Chi-squared statistical test

A statistical analysis was performed using the *Chi-squared test* to compare the number of *de novo* and inherited CNV present in the three groups of CHD patients (CoA, TGA and ToF). Results were considered statistically significant when $P < 0.05$.

Idiographica

Idiographica version 2.2 (July 2013) (<http://www.ncrna.org/idiographica/>) web server was used to create our own ideogram including all the validated CNV regions found in our CHD affected cohort.

DECIPHER and ISCA database

Validated CNVs were compared to published CNV data from DECIPHER (<http://decipher.sanger.ac.uk/>) and ISCA (www.iscaconsortium.org) databases collecting clinical information about chromosomal microdeletions/duplications/insertions, translocations and inversions.

Tools for gene prioritization analysis

Knime4Bio

Validated CNVs were analyzed by Knime4Bio, a set of custom nodes for the interpretation of large datasets (Lindenbaum et al., 2011) (figure 2a). This tool allowed us to quickly find which genes included in our rare CNV regions were implicated in heart development (GO:0007507) according to GeneOntology database (<http://www.geneontology.org/>).

DAPPLE

DAPPLE (<http://www.broadinstitute.org/mpg/dapple/dapple.php>) stands for Disease Association Protein-Protein Link Evaluator (Rossin et al., 2011). DAPPLE takes as input a list of seed genomic regions that are then converted into genes based on overlap. Afterwards it looks for significant physical connectivity among proteins encoded by genes in loci associated to disease according to protein-protein interactions reported

in the literature. The hypothesis behind DAPPLE is that causal genetic variation affects a limited set of underlying mechanisms that are detectable by protein-protein interactions.

GRAIL

GRAIL (<http://www.broadinstitute.org/mpg/grail/>) is a tool to examine relationships between genes in different disease associated loci (Raychaudhuri et al., 2009). Given several genomic regions associated with a particular phenotype or disease, GRAIL looks for similarities in the published scientific text among the associated genes. As input, we uploaded genomic regions considered as rare copy number variants.

Genomic regions were introduced in DAPPLE and GRAIL according to patient's disease (CoA, TGA and ToF).

TFBS

HMR Conserved Transcription Factor Binding Sites (TFBS) from UCSC (<http://genome.ucsc.edu/>). This track containing the location and score of transcription factor binding sites conserved in the human/mouse/rat alignment was used to find the predicted transcription factors binding the genes included in our selected CNV regions. The score and threshold are computed with the Transfac Matrix Database (v7.0) created by Biobase (<http://www.biobase-international.com/>).

Results

CNV inventory and disease association

316 family trios (affected child and normal parents) were analysed by aCGH and led to the identification to 152 rare CNV that fulfilled the selection criteria (38 in trios of the CoA group, 72 in trios of the TGA group and 42 in trios of the ToF group). qPCR analysis performed in the children and their parents confirmed that 71 CNV (50

duplications and 21 deletions) were present in the affected children (15 for CoA, 29 for TGA and 27 for ToF), 78 were present in one of the parents but showed that 3 CNV were aCGH false positive results (table 1) (table 2) (Supplementary table 3). The size of the CNVs ranged from 2.9 kb to 1,845.3 kb except for two large CNVs of 7,962.9 kb and 14,587.4 kb. Among the 71 CNVs found in the affected child, 63 were inherited from one of the parents while 8/71 CNVs were *de novo*: 3 (1 duplication + 2 deletions) for CoA (4,1%) and 5 (2 duplications + 3 deletions) for ToF (6,6%) (table 1). No *de novo* CNV was present in patients with TGA. The difference between the absence of *de novo* CNV in the patients with TGA compared to patients with CoA and ToF was statistically significant (*p=0, 01) (table 1).

CNV genomic distribution

Ideographica results show a spread distribution of the 71 rare CNV regions present in our CHD affected cohort all over the genome (figure 3). CNV overlapping has been observed only in three genomic regions: at 10q24.32, 11p11.2 and 20p11.23. An overlapping duplication in 10q24.32 was identified in two unrelated children with TGA. Both CNVs were inherited and included a portion of the *BTRC* gene (figure 4a). The second overlapping region in 11p11.2 was observed in two unrelated children presenting with two different CHD. One patient presented with CoA and carried a paternally inherited duplication. The other patient presented with ToF and carried a paternally inherited duplication. Both CNVs included a portion of the *PTPRJ* gene (figure 4b). The third overlapping region in 20p11.23 was detected in two unrelated children presenting with two different CHD. One patient presented with ToF and carried a *de novo* CNV duplication. The other patient presented with TGA and carried a maternally inherited duplication. Both CNVs includes the *ZNF133*, *POLR3F*, *RBBP9*, *DZANK1* and *CSRP2BP* genes (figure 4c).

In family 342 (Figure 5a), we identified a homozygous deletion inherited from both consanguineous parents. The deletion was identical and contained a portion of the *CHRNA3* gene. In family 417 (Figure 5b), we identified a rare microduplication in two children (II.1 and II.2). The duplication was inherited from a phenotypically normal mother and contained a portion of the *TCF7L3* gene and the *ERMARD* gene.

Gene prioritization

The two largest CNVs (7962,9 kb and 14587,4 kb) were removed for most of the bioinformatic analyses because they biased the analysis.

Knime4Bio

Knime4Bio allowed us to prioritize a list of genes included in our rare CNV regions and involved in heart development (GO:0007507) according to GeneOntology database (Figure 2b). We obtained the following list of candidate genes for CHD found in our patients: *TCAP*, *ERBB2*, *RB1CC*, *FBN1*, *GJA5*, *PTPRJ*, *NFATC3*, *PLN* and *OVOL2*.

DAPPLE and GRAIL

DAPPLE built direct and indirect interaction networks from proteins encoded by our seeded genes. That assessed the statistical significance of a number of network connectivity parameters as well as of the connectivity of individual proteins to other seed proteins using a within-degree node-label permutation method. The individual protein scores were then used to propose candidate genes in large loci. Based on textual relationships between genes, GRAIL assigned a p-value to each region suggesting its degree of functional connectivity, and picked up the best candidate gene (Supplementary table 5). No significant results were obtained for CoA group. *WDR33*, *CPSF3* and *NUSAP1* genes included in CNVs belonging to TGA group were statistically significant ($p < 0.05$). In the same way *GJA5*, *TJP1*, *GJA8*, *MSH2*, *DSG1*, *MCM3AP*, *DSG3* and *CTSS* genes from ToF group were also statistically significant ($p < 0.05$).

TFBS

According to TFBS tool, an overrepresentation of FOXC1 transcription factor binding sites ($p < 0,001$) was obtained in our CNVs regarding a normal distribution (Table

4). At least one gene containing FOXC1 binding sites was found in 54/69 (77%) of our CNV regions.

Discussion

To date few CNV studies of sporadic cases of CHD have been published (van Bon et al., 2008; Breckpot et al., 2010; Derwińska et al., 2012; Fakhro et al., 2011; Goldmuntz et al., 2011; Hartman et al., 2011; Lalani et al., 2013; Lu et al., 2008; Mefford et al., 2008; Serra-Juhé et al., 2012; Silversides et al., 2012; Thienpont et al., 2007), but there are even less studies reporting family trios composed by a child presenting with CHD and normal parents (Greenway et al., 2009; Hitz et al., 2012; Soemedi et al., 2012; Warburton et al., 2014). As it has been demonstrated in the literature, rare CNVs may account for CHD.

Here, we report a carefully phenotyped family trio-based CNV study in CoA, TGA and ToF. Our results evidence a strong association of rare CNV regions with CHD. Choosing a cohort based on family trios allowed us to find out rare CNV regions in the affected patients that were *de novo* or inherited. Seventy-one rare CHD related CNVs were identified (15 CoA, 29 TGA and 27 ToF) (Table 1). Rare CNV regions of variable size were detected in 22.5% of the children presenting with CHD. Duplications were more prevalent (70.5%) in our patient's cohort than deletions (29.5%). Regarding the transmission, 88.7% of the rare CNVs detected were inherited while 11.3% were *de novo*. The fact that some of our rare CNVs, potentially linked to the disease, are found in an affected child and one of his healthy parent contributes to the variable expressivity and the reduced penetrance characteristic of structural rearrangements (Girirajan and Eichler, 2010). The existence of healthy carriers evidences that rare rearrangements may not be the only cause of CHD in our patients. Given the multifactorial etiology of congenital malformations, these CNVs may be considered as one of the several factors involved.

The proportion of inherited CNVs was similar between the three CHD studied, however there was a statistically significant difference (*p=0,01) in the *de novo* CNVs

proportion found in CoA (4.1%) and ToF (6.6%) vs TGA (0%). The statistically significant difference in the number of *de novo* CNVs found in CoA and ToF vs TGA suggests that highly penetrant CNV's might be less frequent in the origin of TGA than in the ToF and CoA. Then, TGA may result from a genetic predisposition linked to low-impact multiple variants associated to environmental factors.

The results regarding CNV *de novo* regions found in our ToF cohort (6.6%) are broadly similar to previously published studies (Greenway et al., 2009; Soemedi et al., 2012) considering the differences in the analysis pipelines between the studies. The group of Greenway described a proportion of about 10% of *de novo* CNVs in their ToF patient's cohort; however some of these *de novo* CNVs described in their work were usually related to syndromic forms of CHD. Therefore, it is likely that our stringent patients' recruitment selection avoided us to find these CNVs (as 1q21.1 microdeletion) related to syndromic forms of CHD.

Looking at ISCA and DECIPHER public databases, we found 25 individuals presenting with CHD and carrying at least one CNV overlapping any of our patients' CNV regions (Supplementary table 4). Some of database individuals carried a CNV region much larger than the CNVs found in our described patients; therefore a second reduced list was created adjusting the CNVs sizes and the interest of the region (table 3). By this way, we ended up with a shorter list of 6 database CNVs overlapping 3 CNVs from 3 of our patients. Three database CNVs deletions corresponding to patients presenting with patent ductus arteriosus (PDA) and CoA (DECIPHER 1578, ISCA nssv706487, ISCA nssv706596) shared a genomic region with one of our described CNVs duplication, including the *GLIS3* gene. Another ISCA CNV deletion present in a patient with Dextrocardia (nssv580437) overlapped with one of our described CNVs duplications found in a patient presenting with TGA and comprising the *RLN3* gene. Also DECIPHER CNV duplication from a patient presenting with TGA (250627) overlapped with one of our CNV duplications, corresponding to a patient with TGA, in a region including the *PLN* gene. Finally, a last ISCA duplication CNV from a patient presenting with total anomalous pulmonary venous return (TAPVR) (nssv578742) overlapped with one of our *de novo* CNVs, corresponding to a patient with ToF, in a region including the *ERBB2* gene.

TFBS approach led us to identify FOXC1 binding sites in 54 out of 69 CNV regions (77%) (Table 4). FOXC1 belongs to the forkhead family of transcription factors.

The specific function of this gene has not yet been determined; however, it has been shown to play an essential role in the regulation of embryonic development in different model organisms (Kume et al., 2000; Siegenthaler et al., 2013; Wilm et al., 2004), as well as in cardiovascular development and concretely in cardiac outflow tract morphogenesis (Kume, 2009; Kume et al., 2001; Seo and Kume, 2006). A *de novo* CNV deletion of ~45 kb on chromosome 6q25.3 involving the *FOXC1* gene has been reported in a patient presenting with a syndromic form of CHD (Breckpot et al. 2010). In that study, the authors consider that the cause of the patient's phenotype (presenting with left ventricle outflow tract obstruction) is an haploinsufficiency of the *FOXC1* gene. Thus, the existence of microduplications/microdeletions including *FOXC1* binding sites can alter the expression of genes regulated by *FOXC1*. *FOXC1* becomes, then, in a strong candidate gene that may play a role in the cardiac malformations found in our cohort of patients.

The different strategies used to analyze our CNV data let us to build a list (Supplementary table 6) containing 124 candidate genes ready to be sequenced in a large cohort of CHD patients. Most of the genes included in that list are related or implicated in developmental networks; thus the alteration of these genes might cause congenital malformations. Developmental deficiencies resulting in conotruncal defects have been associated with distinct changes in gene expression (Bittel et al., 2011), describing a pattern of expression of developmentally important networks, supporting the hypothesis that converging and accumulating rare genomic and epigenetic variants may disrupt regulatory networks controlling heart development during embryological development ultimately leading to CHD. This hypothesis is also supported by other studies (Hitz et al., 2012; Serra-Juhé et al., 2012). According to the multifactorial origin of CHD, environmental factors during embryo development may also be considered as contributing elements to CHD in addition to the genetic abnormalities. The combinatorial effect of maternal diabetes and a partial deletion of the *FOX1* gene has been described as the cause of complex CHD in an affected newborn (Lale et al., 2011). Further work need to be done to determine more precisely the origin of non syndromic CHD and consequently helping in their diagnostic.

References

- Argaves, W.S., and Drake, C.J. (2005). Genes critical to vasculogenesis as defined by systematic analysis of vascular defects in knockout mice. *Anat. Rec. A. Discov. Mol. Cell. Evol. Biol.* 286, 875–884.
- Bittel, D.C., Butler, M.G., Kibiryeva, N., Marshall, J.A., Chen, J., Lofland, G.K., and O'Brien, J.E., Jr (2011). Gene expression in cardiac tissues from infants with idiopathic conotruncal defects. *BMC Med. Genomics* 4, 1.
- Blue, G.M., Kirk, E.P., Sholler, G.F., Harvey, R.P., and Winlaw, D.S. (2012). Congenital heart disease: current knowledge about causes and inheritance. *Med. J. Aust.* 197, 155–159.
- van Bon, B.W.M., Koolen, D.A., Pfundt, R., van der Burgt, I., de Leeuw, N., and de Vries, B.B.A. (2008). Transposition of the great vessels in a patient with a 2.9 Mb interstitial deletion of 9q31.1 encompassing the inversin gene: clinical report and review. *Am. J. Med. Genet. A.* 146A, 1225–1229.
- Botto, L.D., and Correa, A. (2003). Decreasing the burden of congenital heart anomalies: an epidemiologic evaluation of risk factors and survival. *Prog. Pediatr. Cardiol.* 18, 111–121.
- Breckpot, J., Thienpont, B., Peeters, H., de Ravel, T., Singer, A., Rayyan, M., Allegaert, K., Vanhole, C., Eyskens, B., Vermeesch, J.R., et al. (2010). Array Comparative Genomic Hybridization as a Diagnostic Tool for Syndromic Heart Defects. *J. Pediatr.* 156, 810–817.e4.
- Breckpot, J., Thienpont, B., Arens, Y., Tranchevent, L.C., Vermeesch, J.R., Moreau, Y., Gewillig, M., and Devriendt, K. (2011). Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenet. Genome Res.* 135, 251–259.
- Bruneau, B.G. (2008). The developmental genetics of congenital heart disease. *Nature* 451, 943–948.
- Burn, J., Brennan, P., Little, J., Holloway, S., Coffey, R., Somerville, J., Dennis, N.R., Allan, L., Arnold, R., Deanfield, J.E., et al. (1998). Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* 351, 311–316.
- Chan, R., Hardy, W.R., Laing, M.A., Hardy, S.E., and Muller, W.J. (2002). The catalytic activity of the ErbB-2 receptor tyrosine kinase is essential for embryonic development. *Mol. Cell. Biol.* 22, 1073–1078.
- Christiansen, J., Dyck, J.D., Elyas, B.G., Lilley, M., Bamforth, J.S., Hicks, M., Sprysak, K.A., Tomaszewski, R., Haase, S.M., Vicens-Wyhony, L.M., et al. (2004). Chromosome 1q21.1 contiguous gene deletion is associated with congenital heart disease. *Circ. Res.* 94, 1429–1435.

- Climent, S., Sarasa, M., Villar, J.M., and Murillo-Ferrol, N.L. (1995). Neurogenic cells inhibit the differentiation of cardiogenic cells. *Dev. Biol.* **171**, 130–148.
- Conrad, D.F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y., Aerts, J., Andrews, T.D., Barnes, C., Campbell, P., et al. (2010). Origins and functional impact of copy number variation in the human genome. *Nature* **464**, 704–712.
- Conti, V., Carabalona, A., Pallesi-Pocachard, E., Parrini, E., Leventer, R.J., Buhler, E., McGillivray, G., Michel, F.J., Striano, P., Mei, D., et al. (2013). Periventricular heterotopia in 6q terminal deletion syndrome: role of the C6orf70 gene. *Brain J. Neurol.* **136**, 3378–3394.
- Derwińska, K., Bartnik, M., Wiśniowiecka-Kowalik, B., Jagła, M., Rudziński, A., Pietrzyk, J.J., Kawalec, W., Ziółkowska, L., Kutkowska-Kaźmierczak, A., Gambin, T., et al. (2012). Assessment of the role of copy-number variants in 150 patients with congenital heart defects. *Med. Wieku Rozwoj.* **16**, 175–182.
- Digilio, M.C., Casey, B., Toscano, A., Calabrò, R., Pacileo, G., Marasini, M., Banaudi, E., Giannotti, A., Dallapiccola, B., and Marino, B. (2001). Complete transposition of the great arteries: patterns of congenital heart disease in familial precurrence. *Circulation* **104**, 2809–2814.
- Edgar, R., Mazor, Y., Rinon, A., Blumenthal, J., Golan, Y., Buzhor, E., Livnat, I., Ben-Ari, S., Lieder, I., Shitrit, A., et al. (2013). LifeMap DiscoveryTM: the embryonic development, stem cells, and regenerative medicine research portal. *PloS One* **8**, e66629.
- Fakhro, K.A., Choi, M., Ware, S.M., Belmont, J.W., Towbin, J.A., Lifton, R.P., Khokha, M.K., and Brueckner, M. (2011). Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2915–2920.
- Ferencz, C., Rubin, J.D., McCarter, R.J., Brenner, J.I., Neill, C.A., Perry, L.W., Hepner, S.I., and Downing, J.W. (1985). Congenital heart disease: prevalence at livebirth. The Baltimore-Washington Infant Study. *Am. J. Epidemiol.* **121**, 31–36.
- Ferrero, G.B., Biamino, E., Sorasio, L., Banaudi, E., Peruzzi, L., Forzano, S., di Cantogno, L.V., and Silengo, M.C. (2007). Presenting phenotype and clinical evaluation in a cohort of 22 Williams-Beuren syndrome patients. *Eur. J. Med. Genet.* **50**, 327–337.
- De Galan-Roosen, A.E., Kuijpers, J.C., Meershoek, A.P., and van Velzen, D. (1998). Contribution of congenital malformations to perinatal mortality. A 10 years prospective regional study in The Netherlands. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **80**, 55–61.
- Gessert, S., and Kühl, M. (2010). The multiple phases and faces of wnt signaling during cardiac differentiation and development. *Circ. Res.* **107**, 186–199.
- Gillum, R.F. (1994). Epidemiology of congenital heart disease in the United States. *Am. Heart J.* **127**, 919–927.
- Girirajan, S., and Eichler, E.E. (2010). Phenotypic variability and genetic susceptibility to genomic disorders. *Hum. Mol. Genet.* **19**, R176–187.

- Goldmuntz, E., Paluru, P., Glessner, J., Hakonarson, H., Biegel, J.A., White, P.S., Gai, X., and Shaikh, T.H. (2011). Microdeletions and microduplications in patients with congenital heart disease and multiple congenital anomalies. *Congenit. Heart Dis.* 6, 592–602.
- Greenway, S.C., Pereira, A.C., Lin, J.C., DePalma, S.R., Israel, S.J., Mesquita, S.M., Ergul, E., Conta, J.H., Korn, J.M., McCarroll, S.A., et al. (2009). De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* 41, 931–935.
- Guelman, S., Kozuka, K., Mao, Y., Pham, V., Solloway, M.J., Wang, J., Wu, J., Lill, J.R., and Zha, J. (2009). The double-histone-acetyltransferase complex ATAC is essential for mammalian development. *Mol. Cell. Biol.* 29, 1176–1188.
- Hartman, R.J., Rasmussen, S.A., Botto, L.D., Riehle-Colarusso, T., Martin, C.L., Cragan, J.D., Shin, M., and Correa, A. (2011). The contribution of chromosomal abnormalities to congenital heart defects: a population-based study. *Pediatr. Cardiol.* 32, 1147–1157.
- Henderson, D.J., Phillips, H.M., and Chaudhry, B. (2006). Vang-like 2 and noncanonical Wnt signaling in outflow tract development. *Trends Cardiovasc. Med.* 16, 38–45.
- Hitz, M.-P., Lemieux-Perreault, L.-P., Marshall, C., Feroz-Zada, Y., Davies, R., Yang, S.W., Lionel, A.C., D'Amours, G., Lemyre, E., Cullum, R., et al. (2012). Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet.* 8, e1002903.
- Iafrate, A.J., Feuk, L., Rivera, M.N., Listewnik, M.L., Donahoe, P.K., Qi, Y., Scherer, S.W., and Lee, C. (2004). Detection of large-scale variation in the human genome. *Nat. Genet.* 36, 949–951.
- Kim, Y.-S., Nakanishi, G., Lewandoski, M., and Jetten, A.M. (2003). GLIS3, a novel member of the GLIS subfamily of Krüppel-like zinc finger proteins with repressor and activation functions. *Nucleic Acids Res.* 31, 5513–5525.
- Kirchhoff, M., Bisgaard, A.-M., Stoeva, R., Dimitrov, B., Gillessen-Kaesbach, G., Fryns, J.-P., Rose, H., Grozdanova, L., Ivanov, I., Keymolen, K., et al. (2009). Phenotype and 244k array-CGH characterization of chromosome 13q deletions: an update of the phenotypic map of 13q21.1-qter. *Am. J. Med. Genet. A.* 149A, 894–905.
- Kodo, K., Nishizawa, T., Furutani, M., Arai, S., Yamamura, E., Joo, K., Takahashi, T., Matsuoka, R., and Yamagishi, H. (2009). GATA6 mutations cause human cardiac outflow tract defects by disrupting semaphorin-plexin signaling. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13933–13938.
- Kume, T. (2009). The cooperative roles of Foxc1 and Foxc2 in cardiovascular development. *Adv. Exp. Med. Biol.* 665, 63–77.
- Kume, T., Deng, K., and Hogan, B.L. (2000). Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract. *Dev. Camb. Engl.* 127, 1387–1395.

- Kume, T., Jiang, H., Topczewska, J.M., and Hogan, B.L. (2001). The murine winged helix transcription factors, *Foxc1* and *Foxc2*, are both required for cardiovascular development and somitogenesis. *Genes Dev.* 15, 2470–2482.
- Lalani, S.R., Ware, S.M., Wang, X., Zapata, G., Tian, Q., Franco, L.M., Jiang, Z., Bucasas, K., Scott, D.A., Campeau, P.M., et al. (2013). *MCTP2* is a dosage-sensitive gene required for cardiac outflow tract development. *Hum. Mol. Genet.* 22, 4339–4348.
- Lale, S., Yu, S., and Ahmed, A. (2011). Complex congenital heart defects in association with maternal diabetes and partial deletion of the *A2BP1* gene. *Fetal Pediatr. Pathol.* 30, 161–166.
- Lambertini, L., Marsit, C.J., Sharma, P., Maccani, M., Ma, Y., Hu, J., and Chen, J. (2012). Imprinted gene expression in fetal growth and development. *Placenta* 33, 480–486.
- Lavallée, G., Andelfinger, G., Nadeau, M., Lefebvre, C., Nemer, G., Horb, M.E., and Nemer, M. (2006). The Kruppel-like transcription factor *KLF13* is a novel regulator of heart development. *EMBO J.* 25, 5201–5213.
- Lee, C.-H., Wai, Y.-Y., and Wu, T. (2011). Periventricular nodular heterotopia and cardiovascular defects. *Chang Gung Med. J.* 34, 628–635.
- Lee, K.F., Simon, H., Chen, H., Bates, B., Hung, M.C., and Hauser, C. (1995). Requirement for neuregulin receptor *erbB2* in neural and cardiac development. *Nature* 378, 394–398.
- Lindenbaum, P., Le Scouarnec, S., Portero, V., and Redon, R. (2011). Knime4Bio: a set of custom nodes for the interpretation of next-generation sequencing data with KNIME. *Bioinforma. Oxf. Engl.* 27, 3200–3201.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif* 25, 402–408.
- Lu, X.-Y., Phung, M.T., Shaw, C.A., Pham, K., Neil, S.E., Patel, A., Sahoo, T., Bacino, C.A., Stankiewicz, P., Kang, S.-H.L., et al. (2008). Genomic Imbalances in Neonates With Birth Defects: High Detection Rates by Using Chromosomal Microarray Analysis. *Pediatrics* 122, 1310–1318.
- Maniatis, T. (1999). A ubiquitin ligase complex essential for the NF-kappaB, Wnt/Wingless, and Hedgehog signaling pathways. *Genes Dev.* 13, 505–510.
- Mefford, H.C., Sharp, A.J., Baker, C., Itsara, A., Jiang, Z., Buysse, K., Huang, S., Maloney, V.K., Crolla, J.A., Baralle, D., et al. (2008). Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.* 359, 1685–1699.
- Nogueira, T.C., Paula, F.M., Villate, O., Colli, M.L., Moura, R.F., Cunha, D.A., Marselli, L., Marchetti, P., Cnop, M., Julier, C., et al. (2013). *GLIS3*, a susceptibility gene for type 1 and type 2 diabetes, modulates pancreatic beta cell apoptosis via regulation of a splice variant of the BH3-only protein Bim. *PLoS Genet.* 9, e1003532.

- Nora, J.J. (1968). Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* 38, 604–617.
- Øyen, N., Poulsen, G., Boyd, H.A., Wohlfahrt, J., Jensen, P.K.A., and Melbye, M. (2009). Recurrence of congenital heart defects in families. *Circulation* 120, 295–301.
- Petrini, J., Damus, K., and Johnston, R.B., Jr (1997). An overview of infant mortality and birth defects in the United States. *Teratology* 56, 8–10.
- Raychaudhuri, S., Plenge, R.M., Rossin, E.J., Ng, A.C.Y., International Schizophrenia Consortium, Purcell, S.M., Sklar, P., Scolnick, E.M., Xavier, R.J., Altshuler, D., et al. (2009). Identifying relationships among genomic disease regions: predicting genes at pathogenic SNP associations and rare deletions. *PLoS Genet.* 5, e1000534.
- Richards, E.G., Zaveri, H.P., Wolf, V.L., Kang, S.-H.L., and Scott, D.A. (2011). Delineation of a less than 200 kb minimal deleted region for cardiac malformations on chromosome 7p22. *Am. J. Med. Genet. A.* 155A, 1729–1734.
- Rivera, H., Merinero, B., Martinez-Pardo, M., Arroyo, I., Ruiz-Sala, P., Bornstein, B., Serra-Suhe, C., Gallardo, E., Marti, R., Moran, M.J., et al. (2010). Marked mitochondrial DNA depletion associated with a novel SUCLG1 gene mutation resulting in lethal neonatal acidosis, multi-organ failure, and interrupted aortic arch. *Mitochondrion* 10, 362–368.
- Rosenthal, E. (2005). Coarctation of the aorta from fetus to adult: curable condition or life long disease process? *Heart Br. Card. Soc.* 91, 1495–1502.
- Rossin, E.J., Lage, K., Raychaudhuri, S., Xavier, R.J., Tatar, D., Benita, Y., International Inflammatory Bowel Disease Genetics Consortium, Cotsapas, C., and Daly, M.J. (2011). Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. *PLoS Genet.* 7, e1001273.
- Seo, S., and Kume, T. (2006). Forkhead transcription factors, Foxc1 and Foxc2, are required for the morphogenesis of the cardiac outflow tract. *Dev. Biol.* 296, 421–436.
- Serra-Juhé, C., Rodríguez-Santiago, B., Cuscó, I., Vendrell, T., Camats, N., Torán, N., and Pérez-Jurado, L.A. (2012). Contribution of rare copy number variants to isolated human malformations. *PloS One* 7, e45530.
- Shuib, S., McMullan, D., Rattenberry, E., Barber, R.M., Rahman, F., Zatyka, M., Chapman, C., Macdonald, F., Latif, F., Davison, V., et al. (2009). Microarray based analysis of 3p25-p26 deletions (3p- syndrome). *Am. J. Med. Genet. A.* 149A, 2099–2105.
- Siegenthaler, J.A., Choe, Y., Patterson, K.P., Hsieh, I., Li, D., Jaminet, S.-C., Daneman, R., Kume, T., Huang, E.J., and Pleasure, S.J. (2013). Foxc1 is required by pericytes during fetal brain angiogenesis. *Biol. Open* 2, 647–659.
- Silversides, C.K., Lionel, A.C., Costain, G., Merico, D., Migita, O., Liu, B., Yuen, T., Rickaby, J., Thiruvahindrapuram, B., Marshall, C.R., et al. (2012). Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* 8, e1002843.

Soemedi, R., Wilson, I.J., Bentham, J., Darlay, R., Töpf, A., Zelenika, D., Cosgrove, C., Setchfield, K., Thornborough, C., Granados-Riveron, J., et al. (2012). Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am. J. Hum. Genet.* **91**, 489–501.

Takahashi, T., Takahashi, K., St John, P.L., Fleming, P.A., Tomemori, T., Watanabe, T., Abrahamson, D.R., Drake, C.J., Shirasawa, T., and Daniel, T.O. (2003). A mutant receptor tyrosine phosphatase, CD148, causes defects in vascular development. *Mol. Cell. Biol.* **23**, 1817–1831.

Thienpont, B., Mertens, L., de Ravel, T., Eyskens, B., Boshoff, D., Maas, N., Fryns, J.-P., Gewillig, M., Vermeesch, J.R., and Devriendt, K. (2007). Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. *Eur. Heart J.* **28**, 2778–2784.

Tian, Y., Yuan, L., Goss, A.M., Wang, T., Yang, J., Lepore, J.J., Zhou, D., Schwartz, R.J., Patel, V., Cohen, E.D., et al. (2010). Characterization and in vivo pharmacological rescue of a Wnt2-Gata6 pathway required for cardiac inflow tract development. *Dev. Cell* **18**, 275–287.

Tomita-Mitchell, A., Mahnke, D.K., Struble, C.A., Tuffnell, M.E., Stamm, K.D., Hidestrand, M., Harris, S.E., Goetsch, M.A., Simpson, P.M., Bick, D.P., et al. (2012). Human gene copy number spectra analysis in congenital heart malformations. *Physiol. Genomics* **44**, 518–541.

Tyreman, M., Abbott, K.M., Willatt, L.R., Nash, R., Lees, C., Whittaker, J., and Simoncic, I. (2009). High resolution array analysis: diagnosing pregnancies with abnormal ultrasound findings. *J. Med. Genet.* **46**, 531–541.

Unemori, E.N., Lewis, M., Constant, J., Arnold, G., Grove, B.H., Normand, J., Deshpande, U., Salles, A., Pickford, L.B., Erikson, M.E., et al. (2000). Relaxin induces vascular endothelial growth factor expression and angiogenesis selectively at wound sites. *Wound Repair Regen. Off. Publ. Wound Heal. Soc. Eur. Tissue Repair Soc.* **8**, 361–370.

Veltman, J.A., Schoenmakers, E.F.P.M., Eussen, B.H., Janssen, I., Merkx, G., van Cleef, B., van Ravenswaaij, C.M., Brunner, H.G., Smeets, D., and van Kessel, A.G. (2002). High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. *Am. J. Hum. Genet.* **70**, 1269–1276.

Venturin, M., Guarnieri, P., Natacci, F., Stabile, M., Tenconi, R., Clementi, M., Hernandez, C., Thompson, P., Upadhyaya, M., Larizza, L., et al. (2004). Mental retardation and cardiovascular malformations in NF1 microdeleted patients point to candidate genes in 17q11.2. *J. Med. Genet.* **41**, 35–41.

Warburton, D., Ronemus, M., Kline, J., Jobanputra, V., Williams, I., Anyane-Yeboah, K., Chung, W., Yu, L., Wong, N., Awad, D., et al. (2014). The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. *Hum. Genet.* **133**, 11–27.

Wessels, M.W., and Willems, P.J. (2010). Genetic factors in non-syndromic congenital heart malformations. *Clin. Genet.* 78, 103–123.

Wilm, B., James, R.G., Schultheiss, T.M., and Hogan, B.L.M. (2004). The forkhead genes, *Foxc1* and *Foxc2*, regulate paraxial versus intermediate mesoderm cell fate. *Dev. Biol.* 271, 176–189.

De Wit, M.C.Y., Kros, J.M., Halley, D.J.J., de Coo, I.F.M., Verdijk, R., Jacobs, B.C., and Mancini, G.M.S. (2009). Filamin A mutation, a common cause for periventricular heterotopia, aneurysms and cardiac defects. *J. Neurol. Neurosurg. Psychiatry* 80, 426–428.

Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 498, 220–223.

ZeRuth, G.T., Takeda, Y., and Jetten, A.M. (2013). The Krüppel-like protein Gli-similar 3 (*Glis3*) functions as a key regulator of insulin transcription. *Mol. Endocrinol. Baltim. Md* 27, 1692–1705.

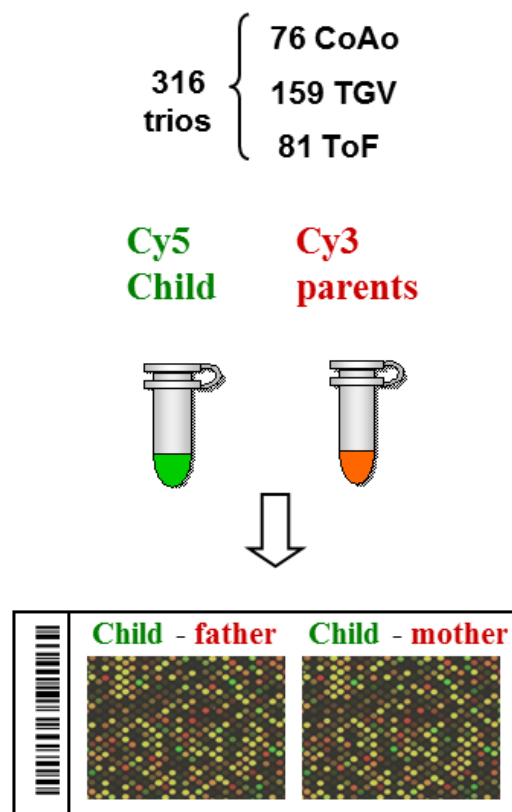
Zhou, W., Lin, L., Majumdar, A., Li, X., Zhang, X., Liu, W., Etheridge, L., Shi, Y., Martin, J., Van de Ven, W., et al. (2007). Modulation of morphogenesis by noncanonical Wnt signaling requires ATF/CREB family-mediated transcriptional activation of *TGFbeta2*. *Nat. Genet.* 39, 1225–1234.

Van der Zwaag, P.A., van Rijsingen, I.A.W., Asimaki, A., Jongbloed, J.D.H., van Veldhuisen, D.J., Wiesfeld, A.C.P., Cox, M.G.P.J., van Lochem, L.T., de Boer, R.A., Hofstra, R.M.W., et al. (2012). Phospholamban R14del mutation in patients diagnosed with dilated cardiomyopathy or arrhythmogenic right ventricular cardiomyopathy: evidence supporting the concept of arrhythmogenic cardiomyopathy. *Eur. J. Heart Fail.* 14, 1199–1207.

Acknowledgments

C.L.C and M.S.C. were supported by grants from PHRC Inter regional (2008), Société Française de Cardiologie / Fédération Française de Cardiologie (2009) and Translational Research of Région des Pays de la Loire (2009). M.S.C. was also supported by École nationale supérieure des mines de Nantes and Genavie enterprise foundation. We are most grateful to the Genomics platform of Nantes (Biogenouest Genomics) core facility for its technical support.

Figures



Agilent 2x400K custom-made array:

-300.000 probes covering 95% exonic regions

-100.000 probes densifying 300 candidate genes

Figure 1. Custom aCGH. 316 family trios (76 CoAo, 159 TGA and 81 ToF) were selected for aCGH. A different color fluorescent dye was used to label children's DNA in green (Cy5) and parent's DNA in red (Cy3). Labeled DNAs (affected child-father, affected child-mother) were put to cohybridate on the double chambered aCGH containing 300K probes covering >95% of the genome plus 100K additional probes densifying 297 candidate genes known to be responsible for CHD in humans and in animals.

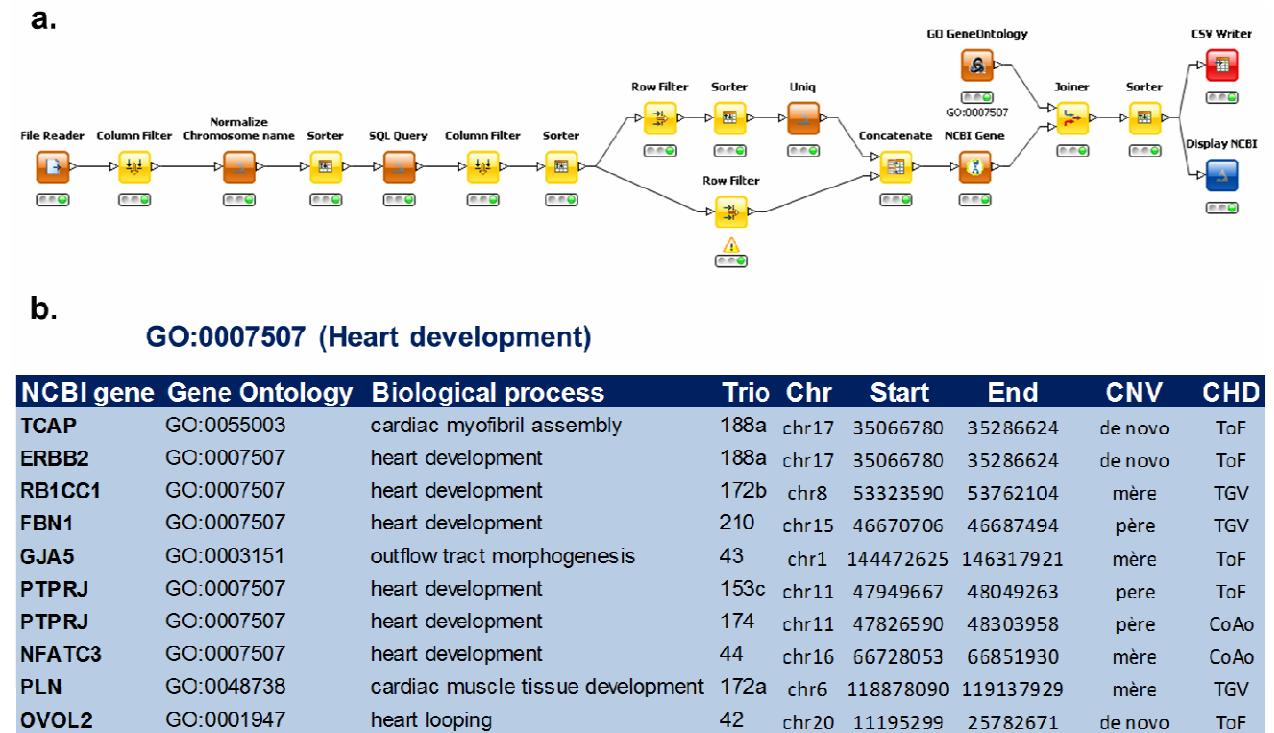


Figure 2. Knime4Bio. Developed by UMR1087 Institut de Thorax of Nantes and based in a set of custom nodes for the interpretation of large datasets (Lindenbaum et al., 2011). This tool allowed us to quickly find the genes inside of our rare validated CNV regions that were implicated in heart development (GO:0007507) according to GeneOntology database (<http://www.geneontology.org/>).

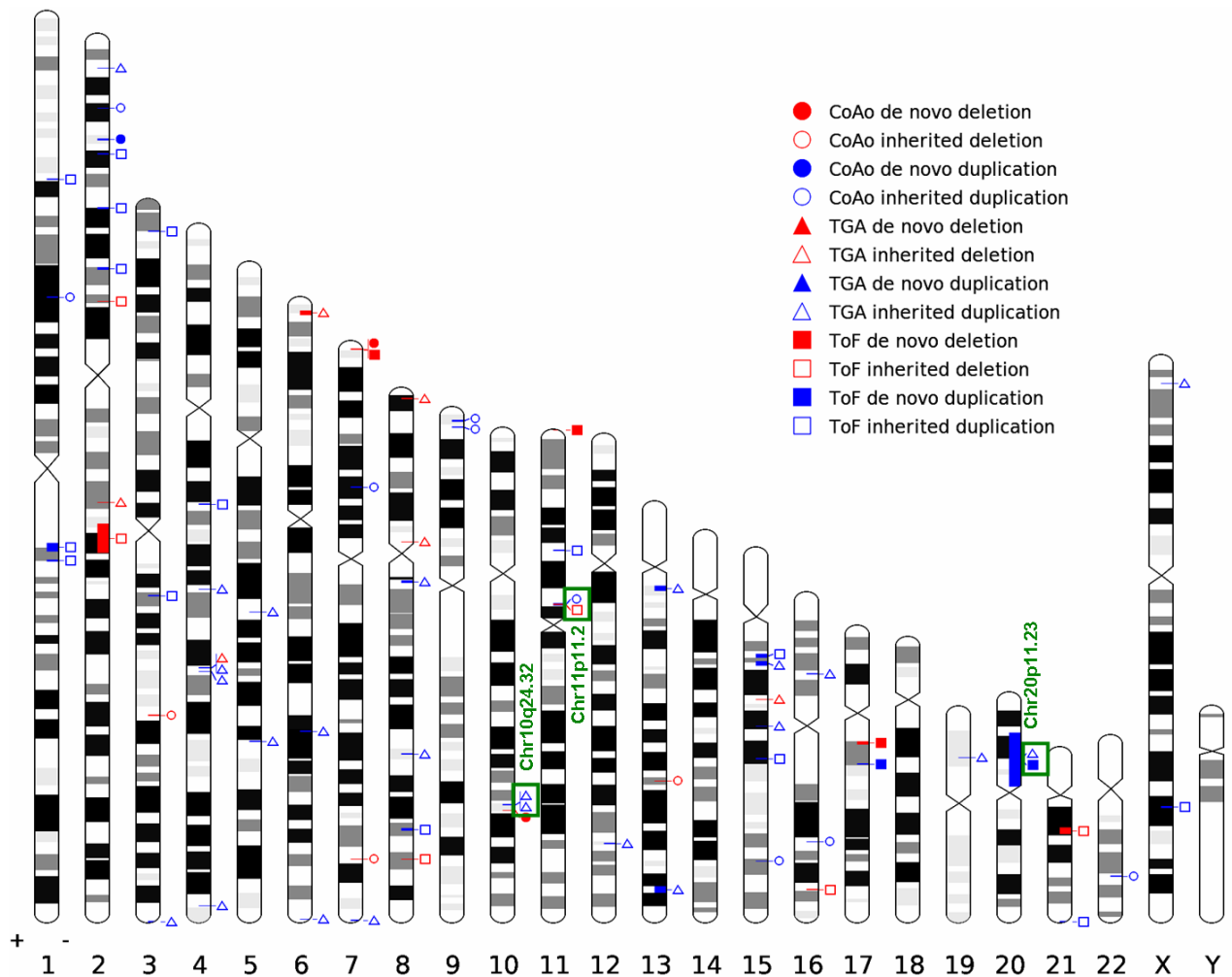


Figure 3. Ideographica (<http://www.ncrna.org/idiographica/>) results show a spread distribution of the 71 rare CNV regions present in our CHD affected cohort all over the genome. CoA is represented by a circle, TGA by a triangle and ToF by a square. Filled geometric figures represent *de novo* rearrangements and empty geometric figures represent inherited rearrangements. Geometric figures colored in red refer to deletions and the ones colored in blue refer to duplications. Three overlaps have been observed in three regions (green rectangles). The first region Chr10q24.32 has an overlap with two inherited duplications found in two patients presenting with TGA. The second region Chr11p11.2 overlaps with an inherited duplication found in a patient presenting with CoA and an inherited deletion found in a patient presenting with ToF. The last overlap Chr20p11.23 is composed by an inherited duplication found in a patient presenting with TGA and a *de novo* duplication found in a patient presenting with ToF.

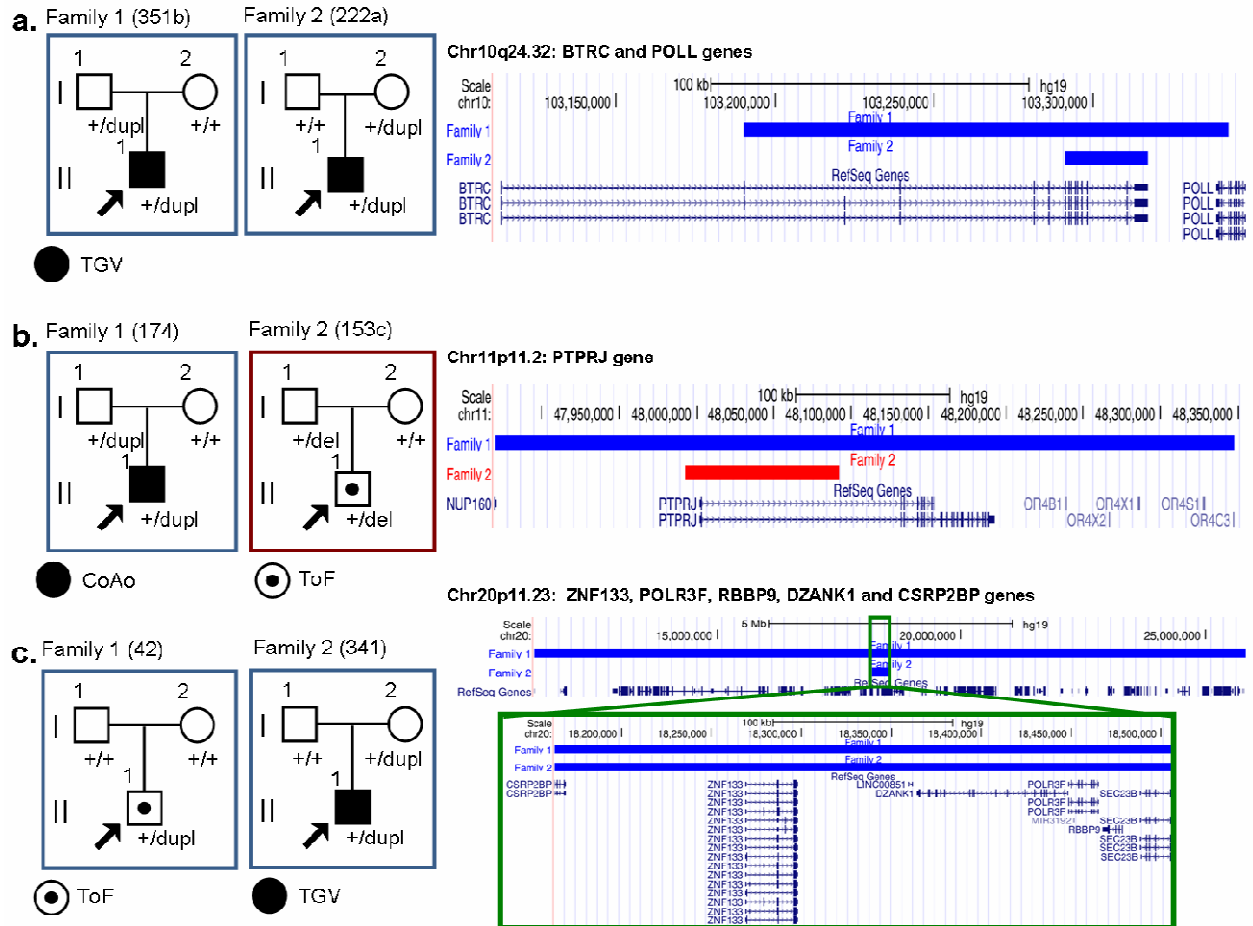


Figure 4. a. CNV overlapping observed in Chr10q24.32, corresponding to two unrelated children presenting with TGA and carrying an inherited CNV duplication from a different progenitor in each patient. Both CNVs include a part of the *BTRC* gene. **b.** CNV overlapping observed in Chr11p11.2 corresponding to two unrelated children presenting with different CHD. One patient presented with CoA and carried an inherited CNV duplication from the father. The other patient presented with ToF and carried an inherited CNV deletion also from the father. Both CNVs included a part of the *PTPRJ* gene. **c.** CNV overlapping observed in Chr20p11.23 corresponding to two unrelated children presenting with different CHD. Patient from family 1 presented with ToF and carried a *de novo* CNV duplication. Patient from family 2 presented with TGA and carried an inherited CNV duplication from the mother. The zoomed in picture shows that both CNVs include *ZNF133*, *POLR3F*, *RBBP9*, *DZANK1* and *CSRP2BP* genes.

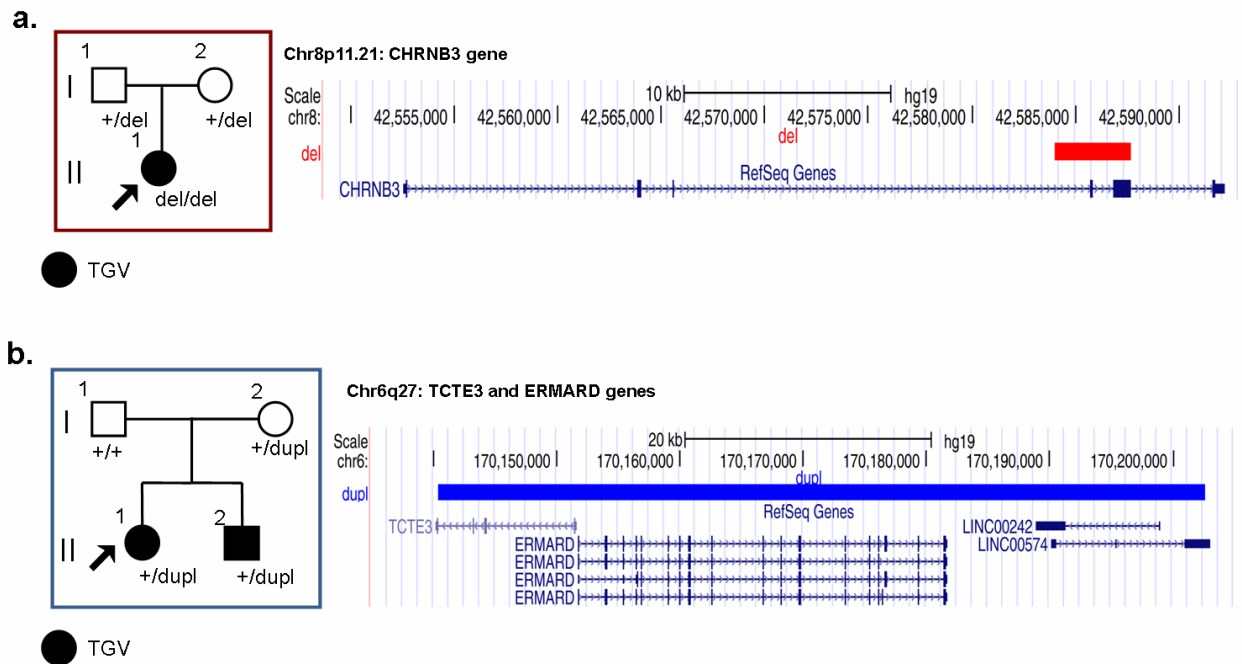


Figure 5. a. Family pedigree of a consanguineous family trio having a child presenting with TGA and carrying a rare CNV deletion inherited from both parents. The child carries a homozygous microdeletion including a part of the *CHRN3* gene. **b.** Family pedigree showing 2 normal parents with 2 children presenting a TGA and carrying a maternally inherited CNV rare duplication which contains the *ERMARD* gene and part of the *TCTE3* gene.

Tables

Table 1. qPCR validation. From 152 selected regions only 71 CNVs were confirmed to be present in the affected children (15 for CoA, 29 for TGA and 27 for ToF). In most of the cases (63/71) the CNV regions were inherited from one of the parents. Only 8 CNVs found in affected children were *de novo*: 3 with CoA (4.1%), 5 with ToF (6.6%) and 0* with TGA. The results obtained for the *de novo* CNVs in the TGA patients were statistically significant (*p=0, 01) in comparison with the other two pathologies.

	Trios CGH	# <i>de novo</i> CNVs		# inherited CNVs
➤ 85 CoAo	76	3	4.1%*	12
➤ 291 TGV	159	0	0%	29
➤ 92 ToF	81	5	6.6%*	22
➤ Total	316	8	2.6%	63

* p=0.01

Table 2. Table of 71 validated rare CNV regions (50 duplications and 21 deletions) in children presenting with outflow tract defect. Generally, CNV sizes ranged from 2,9kb to 1845,3kb, except for two much bigger CNVs of 7962,9kb and 14587,4kb. Most of CNVs (63/71) were inherited from one of the parents, only 8 CNVs were found to be *de novo*. The frequencies of these CNV in general population went from 0% to 1% in some cases (according to the following databases: 42 million probes study (Conrad et al., 2010), WTCCC study <http://www.wtccc.org.uk/>, 1000 genomes project <http://www.1000genomes.org/> and DDD controls project <http://www.ddduk.org/>).

Case #	CHD	Gain / Loss	Region	Length (Kb)	Start	End	Array used	Inheritance	RefSeq Genes in the region	Frequency in general population
212	CoAo	Gain	2p32.2	524,815	28796008	29320823	400K oligonucleotide	<i>De novo</i>	PLB1, PPP1CB, SPDYA, TRMT61B, WDR43, FAM179A, C2orf71, CLIP4	0%
163	CoAo	Loss	7p22.3	19,527	2255641	2275168	400K oligonucleotide	<i>De novo</i>	MAD1L1, FTSJ2	1 singleton observation
6	CoAo	Loss	10q24.32	28,683	104828409	104857092	400K oligonucleotide	<i>De novo</i>	CNNM2, NT5C2	0%
179	CoAo	Gain	7p14.1	109,041	40027276	40136317	400K oligonucleotide	Paternal	CDK13	0%
174	CoAo	Gain	11p11.2	477,368	47870014	48347382	400K oligonucleotide	Paternal	NUP160, PTPRJ, OR4B1, OR4X2, OR4X1, OR4S1, OR4C3	0%
44	CoAo	Gain	16q22.1	123,877	68170552	68294429	400K oligonucleotide	Maternal	NFATC3, ESRP2, MIR6773, PLA2G15	0%
59	CoAo	Loss	7q34	35,204	141705389	141740593	400K oligonucleotide	Paternal	MGAM	0%
23	CoAo	Gain	2p24.1	47,889	20403619	20451508	400K oligonucleotide	Paternal	SDC1, PUM2	0%
90a	CoAo	Gain	1p31.1	75,025	78308983	78384008	400K oligonucleotide	Paternal	FAM73A, NEXN-AS1, NEXN	0%
77a	CoAo	Loss	13q22.2	17,101	76391357	76408458	400K oligonucleotide	Maternal	LMO7	0%
77b	CoAo	Gain	22q13.1	17,5	38609792	38627292	400K oligonucleotide	Paternal	MAFF, TMEM184B	0%
211	CoAo	Loss	3q23	112,858	141272742	141385600	400K oligonucleotide	Maternal	RASA2	0%
226a	CoAo	Gain	15q25.3	76,111	85606120	85682231	400K oligonucleotide	Maternal	PDE8A	1 singleton observation
437b	CoAo	Gain	9p24.2	379,713	3781683	4161396	400K oligonucleotide	Paternal	GLIS3, GLIS3-AS1	0%
437c	CoAo	Gain	9p24.1	231,015	5534835	5765850	400K oligonucleotide	Paternal	PDCD1LG2, KIAA1432	0%
255	TGV	Gain	4q23	15,292	100226779	100242071	400K oligonucleotide	Maternal	ADH1B	0%
194b	TGV	Gain	15q13.2-q13.3	1207,316	31196707	32404023	400K oligonucleotide	Maternal	FAN1, MTMR10, TRPM1, MIR211, RP11-16E12.2, KLF13, OTUD7A, CHRNA7	doubleton to 1%
231b	TGV	Loss	4q27	82,199	121616153	121698352	400K oligonucleotide	Paternal	PRDM5	0%
252b	TGV	Loss	8q22.2	198,691	100205095	100403786	400K oligonucleotide	Maternal	VPS13B	0%
210	TGV	Gain	15q21.1	16,788	48883414	48900202	400K oligonucleotide	Paternal	FBN1	0%
196a	TGV	Loss	2q14.3	7,523	128476795	128484318	400K oligonucleotide	Maternal	WDR33	0%
196b	TGV	Loss	8p23.2	250,895	2975904	3226799	400K oligonucleotide	Paternal	CSMD1	0%
265	TGV	Gain	2p25.1	124,867	9525380	9650247	400K oligonucleotide	Paternal	ASAP2, ITGB1BP1, CPSF3, IAH1, ADAM17	0%

277	TGV	Gain	3q29	241,257	197566221	197807478	400K oligonucleotide	Paternal	LRCH3, IQCG, RPL35A, LMLN, ANKRD18DP	doubleton to 1%
335a	TGV	Gain	19p13.12	21,217	14138589	14159806	400K oligonucleotide	Paternal	RLN3, IL27RA	0%
335b	TGV	Gain	Xp22.31	56,998	7866965	7923963	400K oligonucleotide	Paternal	PNPLA4	0%
351b	TGV	Gain	10q24.32	152,533	103190059	103342592	400K oligonucleotide	Paternal	BTRC, POLL	0%
358c	TGV	Gain	13q33.2	1647,761	105161812	106809573	400K oligonucleotide	Maternal	DAOA-AS1, DAOA, LINC00343	0%
222a	TGV	Gain	10q24.32	26,002	103291033	103317035	400K oligonucleotide	Maternal	BTRC	doubleton to 1%
172a	TGV	Gain	6q22.31	259,839	118771397	119031236	400K oligonucleotide	Maternal	CEP85L, BRD7P3, PLN	doubleton to 1%
172b	TGV	Gain	8q11.23	438,514	53161037	53599551	400K oligonucleotide	Maternal	ST18, FAM150A, RB1CC1	0%
297	TGV	Gain	4q27	4,621	122735142	122739763	400K oligonucleotide	Maternal	EXOSC9, CCNA2	0%
165b	TGV	Loss	15q15.1	25,534	41663852	41689386	400K oligonucleotide	Paternal	NUSAP1, NDUFAF1	0%
284a	TGV	Gain	4q35.1	66,434	186558702	186625136	400K oligonucleotide	Paternal	SORBS2	1 singleton observation
302	TGV	Gain	5q31.1	64,839	131247533	131312372	400K oligonucleotide	Paternal	ACSL6, LOC728637	0%
341	TGV	Gain	20p11.23	343,15	18162446	18505596	400K oligonucleotide	Maternal	CSRP2BP, ZNF133, LINC00851, DZANK1, POLR3F, MIR3192, RBBP9,	0%
356	TGV	Gain	16p12.2	74,101	22357397	22431498	400K oligonucleotide	Maternal	CDR2, PRN3P3	doubleton to 1%
417	TGV	Gain	6q27	62,207	170140381	170202588	400K oligonucleotide	Maternal	TCTE3, ERMARD, LINC00242, LINC00574	0
393	TGV	Gain	5q15	36,42	96107398	96143818	400K oligonucleotide	Maternal	CAST, ERAP1	doubleton to 1%
294a	TGV	Gain	13q12.11-q12.12	1280,378	23235248	24515626	400K oligonucleotide	Maternal	BASP1P1, SGCG, SACS, SACS-AS1, LINC00327, TNFRSF19, MIPPEP, C1QTNF9B-AS1, ANKRD20A19P	0%
295	TGV	Gain	7q36.3	279,466	158331982	158611448	400K oligonucleotide	Maternal	PTPRN2, MIR5707, NCAPG2, ESYT2	0
329a	TGV	Gain	12q24.12-q24.13	126,387	112182542	112308929	400K oligonucleotide	Paternal	ACAD10, ALDH2, MIR6761, MAPKAPK5-AS1, MPKAPK5	doubleton to 1%
342	TGV	Loss	8p11.21	3,712	42583985	42587697	400K oligonucleotide	Maternal / Paternal	CHRNA3	0%
42	ToF	Gain	20p12.2-p11.1	14587,372	11247299	25834671	400K oligonucleotide	De novo	LOC339593, LINC00687, BTBD3, RP5-1069C8.2, LOC102606466, LOC100505515, SPTLC3, ISM1, ISM1-AS1, TASP1, ESF1, NDUFAF5, SEL1L2, MACROD2, FLRT3, MACROD2-IT1, MACROD2-AS1, KIF16B, SNRPB2, OTOR, PCSK2, BFSP1, DSTN, RRPB1, BANF2, SNX5, SNORD17, MGME1, OVOL2, CSRP2BP, ZNF133, LINC00851, DZANK1, POLR3F, MIR3192, RBBP9, SEC23B, LINC00493, DTD1, RP11-379J5.5, LINC00652, LOC100270804, C20orf78, SCP2D1, SLC24A3, RP5-1027G4.3, RIN2, NAA20, CRNKL1, C20orf26, INSM1, RALGAP2, KIZ, RP4-777D9.2, XRN2, NKX2-4, NKX2-2, LOC101929625, LOC1019608, PAX1, RP11-125P18.1, RP5-828K20.2, RP5-1004I9.1, LINC00261, FOXA2, RP4-788L20.3, SSTR4, THBD, CD93, LINC00656, NXT1, RP3-322G13.5, GZF1, NAPB, CSTL1, CST11, CST8, CST13P, CST9L, CST9, CST3, CST4, CST1, CST2, CST5, GGTLC1, FLJ33581, SYNDIG1, CST7, APMAP, ACSS1, VSX1, RP4-738P15.1,	0%

188a	ToF	Gain	17q12	219,844	37813254	38033098	400K oligonucleotide	De novo	STARD3, TCAP, PNMT, PGAP3, ERBB2, MIR4728, MIEN1, GRB7, IKZF3, ZPBP2	0%
24	ToF	Loss	17q12	741,259	31856936	32598195	400K oligonucleotide	De novo	ASIC2, AA06, RP11-215E13.1, CCL2, CCL7	0%
11	ToF	Loss	7p22.3	19,967	2400442	2420409	400K oligonucleotide	De novo	EIF3B	0%
515	ToF	Loss	11p15.5	127,12	193816	320936	400K oligonucleotide	De novo	LOC653486, SCGB1C1, ODF3, BETIL1, RIC8A, MIR6743, SIRT3, PSMD13, NLRP6, ATHL1, IFITM5, IFITM2, IFITM1, IFITM3	doubleton to 1%
205	ToF	Gain	2p14	457,802	64107729	64565531	400K oligonucleotide	Paternal	UGP2, VPS54, PELI1, LINC00309	doubleton to 1%
218	ToF	Loss	2p13.2	56,753	73285631	73342384	400K oligonucleotide	Paternal	SFXN5, RAB11FIP5	1 singleton observation
188b	ToF	Loss	8q24.21	2,879	129108394	129111273	400K oligonucleotide	Maternal	PVT1	0%
193a	ToF	Gain	3q13.13	277,172	108626684	108903856	400K oligonucleotide	Paternal	GUCA1C, MORC1, FLJ22763, LINC00488	1 singleton observation
167	ToF	Gain	8q24.12	551,841	120628587	121180428	400K oligonucleotide	Maternal	ENPP2, TAF2, DSCC1, DEPTOR, COL14A1	0
184	ToF	Gain	Xq25	361,263	123509834	123871097	400K oligonucleotide	Maternal	TENM1	0%
148	ToF	Gain	11p13	187,151	33191299	33378450	400K oligonucleotide	Paternal	CSTF3-AS1, HIPK3	0%
146a	ToF	Gain	2p21-p16.3	398,949	47537252	47936201	400K oligonucleotide	Maternal	AC073283.4, EPCAM, MIR559, MSH2, KCNK12	0%
149a	ToF	Loss	2q21.2-q22.2	7962,887	134275082	142237969	400K oligonucleotide	Maternal	NCKAP5, MIR3679, MGAT5, TMEM163, ACMSD, CCNT2-AS1, CCNT2, MAP3K19, RAB3GAP1, ZRANB3, R3HDM1, MIR128-1, UBXN4, LCT, LOC100507600, MCM6, DARS, AC093391.2, CXCR4, THSD7B, HNMT, SPOPL, NXPH2, YY1P2, LRP1B	0%
153b	ToF	Gain	1p34.1	12,36	46072235	46084595	400K oligonucleotide	Paternal	NASP	0%
5	ToF	Gain	3p25.3	132,049	8919991	9052040	400K oligonucleotide	Paternal	RAD18, SRGAP3	0%
29b	ToF	Loss	16q23.2	211,72	81134814	81346534	400K oligonucleotide	Paternal	PKD1L2, BCMO1	0%
36	ToF	Gain	21q22.3	165,777	47690060	47855837	400K oligonucleotide	Maternal	MCM3AP, YBEY, C21orf58, PCNT	0%
43	ToF	Gain	1q21.2-q22.2	2090,029	145761268	147851297	400K oligonucleotide	Maternal	NBPF20, NBPF10, PDZK1, GPR89A, NBPF25P, GPR89C, PDZK1P, NBPF11, NBPF12, LOC728989, NBPF13P, PRKAB2, PDIA3P1, FMO5, CHD1L, LINC00624, BCL9, ACP6, GJA5, GJA8, GPR89B, PDZK1P1, NBPF8, NBPF11, MIR5087	0%
84	ToF	Gain	2p22.3	234,773	32815822	33050595	400K oligonucleotide	Maternal	BIRC6, TTC27, MIR4765, LINC00486	doubleton to 1%
93b	ToF	Loss	21q21.1	1570,572	22115159	23685731	400K oligonucleotide	Paternal	LINC00320, NCAM2, RNU6-67P, LINC00317, AP000475.2, LINC00308	0%
106	ToF	Gain	15q21.3	64,885	57689066	57753951	400K oligonucleotide	Paternal	CGNL1	doubleton to 1%
217	ToF	Gain	4q21.1	68,634	77033621	77102255	400K oligonucleotide	Maternal	ART3, NUP54, SCARB2	doubleton to 1%
214	ToF	Gain	15q13.1	1013,794	29212452	30226246	400K oligonucleotide	Maternal	APBA2, FAM189A1, NDNL2, TJP1	0%
153c	ToF	Loss	11p11.2	99,596	47993091	48092687	400K oligonucleotide	Paternal	PTPRJ	0%
53	ToF	Gain	1q21.3	358,831	150413398	150772229	400K oligonucleotide	Maternal	RPRD2, TARS2, MIR6878, ECM1, LINC00568, ADAMTSL4, MIR4257, ADAMTSL4-AS1, MCL1, ENSA, GOLPH3L, HORMAD1, CTSS, CTSK	0%
158a	ToF	Gain	18q12.1	112,124	28937235	29049359	400K oligonucleotide	Maternal	DSG1, RP11-534N16.1, DSG4, DSG3	0%
161	TGV	Loss	6p25.2-p25.1	1020,47	3983732	5004202	400K oligonucleotide	Maternal	PRPF4B, FAM217A, C6orf201, ECI2, RP3-400B16.1, CDYL, RPP40	0%

Table 3. ISCA and DECIPHER public databases collecting clinical information about chromosomal rearrangements. We found 6 described individuals presenting with CHD and carrying at least a CNV overlapping any of our patients' CNV regions.

Patient Data						ISCA/DECIPHER					
TRIO	CHD	Inheritance	CNV	chr	CNV size (Mb)	N°	CHD	Inheritance	CNV	CNV size (Mb)	
437b	CoAo	father	duplication	9p24.2	0,38	1578	PDA	unknown	deletion	11,34	GLIS3
						nssv706487	CoAo	unknown	deletion	11,1	
						nssv706596	CoAo	unknown	deletion	16,7	
335a	TGV	father	duplication	19 p13.12	0,021	nssv580437	Dextrocardia	unknown	deletion	0,7	RLN3
172a	TGV	mother	duplication	6 q22.31	0,26	250627	TGV	unknown	duplication	0,85	PLN
188a	ToF	de novo	duplication	17 q12	0,22	nssv578742	TAPVR	unknown	duplication	6,4	ERBB2

Table 4. List of genes included in our CNVs regions that are predicted to have binding sites for FOXC1 transcription factor. Most of the selected CNVs found in CHD affected patients (54/69=77%) share binding sites for FOXC1 according to TFBS tool.

Trio ID	CNV			TFBS
	chr	start	end	FOXC1 Binding Sites
212	chr2	28649512	29174327	PPP1CB, WDR43, FAM179A
163	chr7	2222167	2241694	near to MAD1L1
6	chr10	104818399	104847082	CNNM2
179	chr7	39993801	40102842	CDK13
174	chr11	47826590	48303958	PTPRJ, near to NUP160
44	chr16	66728053	66851930	NFATC3,ESRP2
59	chr7	141351858	141387062	MGAM
23	chr2	20267100	20314989	PUM2
77a	chr13	75289358	75306459	LMO7
77b	chr22	36939738	36957238	MAFF, TMEM184B
211	chr3	142755432	142868290	RASA2
226a	chr15	83407124	83483235	PDE8A
437b	chr9	3771683	4151396	GLIS3
437c	chr9	5524835	5755850	KIAA1432
194b	chr15	28983999	30191315	FAN1, MTMR10, LOC283710, KLF13
231b	chr4	121835603	121917802	PRDM5
252b	chr8	100274271	100472962	VPS13B
265	chr2	9442831	9567698	ASAP2, ADAM17
277	chr3	199050618	199291875	LRCH3
335a	chr19	13999589	14020806	RLN3
351b	chr10	103180049	103332582	BTRC
358c	chr13	103959813	105607574	LINC00343, near to DAOA
172a	chr6	118878090	119137929	CEP85L
172b	chr8	53323590	53762104	ST18, RB1CC1
165b	chr15	39451144	39476678	NDUFAF1
284a	chr4	186795696	186862130	SORBS2
302	chr5	131275432	131340271	ACSL6
341	chr20	18110446	18453596	CSRP2BP, DZANK1, SEC23B
356	chr16	22264898	22338999	CDR2
393	chr5	96133154	96169574	ERAP1
294a	chr13	22133248	23413626	SACS, TNFRSF19, MIPEP, C1QTNF9B-AS1, C1QTNF9B
295	chr7	158024743	158304209	PTPRN2, NCAPG2, ESYT2
329a	chr12	110666925	110793312	MAPKAPK5
188a	chr17	35066780	35286624	ERBB2, GRB7, IKZF3
24	chr17	28881049	29622308	ASIC2
205	chr2	63961233	64419035	PELI1
218	chr2	73139139	73195892	RAB11FIP5
193a	chr3	110109374	110386546	MORC1
167	chr8	120697768	121249609	ENPP2, TAF2, DEPTOR
184	chrX	123337515	123698778	ODZ1
148	chr11	33147875	33335026	HIPK3
146a	chr2	47390756	47789705	KCNK12
5	chr3	8894991	9027040	RAD18, SRGAP3
36	chr21	46514488	46680265	YBEY, PCNT
43	chr1	144472625	146317921	GPR89A, GPR89C, NBPF10, NBPF20, NBPF25P, LOC728989, FMO5, CHD1L, BCL9, ACP6, GPR89C, GPR89B
84	chr2	32669326	32904099	BIRC6, TTC27
93b	chr21	21037030	22607602	NCAM2
106	chr15	55476358	55541243	CGNL1
217	chr4	77252645	77321279	NUP54
214	chr15	26999744	28013538	APBA2, FAM189A1, TJP1
53	chr1	148680022	149038853	GOLPH3L, HORMAD1, CTSS, CTSK
158a	chr18	27191233	27303357	LOC101927718, near to DSG4
161	chr6	3928731	4949201	PRPF4B, C6orf201, CDYL

Supplementary data

Supplementary Table 1. List of 297 candidate genes. These candidate genes were chosen because they are reported to have a role in heart development in humans as well as in animals. Therefore, they were densified using 100.000 extra oligonucleotide probes in the custom aCGH with the aim to be able to detect any rearrangement in these interesting regions.

Gene name	Mapping (hg19)	Gene name	Mapping (hg19)
MTHFR	chr1:11845786-11866115	Notch1	chr9:139388895-139440238
NPPA	chr1:11905768-11907840	Itgb1	chr10:33189245-33247293
NPPB	chr1:11917520-11918992	NRP1	chr10:33466418-33623833
Ece1	chr1:21543739-21672034	Cdc2	chr10:62538219-62554604
Hspg2	chr1:22148736-22263750	NODAL	chr10:72191691-72201465
WNT4	chr1:22443797-22469519	NRG3	chr10:83635069-84746935
Id3	chr1:23884408-23886322	Bmpr1a	chr10:88516395-88684945
GJA4	chr1:35258598-35261348	PTEN	chr10:89623194-89728532
PTCH2	chr1:45288086-45308616	HHEX	chr10:94449680-94455408
TGFBR3	chr1:92145901-92351787	NKX2-3	chr10:101292689-101296280
VCAM1	chr1:101185296-101204601	Lbx1	chr10:102986732-102988717
Rbm15	chr1:110881944-110889303	Fgf8	chr10:103529886-103535827
NOTCH2	chr1:120454175-120612276	Fgfr2	chr10:123237843-123357972
GJA5	chr1:147228331-147245484	HRAS	chr11:532241-535550
GJA5	chr1:147228331-147245484	Tead1	chr11:12695968-12966299
Shc1	chr1:154934773-154946959	Sox6	chr11:15987995-16497935
VANGL2	chr1:160370366-160398464	MYOD1	chr11:17741109-17743678
Pbx1	chr1:164528801-164821045	Csrp3	chr11:19203577-19223589
RXRG	chr1:165370349-165414430	WT1	chr11:32409324-32457087
Tnnt2	chr1:201328141-201346805	NR1H3	chr11:47270448-47290401
csrp1	chr1:201452659-201476387	Mybpc3	chr11:47352956-47374253
Prox1	chr1:214161859-214209762	Ptprj	chr11:48002109-48192394
Cenpf	chr1:214776531-214837914	FGF19	chr11:69513005-69519106
TGFB2	chr1:218519390-218617959	Fgf4	chr11:69587796-69590171
LEFTY1	chr1:226073981-226076836	Wnt11	chr11:75897369-75917574
LEFTY2	chr1:226124302-226128920	SLN	chr11:107578100-107582787
Mixl1	chr1:226411382-226413513	ETS1	chr11:128328655-128457453
WNT3A	chr1:228194751-228248961	CCND2	chr12:4382901-4414521
MTR	chr1:236958580-237067281	NTF3	chr12:5541279-5604465
Id2	chr2:8822112-8824583	Kcnj8	chr12:21917888-21927747
MYCN	chr2:16080682-16087129	KRAS	chr12:25358179-25403854
Osr1	chr2:19551245-19558372	Col2a1	chr12:48366747-48398285
SOS1	chr2:39208689-39347604	WNT1	chr12:49372235-49376395
SLC8A1	chr2:40339285-40739575	ErbB3	chr12:56473891-56497128
meis1	chr2:66662531-66799891	Stat6	chr12:57489192-57505161
Bmp10	chr2:69092612-69098649	Cdk4	chr12:58142002-58146164
TGFB3	chr2:75640686-75664337	Irrc10	chr12:70002344-70004942
SMYD1	chr2:88367381-88412902	NR1H4	chr12:100867678-100957643
Gli2	chr2:121554866-121750229	ATP2A2	chr12:110719031-110788897

CFC1B	chr2:131278666-131285565	MYL2	chr12:111348623-111358404
CFC1	chr2:131278835-131357082	Myl2	chr12:111348623-111358404
Acvr1	chr2:158592958-158732374	PTPN11	chr12:112856535-112947717
Dlx2	chr2:172964165-172967478	TBX5	chr12:114791734-114846247
TTN	chr2:179390717-179672150	TBX3	chr12:115108058-115121969
ITGA4	chr2:182321618-182402468	MED13L	chr12:116396382-116714991
Calcr1	chr2:188207848-188313021	NOS1	chr12:117650978-117799582
Col3a1	chr2:189839098-189877472	IFT88	chr13:21141207-21265576
Casp8	chr2:202098165-202152434	Fgf9	chr13:22245214-22278640
Bmpr2	chr2:203241049-203432474	Flt1	chr13:28874482-29069265
ErbB4	chr2:212240441-213403352	SMAD9	chr13:37422206-37494409
Wnt6	chr2:219724545-219738954	APEX1	chr14:20923289-20925926
Pax3	chr2:223064606-223163700	NDRG2	chr14:21484921-21493935
Oxtr	chr3:8792094-8811300	MYH6	chr14:23851198-23877482
CRELD1	chr3:9975505-9987090	MYH7	chr14:23881946-23904870
RAF1	chr3:12625099-12705700	Nfatc4	chr14:24836144-24848810
TGFBR2	chr3:30647993-30735633	NKX2-8	chr14:37049215-37051786
ACVR2B	chr3:38495789-38534633	SIP1	chr14:39583487-39606177
Myl3	chr3:46899356-46904973	Bmp4	chr14:54416454-54423554
FlnB	chr3:57994126-58157977	MNAT1	chr14:61201469-61435398
Foxp1	chr3:71004735-71633140	Hif1a	chr14:62162118-62214977
Nphp3	chr3:132399453-132441276	MTHFD1	chr14:64854758-64926725
SOX14	chr3:137483578-137484396	SLC8A3	chr14:70510933-70655787
ATR	chr3:142168076-142297668	dpf3	chr14:73136659-73360809
SHOX2	chr3:157813799-157823936	Psen1	chr14:73603142-73690398
prkci	chr3:169940219-170023770	LGMN	chr14:93170154-93215012
DVL3	chr3:183873283-183891314	rtf1	chr15:41709301-41775760
Ece2	chr3:183967444-184010819	FBN1	chr15:48700504-48937918
Fgfr3	chr4:1795038-1810599	Aldh1a2	chr15:58245626-58357906
Msx1	chr4:4861391-4865660	Smad6	chr15:66994673-67074335
EVC2	chr4:5564151-5710294	STRA6	chr15:74471809-74501371
EVC	chr4:5712923-5816031	NRG4	chr15:76235843-76304785
NKX3-2	chr4:13542453-13546114	Chrna3	chr15:78887651-78913322
RBPJ	chr4:26321331-26433278	Mesp1	chr15:90293099-90294540
Pdgfra	chr4:55095263-55164412	NR2F2	chr15:96869156-96883490
Pitx2	chr4:111538579-111558508	IGF1R	chr15:99192760-99507759
BBS7	chr4:122745634-122791642	Mkl2	chr16:14165195-14360630
Fgf2	chr4:123747862-123819390	DOC2A	chr16:30016834-30022401
Gab1	chr4:144257982-144395717	TBX6	chr16:30097116-30103205
Smad1	chr4:146402950-146480325	Irx5	chr16:54965110-54968395
EDNRA	chr4:148401906-148466106	MMP2	chr16:55513080-55540586
TLL1	chr4:166794409-167024993	ndrg4	chr16:58497548-58547523
HAND2	chr4:174447651-174451378	Nfatc3	chr16:68119374-68260837
Casp3	chr4:185548849-185570629	CDH1	chr16:68771194-68869444
Irx4	chr5:1877540-1882880	Foxc2	chr16:86600856-86602535
MTRR	chr5:7869216-7901235	FOXC2	chr16:86600856-86602535
isl1	chr5:50678957-50690563	Zfp1	chr16:88520013-88601574
Hmgcr	chr5:74632992-74657924	CYBA	chr16:88709696-88717457
dhfr	chr5:79922044-79950800	SNAI3	chr16:88744089-88752882
VCAN	chr5:82767529-82877800	DVL2	chr17:7128660-7137863
MEF2C	chr5:88014058-88199869	Dvl2	chr17:7128660-7137863
Kif3A	chr5:132028322-132073265	MYH10	chr17:8377529-8534036
Smad5	chr5:135468535-135518422	MAPK7	chr17:19281033-19286856
NRG2	chr5:139227259-139422879	NOS2	chr17:26083792-26127555

Adam19	chr5:156904311-157002768	Adap2	chr17:29248753-29286211
FGF18	chr5:170846666-170884164	NF1	chr17:29421994-29704695
NKX2-5	chr5:172659137-172662262	MED1	chr17:37560537-37607527
Msx2	chr5:174151574-174157902	ErbB2	chr17:37844392-37884915
Foxc1	chr6:1610680-1614129	Hdac5	chr17:42154120-42201014
Bmp6	chr6:7727010-7881961	FZD2	chr17:42634924-42636907
Edn1	chr6:12290528-12297427	Hexim1	chr17:43224683-43229468
JARID2	chr6:15246526-15522253	MYL4	chr17:45286427-45301045
Sox4	chr6:21593971-21598849	HOXB6	chr17:46673098-46682334
Pbx2	chr6:32152509-32157963	Dlx3	chr17:48067368-48072588
RXRb	chr6:33161364-33168432	NOG	chr17:54671059-54672951
Srf	chr6:43138919-43149244	TBX2	chr17:59477256-59486827
TFAP2B	chr6:50786438-50815326	MAP3K3	chr17:61699800-61773670
Bmp5	chr6:55620236-55740375	Prkca	chr17:64298925-64806862
TBX18	chr6:85444156-85473899	SOX9	chr17:70117160-70122560
PLN	chr6:118869441-118881587	Jmjd6	chr17:74708913-74722881
GJA1	chr6:121756744-121770873	GATA6	chr18:19749415-19782227
HEY2	chr6:126070731-126082415	DTNA	chr18:32073253-32471808
Eya4	chr6:133562494-133853258	Smad2	chr18:45359465-45457515
Cited2	chr6:139693396-139695785	Smad7	chr18:46446222-46477081
MAP3K7IP2	chr6:149639062-149732747	NFATC1	chr18:77155771-77289323
Sod2	chr6:160100148-160114353	insR	chr19:7112265-7294011
TWIST1	chr7:19155090-19157295	Smarca4	chr19:11071597-11172958
TBX20	chr7:35242041-35293242	GDF1	chr19:18979360-19006953
Gli3	chr7:42000549-42276618	Megf8	chr19:42829760-42882921
MYL7	chr7:44178462-44180916	SLC8A2	chr19:47931279-47975307
ELN	chr7:73442426-73484236	NR1H2	chr19:50879684-50886267
Sema3C	chr7:80371853-80548667	Tnni3	chr19:55663135-55669100
Sema3d	chr7:84624871-84751247	Bmp2	chr20:6748744-6760910
SRI	chr7:87834431-87856308	JAG1	chr20:10618331-10654694
Krit1	chr7:91828282-91875414	Ovol2	chr20:18004795-18038521
cav1	chr7:116164838-116201230	Id1	chr20:30193091-30194313
BRAF	chr7:140433812-140624564	KIF3B	chr20:30865466-30922811
NOS3	chr7:150688143-150711686	ift52	chr20:42219578-42275862
SMARCD3	chr7:150936058-150974231	SNAI1	chr20:48599526-48605420
SHH	chr7:155595557-155604967	SALL4	chr20:50400582-50419048
GATA4	chr8:11561716-11617509	Bmp7	chr20:55743808-55841707
Dlc1	chr8:12940871-13372395	Gata5	chr20:61038552-61051026
NKX3-1	chr8:23536227-23540450	Adamts1	chr21:28208605-28217728
NKX2-6	chr8:23559964-23563922	TBX1	chr22:19744225-19771112
NRG1	chr8:31497267-32622558	Txnrd2	chr22:19863039-19929359
Fgfr1	chr8:38268655-38326352	CHEK2	chr22:29083730-29137822
SNAI2	chr8:49830240-49833988	Cby1	chr22:39052657-39069855
Sox17	chr8:55370494-55373456	Map3k7ip1	chr22:39795758-39833132
CHD7	chr8:61591338-61779467	BCOR	chrX:39910498-40036582
ZFPM2	chr8:106331146-106816765	Cited1	chrX:71521489-71527037
ZFPM2	chr8:106331146-106816765	apln	chrX:128779324-128788914
Foxh1	chr8:145699116-145701718	ZIC3	chrX:136648345-136654259
TEK	chr9:27109146-27230172	FlnA	chrX:153576899-153603006
Barx1	chr9:96713908-96717608	TAZ	chrX:153639876-153650061
PTCH1	chr9:98205263-98279247	hsa-mir-1-1	chr20:61151513-61151583
TGFB1	chr9:101867411-101916473	hsa-mir-1-2	chr18:19408965-19409049
INVS	chr9:102861510-103063426	hsa-mir-133a-1	chr18:19405659-19405746

Pbx3	chr9:128509616-128729653	hsa-mir-133a-2	chr20:61162119-61162220
ENG	chr9:130577290-130617047	hsa-mir-133b	chr6:52013721-52013839
RXRA	chr9:137218315-137332432		

Supplementary Table 2. List of primers designed for qPCR validation of the rare CNV selected regions obtained from aCHG. All primers were designed with Primer3-PCR primer design tool v.0.4.0.

Gene name	Forward	Reverse
Albumine	GCTGTCATCTCTTGTGGGCTGT	AAACTCATGGGAGCTGCTGGTT
TNNI3KF	TCACTTCTCTGCTTCACAGTGG	TGTTTTACAGGTGCCAAAG
PLB1	GCATCCTTTGCCTACAGCTC	AGGGCTATTTTCCCCTTGAA
PPP1CB	GGTCTCTCACAAGGCTCCAA	AATGAAGCATCCCACGTAGC
MAD1L1	TTCCATGGTTGCTTTCCTTC	TGAGCTCCAATGTGCGTTAG
FTSJ2	TTGAAGAATGCGTGAATTGC	AACCTGGGCAACATAGCAAC
NT5C2	AAGAGTGGGGGAAGGAGAAA	ATCGACCATGGCAGTCCTAC
CNNM2	AGCAGGTGGAGAGTCCAGAA	ATTCAAGGTCAGCCACAAGG
CCDC76	TGGCTGCTTAAGCTGTTGTG	ACCTGCTGTTTCAGGTGCTT
SASS6	CAAGCTGGTCCTTATCCTTGA	GGAGCTACAGCGGACTAAGC
CDC2L5.1	CAATTGGGTGGCGAGTAGAT	ATTTATCCACGCAGCGTTTT
CDC2L5.2	TCTGAAAGTGTGGAGGCAA	GAGCACTGCTTTCACCCATC
PTPRJ.1	GGGAATGGTGACTGAAATGG	TCGAAAACACCTCCCTAAC
PTPRJ.2	TGTGTCCTTCCTTCCCTTTG	TCCAACTAACCACCTCTG
ATM	ATTGGTTTGAGTGCCCTTTG	CCTGTTGCATGACAGCATCT
KDELC2	GGCTTTTCCAAGCTCCTTTT	GGTAGAGACAGCCGAGAGGA
CPNE8.1	TTGTTTCGCACAAAATTGCAT	CAAACCTGGCCATTTGCATT
CPNE8.2	TGCATGGCTAAACTGCTCAC	CAGGGTCCTATGGTTGCACT
GABRG3.1	AGTGCCTAGACCCCCAGATT	TATGGTGACTGCATCGTGGT
GABRG3.2	TTTCCACCATGTCCTTGTGA	CGTCAATTACGGTCGGTTTT
NFATC3.1	TGCTCCATCCCACAAACATA	GCTGATAGGAAGACCCCA
NFATC3.2	AACACCAACAGATGCCCACT	ACACCTGCCCCTCTGTAATG
MYO19.1	AGGGTCCTATTGGGAAATGG	CCAATGGACCTTCTCTTGGA
MYO19.2	CACTGTTGAGCCTGTCTGGA	TGTGTGACTCTGGGCAGTTC
MAD1L1	AAAACCCAGGAGAGCTGTGA	TTGTCCTGCATCACAGAAGG
FTSJ2	TCAGGAGCTCAGGCTACGTT	GCCAGCAGGAAAGAGTCATC
LMO7.1	CCCCTGCTGATGTGATTTTT	CTGTCCTGCAGCATTTCTT
LMO7.2	AATCTGGGCTGAAATGTGC	AGCCAAATCCCAAGGAGACT
MGAM.2	CCTTCAGTGGAAATGCTGCT	CCCCCTTCTTCATGCTTACA
XRN1.1	CAGGGTTGGAAGGAAAACAA	TGATTTGAGGTTCTCTACCG
XRN1.2	TGTCTGCTCTGCCATGTTTC	TGCAGGAGCTTCTTTCTCTA
CENPP_OGN	CTGCTGCACTCAGTCTGCTT	CCTGGGAAATCACCAAATG
CENPP_OMD	CAAACCAGGAGGCATTGATT	TCCAAACTGCAGACAAATGC
KANK1.1	ATCCCTTCAGTGGCTTCGTA	GCATGTCCTTCTCTGCTTCC
KANK1.2	GGCTTGCCTCATTCTCTTC	TGGGTTTCTCAAACCTCCAG
SDC1.1	GGGAGCTGTTATTGGGAACA	TTGTTCTTGGGGGTTTTTGA
SDC1.2	TCAGGCTATGAGGGAGGCTA	CAGAAAAGCTTTCACAGGTG
NR2C1.1	GCTTTCCTGAAGGTAGGCTTT	CAGCAGAGGGGTAATCACCTT
NR2C1.2	CCTGCCCAGTCACTTCAAA	GCCAGAACACAAGACACCAA
KLHL12	GGGGGAATCTGTTGCCTTAT	GCAGAAAGGCGATAATGCTT
ADIPOR1	AGGACCTGCTGGAAGATTCA	CTATCGCTGAGGGCTTTGTC
DYDC1.1	TGAGTTAAACAGGCCCAAG	GAAGTGGGAGGGGTGTATG
DYDC1.2	ATCCTGGCCTAGCCTCATTT	TGGAGCACCTAAGTTGAGCA
NME6.1	TCCATTGAAGGGTGAGAAG	GTCTGAGTGGTGGGTCTGT
NME6.2	CCTCCAGAATCAGTGGATGG	ATGTAGGGAAGGCTGGGAGT

LTN1.1	TCCCCTCTATTGAGCCAATTT	CATTTGGCTTCTCTGGGTGT
LTN1.2	AGGCTGGCTCTTCTTCTTCA	TTCTGCTCTTGAGCACGACTT
08-oct	CAATTTCTGTTGCTCCACCTC	TTGGTGGGGTACTGGAATGT
ETS1.1	AGCAGAGTGGCTGGAGAGAG	TGCTTAGGCCAACTCCATCT
ETS1.2	AGGCTGTATGGGCACATTTT	ACACCTCCTGCCTGTACTGC
EML4.1	ATGGGAGTTTTGTGGGACAT	TGTCCTTCTCCCAATTCTG
EML4.2	TTATGAAGAGCCATGCAACG	CACTGCCCATCCTGTTACCT
C1orf25.1	CTGTCCACCACCACTTTCAA	ATGGGATTACAGTGGGCAAA
C1orf25.2	ATTCTGGAGGTGACGACGAA	GCCCCCATTAACTCTTTGT
NEXN	AAAACCTTGGAAGGGTGATG	CCTGCTTTCTCCTGTTCCAT
COCH	CCCAAAGTGGTGGTGGTATT	GAACCATCCCCAGTTCTTCA
TTC19	CCACCGTCAGTCTGGAAGTT	CCCTGGCTGCTGTACTTACC
CDK13.1	CCTGAGCTACCAGGAGGAGA	CCAGAATCACTGCTGGTCAA
DNAJC2	CCACCTCCACCAGTTGATTT	GCCCTAGCCAGGAAGAGAGT
TMEM38B.1	TCATGACTTTTTTGGCAGATGA	CAATTTGGCAAAACACAACAA
TMEM38B.2	GTTTGGGATGGAAGGAGACA	AATGGGTGATTTTTTGGCACT
SH2D4B.1	GGGACAGCCCTTGACCTTAT	AGAAGGCTCTTTCCAACACG
TMEM184B.1	GAGGGAACCTGTTCAATCCA	CCCAGATTATCCCTGGGTTT
PTPRN2.1	GCTCAATGCCAGATTTGTT	ACTCCCCTGACCTCACCTTT
PTPRN2(2).1	GGTCCCCAAGGGAGATGTAT	CAATTCCCTGCCAGTTGTTT
SLN.1	GCCTACTGGGATCAACCAAA	TCTAGGCACATCCTGGACCT
RASA2.1	TGAAGCTGTTTTGTGGTGTGA	ACCAGGCTTGATGAGAGGAA
PDE8A.1	TGACGTCAACATGGGTTTGT	CAGGCACATTTTCAGGCACTA
PGAM1	TGCTTCACTGTGGGCTTCTA	GGCCCCCACCATACTTATTT
TM2D3	GGCTTTTTTCTAAGCCCTGCT	CTGGGAATATGGACGCTGAT
PRKCD	CCATGAAAACATTGCTGTGG	CTGTTACTCCGAGGCTCCAG
RNF115	TAATGGGCTCCATCTGCTCT	TGAAGCTACCTCACCGTTCC
GLIS3.1	CTTGCCACACACACAAATCC	GGTTTTACGCCACAAAATG
KIAA1432	TGGGAATGACCGACACTTCT	TGGGATAGGCAAAAGGGATT
WDR72.1	CTTTCAGGGCAAACCTGTGCT	AGGTTCTCTGTTTTGCTGTG
PRPF4B	GCAGGTCTCCACTTTTGAATG	TCCTCTCTGGTTCTCGTCGT
ADH1B.1	TACGCTCCATGCAAAGACTG	ACCTGCTTCACTCTGGGAAA
TDRD3.2	GCCTGCATCATGAACACTTG	TGCAGTGGCATAAACTTTGAA
TRPM1	AGGCTGGGGAAAGCATCTAT	CCTGTGGCTCTTTCCAGAAG
PVT1.1	ATCCCTTCTGAATGGCACAC	AGGAAACATGGCAAGACTGG
CIITA	CCGACACAGACACCATCAAC	GCAAAGAACTCTTGCCCTTG
VPS13B.1	TCAGCCTCAGAAACGAACAA	GGTGGGGAGACAAAGGAAAT
FMR1-AS1	CAAGGACCCTGTAGGGACTG	CATAGAGGCCCTGCACCTTA
CAMSAP1L1.1	ACTGGAATTCACGTTCTTCA	TGGTGTGCGCAGTTTCAATA
PRKACB.1	TGGGAGAATCCAACCTCAGGT	GGCAAAATTCCCATTTTAAGC
CSPP1	AAACCGAAACCTTCCAAACA	TCTAAGGGACCACCATGCTT
ARSB	TGACACCCACTGGTGAAAGA	GCCATTTAAATCTGGGGTGA
CSMD1.1	TAATGTTCCCCCTGGATCTG	GCCCTTTCCAATCGGTAGT
FBN1.1	TGGGGTCCAGTAAATCCGTA	GTCCTTCCAGAGGACCACAA
WDR33.1	GCTGCAAGACACTGCATGTTA	GAAAGAGGCATGGAGCAGAG
PLEK.1	GGAGGCGAATGGAGTACAAA	TAGCCCCTGCTTTTTCTTTT
BTRC.1	GCACCAAATCAACTGCTCAA	CACTACCAGCCTGTCCCTGT
PLN_C6orf204	CTGTTCCCATAACTGGGTGA	CCAAAGTCAGCGAAATCTGTT
FAM150A	CTTCCCACCCACAGACATTT	TTCACCAGAATGCAGCAAAC
NDUFAF1	GCCTCACCTACCCTTGGAAT	AAGACGATTGGAGGCAGAAG
TNFRSF19	TGGTAGCAAAAGGACCTTGG	GGGTATCCCATTTTCTTGGA
NEUROG2	GCCCGTCTGAATGAAGGATA	GCCTTAGTGGGTTCCCTTTC
ICA1.1	TGCCGAAAAGTCTCCACTTC	AGATTGGCAGAATGTCAAGGA

AGBL2	GGCATCCAATTGGGAGTACA	TGAACTCACCTTGCGAACTG
SORBS2.1	CCCCTCAATGATTTCCCATA	AGCAGGAAGGGAAAGGAGAA
EPHA5.1	GGCACAATTGGTTCCACTG	CCCTGGCAGGATCTCATTTA
ACSL6.1	TGATTACGTGGGCCTCTTTC	TTTGGGTTCTGGATCTGGAG
SGK196.2	CGTTTGGACTATGGCGATTT	CTCCAGCTCAGGTCTCCATC
VPS33A	TGCAAACCAATGATTCAGGA	TTTGGCCTGATTGCACAGTA
IQCG	AATCTGCAGCTCGGTATTGG	CGCTGACTGAACTGAATGGA
ESYT2	GTGTGGATGGAGCTGTGTTG	CCTCTAGGTGCTCCATCTCG
ZMYM4	CTCTCGCATTGAGGAAGAGC	TTCAGAGTCCTGGTCCGTCT
TRIM45	GGCATGCTTCCTAGCAGAAC	GATCCAGCCAAATGTGTCCT
PEX1	AACCCTCATGAGTTCCCAAA	GCGCTTTCCAAGAAAATGAC
PNPLA4.1	CGCTGCAAATGATAGGTTGA	AGAAAGCCTTGGAACGATGA
RLN3	GCTGGAAGTCTGAGGACAGG	CGGAAGAAAAACCAACTTGC
POLR3F	CCACGAAGGTGGTGAGATTT	TGCTTTCAAGGAGATTTTCGTG
CHRN3.1	CTGAGTCCCCAGTAGGGTCA	TTGTGGTCTGTCCATTCTCG
TNFRSF10A.1	AGGAAGCACAGGGCTACAGA	GCCCACAGTGTAAGTCCAGGT
ACAD10	TCACTGATGTGCCTCGAAAG	AGACTCCAGGCTTCTTGACAG
FAM47E	AGACCAAACCAAGCCATGAT	TCCCCTTCTGCAATTACCTG
CCNI.1	ATTCAGGTGGCAGGAAAATG	GGAAAGTTGGTGAGAGTGCTG
KCTD7.1	GAGTGTCCGCTCCTCAACTC	GCAGTGCAGCAGGTCCATAAA
RERG.1	ATCCCAAGATGAACCGTTTG	AGGGTTGTGTGGAGCATTTT
DAOA.1	CTGTTGCAGCAAAGGAGACA	ACTCTGCATAGGGCTGAGGA
PIK3C3.1	CTGAAGGCCACTCTCAAACC	TGCCTGGATGTTGACAAGAA
ERAP1.1	TCTGTACGCACGGCTGATAG	CTGGTCCCTGTTTCCCTGTA
GFPT1	TCAGCTTTTGCCAAGATATGC	TGCCTGTGATGGTGGAATA
CDR2.1	TTTTGAATGGACCTTGAGC	TCCCCTTAGGAGTCCAGGAT
ATG5.1	CATGAAAAACCAAGGCTGA	TTGGCTATATCCTGGCTTGG
VIP.1	TGTTGCAAATGCATTAGCTG	TCGGAATAGCTTCATGGTGA
ARNTL2	ACGAAAGCTGGTTTGCAAGT	TTGAGTTGGGTAAGCATCCA
TNRC6B.1	CCCTTCTTCCAGGTGTGTGT	ACGGGATCTGACTGGTTCTG
NFE2L3	CTAAAGCCCCAACCAGTTCA	TCAGGCTGTGATGAAAGCAA
CHEK2.1	TCCGAAAGTGTTTCTTGCTG	ATTCAACAGCCCTCTGATGC
TCTE3	TTAATTTTGCCAGGGCACTC	GCCTTTCTCTTCCCTCT
WDR41	AAAAGCAGGAGTTGCTTGGA	GCCAGCCAACTGTAAAAA
RAP1GAP2.1	TAAGGAGAACGCCTTGTTGG	AAGGTCCTCCCTACCCAGAA
PRDM5.2	GGAAGCCTCCATGTCAGAAA	CACAGTCCACTGCTGTCTCTG
EXOSC9_CCNA2	ACACAGACCACAGTGCAAA	CCAGGGTTCTCAGAATGGAA
ASAP2	TCAAGATCCTGCCTCTACGG	TCAGAGCTGGGCTATGGACT
TBX20.1	TATGCCATTCCCAAACATCA	TATTTGAGGGCACGTGGAAT
VPS54	AACAACCTTGACGTGCACCAG	GGCCATTGAATGAGAGGAAA
RAB11FIP5.1	AGGTGCGACACACACTTCTG	GGTGGCAGCTGTTTTGTTTT
IKZF3	TGCCACATTGCTTGCTAATC	CCTTTTCTCCCCATGTACCA
PVT1.1	TGCTCTGACGTCCTCATTTG	ATCTGGGAGCCCGTTATTCT
MORC1	TGCTGCAAAGCAGACAGAGT	ATCCCCAGTATTCCAGGTC
DSP	ACGAAGACCACCATCAAGGA	GCTGTCATTGAGCCTGACAA
ENPP2	AGCCAAGGCCTGGATAAAAT	ATCAGCAAATCTTCCCCAAA
EPCAM.1	TGAGTTCCATGGCAGATCAC	TCCACTGGAGTATGCCTTCA
ODZ1.1	TACCAATGAGGCCCACTAC	AGGGCTTTTTACCCAGAGGA
SFXN1	GAACAGGGCAAAGAGAGCAC	CGGCAAGGATCATGAAGAAT
DGS4.1	TCATTGTCCAAGAGGCACAG	TCACGTTCCCTTTTGAGCTT
ZFP106.1	CCATGCTGTCCCAGAGAAGT	CCAGGCATACCCCTAGCATA
HIPK3	AAGGCATTGCTTGTGGAGAG	GCACACCCAGGCAGTGTAG
KCNK12	ATGGGAGGAGGATTGAGGTT	AGAGGAGACATGGGTGGTTG

NUP85	GACCTGTTTCTGGGAGGTGA	GCCACTCTAGAAACGCAAGG
MGAT5	ATTGCTTTGGGATCATCTGC	CAAGACAGTGTGCCCTCTCA
TRAP1	GTGAGCATGTGAGCTCCTGA	CCAGAGCAGGTGTGAACTGA
NASP.1	TAACCGGGATATGCAAGAGC	GCCCTTTCTAATCCAAAGG
PTPRJ(2).2	GACGGTCAAAGCAGAAGAGG	CGTCCATTTCACTGGACCTT
RAD18	ATTTGGCAAAGTGGGAGTTG	CTTTAGCTGGGTGCCTCTTG
ACCN1.1	AACTGGCTTCAGGTTGATCG	AAACACTCACCCCTGTCAGC
BCMO1	TTCTTGGCACTGAGGGAGTT	TTCCCAGGAAAGGTATGTGC
EVC.1	GGCATCACATGGACTGAGTG	CAGACTGCTGGAAGGAGAGG
CSRP2BP	GGCTTTCTTGACAGGAGCAC	GAAGGAGATGGGCTTGTCAG
PCNT	TAGTGCCTGTTCCCTCTTG	CACAGACCACCTGGAAACCT
GJA5	TACAGAGACCAGGCCAATCC	AGCAGGGGCAAGGAAATAGT
GPR83	TGCTGCCTTGTCTCATTCTG	GGATCCCCAGATGGATTTTT
DYDC1	GCCTAACTGACCCTGAGCTG	TGGAGTGGTCTTGGGAAAAC
BIRC6	GTGCGATTCAATCCAAACCT	AAAGAAACCCATGCCCTCTT
NCAM2	TTACATTGTGCTCGGATTGC	ATCGCATTTGCATGTGTTGT
CGNL1	GCAACCCTATGACCCTGAAA	GTTTGGGAAAGCAAGTCCAA
SPATA6	CAGGTGGAATGAAATGTTGCT	AAATATGCCCCCTTGGTGGT
GOLPH3L	CCACCAAATGGGAGAAAATG	ATTTGGGGAGAGCAACTTCA
EIF3B	TTGGATGTACCACAGGCTGA	TACTTCCGGAATGGATCTCG
TTC3	CGTACCGTTCTGGGATCTGT	ATCCCGTTCCAAGGGATTAG
ARPC4	GGCTTCTTTTCCCTGAGACC	ATTAATGGTGCCTGGACTGC
NUP54	CAGGGCCAGGGAATAAAAAG	TTCACTGCGTTTCTTCAGGTT
TJP1	TCGCTACAGCACTCATCTGG	CAACAGCATCCTTCCACCTT
STBD1	CCCCTGGGAGAACAAGTGTA	CTCCACAGTCCAGGCCATAC
PSMD13	CACAGATGCAGCCAGACCTA	TCACTACCCCTCCAGACAGG
DEFB119	TGACCCAGCCCTAGTGACTT	TCAGGTTGCTTAGGGTCCAC
TNN	ATCAAAGGATGTGGCCTGAG	CTCTCTTTCAGTGGGCCAAG
GMDS	ACAGCCAAACCATCATGACA	TGTCAAAGCCCAACAAGTCA
MAP3K4	CGTTTGATGTCTGGGGAGTT	GCCCACATGACACATTTCAG
EXOSC1	TGGGGATTACTAGGGCTTCA	TGGCCAGAGTGGCTCTTATT
TARSL2	GGGCAAAAGTGTCTTTTCA	GACTTCACGGAGGTGCTCTC
PRKCD	CCTAGGTCCTGTTTCGCTCAC	GTCCATGGAGTCGATGAGGT
FMR1-AS1	GACAGGACGCATGACTGCTA	ACTCGCCTTTCCTCAACTCA
DSG3	ATGAGCATTTGGCAGTCACA	TCCCCTTTCCACAAATTATCC
ETS1	GTGGGCAGGCAACTATCAAT	AAATGACTTTGTGGCCTTGG
TBX20	CCCCATCATAACCACATGTCT	TGGCTGCTGTTTTCTTACC
ECI2	TATGACTGGCAACGCTGAAG	TGATTGCAGTGGTCAATGGT

Supplementary table 3. Table of 71 validated rare CNV regions (50 duplications and 21 deletions) in children presenting with outflow tract defect. Generally CNV sizes ranged from 2,9kb to 1845,3kb, except for two much bigger CNVs of 7962,9kb and 14587,4kb (highlighted). Most of CNVs (63/71) were inherited from one of the parents, just 8 CNVs were found to be *de novo*. The frequencies of these CNV in general population went from 0% to 1% in some cases (according to the following databases: 42 million probes study (Conrad et al., 2010), WTCCC study <http://www.wtccc.org.uk/>, 1000 genomes project <http://www.1000genomes.org/> and DDD controls project <http://www.ddduk.org/>).

Trio ID	Mapping (hg19)	size	Inheritance	CHD	Rearrangement	freq in general population
212	chr2:28796008-29320823	524,815	<i>De novo</i>	CoAo	Duplication	0%
163	chr7:2255641-2275168	19,527	<i>De novo</i>	CoAo	Deletion	singleton
6	chr10:104828409-104857092	28,683	<i>De novo</i>	CoAo	Deletion	0%
179	chr7:40027276-40136317	109,041	Father	CoAo	Duplication	0%
174	chr11:47870014-48347382	477,368	Father	CoAo	Duplication	0%
44	chr16:68170552-68294429	123,877	Mother	CoAo	Duplication	0%
59	chr7:141705389-141740593	35,204	Father	CoAo	Deletion	0%
23	chr2:20403619-20451508	47,889	Father	CoAo	Duplication	0%
90a	chr1:78308983-78384008	75,025	Father	CoAo	Duplication	0%
77a	chr13:76391357-76408458	17,101	Mother	CoAo	Deletion	0%
77b	chr22:38609792-38627292	17,5	Father	CoAo	Duplication	0%
211	chr3:141272742-141385600	112,858	Mother	CoAo	Deletion	0%
226a	chr15:85606120-85682231	76,111	Mother	CoAo	Duplication	singleton
437b	chr9:3781683-4161396	379,713	Father	CoAo	Duplication	0%
437c	chr9:5534835-5765850	231,015	Father	CoAo	Duplication	0%
255	chr4:100226779-100242071	15,292	Mother	TGA	Duplication	0%
194b	chr15:31196707-32404023	1207,316	Mother	TGA	Duplication	doubleton to 1%
231b	chr4:121616153-121698352	82,199	Father	TGA	Deletion	0%
252b	chr8:100205095-100403786	198,691	Mother	TGA	Deletion	0%
210	chr15:48883414-48900202	16,788	Father	TGA	Duplication	0%
196a	chr2:128476795-128484318	7,523	Mother	TGA	Deletion	0%

196b	chr8:2975904-3226799	250,895	Father	TGA	Deletion	0%
265	chr2:9525380-9650247	124,867	Father	TGA	Duplication	0%
277	chr3:197566221-197807478	241,257	Father	TGA	Duplication	doubleton to 1%
335a	chr19:14138589-14159806	21,217	Father	TGA	Duplication	0%
335b	chrX:7866965-7923963	56,998	Father	TGA	Duplication	0%
351b	chr10:103190059-103342592	152,533	Father	TGA	Duplication	0%
358c	chr13:105161812-106809573	1647,761	Mother	TGA	Duplication	0%
222a	chr10:103291033-103317035	26,002	Mother	TGA	Duplication	doubleton to 1%
172a	chr6:118771397-119031236	259,839	Mother	TGA	Duplication	doubleton to 1%
172b	chr8:53161037-53599551	438,514	Mother	TGA	Duplication	0%
297	chr4:122735142-122739763	4,621	Mother	TGA	Duplication	0%
165b	chr15:41663852-41689386	25,534	Father	TGA	Deletion	0%
284a	chr4:186558702-186625136	66,434	Father	TGA	Duplication	singelton
302	chr5:131247533-131312372	64,839	Father	TGA	Duplication	0%
341	chr20:18162446-18505596	343,15	Mother	TGA	Duplication	0%
356	chr16:22357397-22431498	74,101	Mother	TGA	Duplication	doubleton to 1%
417	chr6:170140381-170202588	62,207	Mother	TGA	Duplication	0%
393	chr5:96107398-96143818	36,42	Mother	TGA	Duplication	doubleton to 1%
294a	chr13:23235248-24515626	1280,378	Mother	TGA	Duplication	0%
295	chr7:158331982-158611448	279,466	Mother	TGA	Duplication	0%
329a	chr12:112182542-112308929	126,387	Father	TGA	Duplication	doubleton to 1%
342	chr8:42583985-42587697	3,712	<i>De novo</i>	TGA	Deletion	0%
161	chr6:3983732-5004202	1020,47	Mother	TGA	Deletion	0%
42	chr20:11247299-25834671	14587,372	<i>De novo</i>	ToF	Duplication	0%
188a	chr17:37813254-38033098	219,844	<i>De novo</i>	ToF	Duplication	0%
24	chr17:31856936-32598195	741,259	<i>De novo</i>	ToF	Deletion	0%
11	chr7:2400442-2420409	19,967	<i>De novo</i>	ToF	Deletion	0%
515	chr11:193816-320936	127,12	<i>De novo</i>	ToF	Deletion	doubleton to 1%
205	chr2:64107729-64565531	457,802	Father	ToF	Duplication	doubleton to 1%
218	chr2:73285631-73342384	56,753	Father	ToF	Deletion	1 singelton
188b	chr8:129108394-129111273	2,879	Mother	ToF	Deletion	0%
193a	chr3:108626684-108903856	277,172	Father	ToF	Duplication	1 singelton

167	chr8:120628587-121180428	551,841	Mother	ToF	Duplication	0%
184	chrX:123509834-123871097	361,263	Mother	ToF	Duplication	0%
148	chr11:33191299-33378450	187,151	Father	ToF	Duplication	0%
146a	chr2:47537252-47936201	398,949	Mother	ToF	Duplication	0%
149a	chr2:134275082-142237969	7962,887	Mother	ToF	Deletion	0%
153b	chr1:46072235-46084595	12,36	Father	ToF	Duplication	0%
5	chr3:8919991-9052040	132,049	Father	ToF	Duplication	0%
29b	chr16:81134814-81346534	211,72	Father	ToF	Deletion	0%
36	chr21:47690060-47855837	165,777	Mother	ToF	Duplication	0%
43	chr1:145761268-147851297	1845,296	Mother	ToF	Duplication	0%
84	chr2:32815822-33050595	234,773	Mother	ToF	Duplication	doubleton to 1%
93b	chr21:22115159-23685731	1570,572	Father	ToF	Deletion	0%
106	chr15:57689066-57753951	64,885	Father	ToF	Duplication	doubleton to 1%
217	chr4:77033621-77102255	68,634	Mother	ToF	Duplication	doubleton to 1%
214	chr15:29212452-30226246	1013,794	Mother	ToF	Duplication	0%
153c	chr11:47993091-48092687	99,596	Father	ToF	Deletion	0%
53	chr1:150413398-150772229	358,831	Mother	ToF	Duplication	0%
158a	chr18:28937235-29049359	112,124	Mother	ToF	Duplication	0%

Supplementary table 4. ISCA and DECIPHER public databases collecting clinical information about chromosomal rearrangements. We found 25 described individuals presenting with CHD and carrying at least a CNV overlapping any of our patients' CNV regions.

Patient Data						ISCA/DECIPHER				
TRIO	CHD	Inheritance	CNV	chr	CNV size (Mb)	N°	CHD	Inheritance	CNV	CNV size (Mb)
77b	CoAo	father	deletion	22q13.1	0,02	nssv706258	TGA	unknown	duplication	33,75
						250853	ToF	de novo	duplication	13,85
437b	CoAo	father	duplication	9p24.2	0,38	1578	PDA	unknown	deletion	11,34
						nssv706487	CoAo	unknown	deletion	11,1
						nssv706596	CoAo	unknown	deletion	16,7
252b	TGA	mother	deletion	8 q22.2	0,2	nssv706345	ToF	unknown	duplication	50,7
196b	TGA	father	deletion	8 p23.2	0,25	nssv706595	CoAo	unknown	duplication	22,9
335a	TGA	father	duplication	19 p13.12	0,021	nssv580437	Dextrocardia	unknown	deletion	0,7
335b	TGA	father	duplication	X p22.31	0,057	nssv582252	ToF	unknown	deletion	1,5
						nssv578385	Hypoplastic left heart	unknown	deletion	154,9
358c	TGA	mother	duplication	13 q33.2	1,65	4668	ToF	de novo	deletion	9,41
172a	TGA	mother	duplication	6 q22.31	0,26	250627	TGA	unknown	duplication	0,85
294a	TGA	mother	duplication	13 p13-q12.12	1,28	nssv584543	ToF	unknown	duplication	1,05
42	ToF	de novo	duplication	20 p12.2-p11.1	14,59	1692	PDA	unknown	deletion	3,45
						nssv582039	CoAo	unknown	duplication	0,13
188a	ToF	de novo	duplication	17 q12	0,22	nssv578742	TAPVR	unknown	duplication	6,4
188b	ToF	mother	deletion	8 q24.21	0,003	nssv706345	ToF	unknown	duplication	50,7
167	ToF	mother	duplication	8 q24.12	0,55	nssv706345	ToF	unknown	duplication	50,7
184	ToF	mother	duplication	X q25	0,36	nssv578385	Hypoplastic left heart	unknown	deletion	154,9
146a	ToF	mother	duplication	2 p21-p16.3	0,4	251922	PDA	de novo	deletion	6,78
5	ToF	father	duplication	3 p25.3	0,13	248772	PDA	de novo	deletion	3,16
36	ToF	mother	duplication	21q22.3	0,17	nssv576377	CAVCD	unknown	duplication	32,5
93b	ToF	father	deletion	21 q21.1	1,57	nssv576377	CAVCD	unknown	duplication	32,5
106	ToF	father	duplication	15 q21.3	0,065	965	CoAo	unknown	deletion	10,56
						250024	Cardiac Situs Inversus	de novo	deletion	5,99

Supplementary table 5. DAPPLE and GRAIL tool combined *p-value* results for selection of candidate genes ($p < 0.05$) from our CNVs. In this table there are represented the candidate genes obtained for each group of CHD studied. No significant results were obtained for CoA group of patients.

CHD	Gene	Trio ID	GRAIL			DAPPLE	TOTAL
			PubMed 2006	Novartis gene expression	Gene Ontology		
TGA	WDR33	196a	0.015787802	0.15227096	0.010582127	0.323835334105	0.00287207559880234
	CPSF3	265	0.022559575	0.62023193	0.024720905	0.600618530939	0.0305460538041664
	NUSAP1	165b	0.77651924	0.031565095	0.032658305	0.484201145288	0.0467068926678491
ToF	GJA5	43	0.055199803 0.005771034	0.019855906	0.0082752449	0.57770502244	0.00202622780585411
	TJP1	314	5	0.053296526	0.061067656	0.746306118635	0.0043065431777729 0.0063361849401889
	GJA8	43	0.072884468	0.032352244	0.31742415	0.031296957876	9
	MSH2	146	0.12448238	0.027721422	0.073169111	0.357979480341	0.017012384039681
	DSG1	158a	0.015122614	0.088776053	0.11185395	0.802157450069	0.0208663302542409
	MCM3A	36	0.22475795 0.0930385	0.015551175 0.0303605	0.054988285	0.667952600672 0.37305575734	0.0218262604726446 0.02318674708390
	DSG3	158a	47	88	0.13275343	7	32
	CTSS	53	0.0910465 67	0.8869483 4	0.40883486	0.01048760450 29	0.04329491116607 63

Supplementary table 6. List of 124 candidate genes for CHD.

CANDIDATE GENE LIST			
BTRC	SRGAP3	LRCH3	KCNK12
PTPRJ	RIC8A	LINC00343	RAD18
ZNF133	ASIC2/ACCN1	DAOA	YBEY
POLR3F	KLF13	CEP85L	PCNT
RBBP9	BCL9	ST18	GPR89A
DZANK1	LMO7	RB1CC1	GPR89C
CSRP2BP	RASA2	NDUFAF1	NBPF10
CHRNA3	CHRNA7	SORBS2	NBPF20
TCTE3	EIF3B	ACSL6	NBPF25P
ERMARD	MAD1L1	SEC23B	LOC728989
GLIS3	FTSJ2	CDR2	FMO5
RLN3	PPP1CB	ERAP1	CHD1L
PLN	FAM179A	SACS	ACP6
ERBB2	CNNM2	TNFRSF19	GPR89C
TCAP	CDK13	MIPEP	GPR89B
RB1CC	NUP160	C1QTNF9B-AS1	BIRC6
FBN1	NFATC3	C1QTNF9B	TTC27
GJA5	ESRP2	PTPRN2	NCAM2
NFATC3	MGAM	NCAPG2	CGNL1
OVOL2	PUM2	ESYT2	NUP54
WDR33	MAFF	MAPKAPK5	APBA2
CPSF3	TMEM184B	GRB7	FAM189A1
NUSAP1	PDE8A	IKZF3	GOLPH3L
TJP1	KIAA1432	PELI1	HORMAD1
GJA8	FAN1	RAB11FIP5	CTSS
MSH2	MTMR10	MORC1	CTSK
DSG1	LOC283710	ENPP2	LOC101927718
MCM3AP	PRDM5	TAF2	DSG4
DSG3	VPS13B	DEPTOR	PRPF4B
CTSS	ASAP	ODZ1	C6orf201
FOXC1	ADAM17	HIPK3	CDYL

PROJET 2

Congenital heart defects in patients with deletions upstream of SOX9

Marta Sanchez-Castro,^{1,2} Christopher T. Gordon,^{3,4} Florence Petit,⁵ Alex S. Nord,^{6,7} Patrick Callier,⁸ Joris Andrieux,⁵ Patrice Guérin,^{1,2} Olivier Pichon,⁹ Albert David,⁹ Véronique Abadie,¹⁰ Damien Bonnet,¹¹ Axel Visel,^{6,7} Len A. Pennacchio,^{6,7} Jeanne Amiel,^{3,4,12} Stanislas Lyonnet,^{3,4,12} Cédric Le Caignec^{1,2,9*}

Article accepté et publié dans le journal : Human mutation

¹INSERM, UMR1087, l'institut du thorax, Nantes, France

²Université de Nantes, Nantes, France

³INSERM U781, Hôpital Necker-Enfants Malades, Paris, France.

⁴Université Paris Descartes–Sorbonne Paris Cité, Institut Imagine, Paris, France.

⁵Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, France

⁶Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California, USA.

⁷U.S. Department of Energy, Joint Genome Institute, Walnut Creek, California, USA

⁸Département de Génétique, Hôpital Le Bocage, Dijon, France

⁹CHU Nantes, Service de Génétique Médicale, Nantes, France

¹⁰Service de Pédiatrie Générale, Université Paris Descartes, Hôpital Necker-Enfants Malades, Paris, France

¹¹Centre de Référence Malformations Cardiaques Congénitales Complexes – M3C, Hôpital Necker-Enfants Malades, AP-HP, Paris, France

¹²AP-HP, Département de Génétique, Hôpital Necker-Enfants Malades, Paris, France

*Correspondence to: Cédric Le Caignec, MD, PhD, Service de Génétique Médicale, CHU, 9, quai Moncousu 44093 Nantes, France, Tél.: +33 2 40 08 42 84, Fax: +33 2 40 08 39 43, E-mail: cedric.lecaignec@chu-nantes.fr

Congenital Heart Defects in Patients with Deletions Upstream of *SOX9*

Marta Sanchez-Castro,^{1,2} Christopher T. Gordon,^{3,4} Florence Petit,⁵ Alex S. Nord,^{6,7} Patrick Callier,⁸ Joris Andrieux,⁹ Patrice Guérin,^{1,2} Olivier Pichon,¹⁰ Albert David,¹⁰ Véronique Abadie,^{11,12} Damien Bonnet,¹³ Axel Visel,^{6,7} Len A. Pennacchio,^{6,7} Jeanne Amiel,^{3,4,12,14} Stanislas Lyonnet,^{3,4,12,14} and Cédric Le Caignec^{1,2,10*}

¹INSERM, UMR1087, l'institut du thorax, Nantes, France; ²Université de Nantes, Nantes, France; ³INSERM U781, Hôpital Necker-Enfants Malades, Paris, France; ⁴Université Paris Descartes–Sorbonne Paris Cité, Institut Imagine, Paris, France; ⁵Service de Génétique Clinique, Hôpital Jeanne de Flandre, CHRU Lille, France; ⁶Genomics Division, Lawrence Berkeley National Laboratory, Berkeley, California; ⁷U.S. Department of Energy, Joint Genome Institute, Walnut Creek, California; ⁸Département de Génétique, Hôpital Le Bocage, Dijon, France; ⁹Institut de Génétique Médicale, Hôpital Jeanne de Flandre, CHRU de Lille, France; ¹⁰CHU Nantes, Service de Génétique Médicale, Nantes, France; ¹¹Service de Pédiatrie Générale, Université Paris Descartes, Hôpital Necker-Enfants Malades, Paris, France; ¹²CRMR National Reference Centre for Pierre Robin Syndromes and Sucking and Swallowing Congenital Disorders, Paris, France; ¹³Centre de Référence Malformations Cardiaques Congénitales Complexes – M3C, Hôpital Necker-Enfants Malades, AP-HP, Paris, France; ¹⁴AP-HP, Département de Génétique, Hôpital Necker-Enfants Malades, Paris, France

Communicated by Nancy B. Spinner

Received 2 June 2013; accepted revised manuscript 12 September 2013.

Published online 30 September 2013 in Wiley Online Library (www.wiley.com/humanmutation). DOI: 10.1002/humu.22449

ABSTRACT: Heterozygous loss-of-function coding-sequence mutations of the transcription factor *SOX9* cause campomelic dysplasia, a rare skeletal dysplasia with congenital bowing of long bones (campomelia), hypoplastic scapulae, a missing pair of ribs, pelvic, and vertebral malformations, clubbed feet, Pierre Robin sequence (PRS), facial dysmorphism, and disorders of sex development. We report here two unrelated families that include patients with isolated PRS, isolated congenital heart defect (CHD), or both anomalies. Patients from both families carried a very similar ~1 Mb deletion upstream of *SOX9*. Analysis of ChIP-Seq from mouse cardiac tissue for H3K27ac, a marker of active regulatory elements, led us to identify several putative cardiac enhancers within the deleted region. One of these elements is known to interact with Nkx2.5 and Gata4, two transcription factors responsible for CHDs. Altogether, these data suggest that disruption of cardiac enhancers located upstream of *SOX9* may be responsible for CHDs in humans.

Hum Mutat 00:1–4, 2013. © 2013 Wiley Periodicals, Inc.

KEY WORDS: congenital heart defect; pierre robin sequence; *SOX9*; copy number variant; CNV

SOX9 (MIM #608160) is located on chromosome 17q24 and encodes a transcription factor that is expressed in a variety of tissues such as cartilage, testis, heart, glial cells, inner ear, and neural crest cells. Analysis of knockout mice has shown that *Sox9* is involved in heart valve and endocardial cushion (EC) development, male sexual development, and chondrocyte differentiation. Heterozygous loss-of-function coding-sequence mutations of the *SOX9* gene cause campomelic dysplasia (CD; MIM #114290). CD is a severe skeletal dysplasia characterized by congenital bowing of long bones (campomelia), hypoplastic scapulae, a missing pair of ribs, pelvic, and vertebral malformations, clubbed feet, Pierre Robin sequence (PRS), facial dysmorphism and disorders of sex development (DSD) in patients with a 46,XY karyotype. In addition to coding-sequence mutations, a range of translocations, inversions, and deletions that are predicted to disrupt regulatory elements located upstream or downstream of *SOX9* have also been described in patients with CD of varying severity (Supp. Reference List S1).

Several lines of evidence suggest that some regulatory elements in the region surrounding *SOX9* drive tissue-specific expression (Supp. Reference List S2). According to this paradigm, *SOX9* expression in a particular cell population is dependent on a series of enhancers dedicated to that tissue, spread throughout the *SOX9* control region. Deletions and duplications upstream of *SOX9* have been identified in patients with DSD without any other abnormalities (Supp. Reference List S3). Similarly, translocations, deletions, and one point-mutation upstream and downstream of *SOX9* have been detected in patients with isolated PRS, a typical feature of CD [Benko et al., 2009].

We report here two unrelated families with patients with PRS and/or congenital heart defects (CHDs) carrying a ~1 Mb deletion upstream of *SOX9*. The proband of family 1 (patient II-2) was referred to the medical genetics department of CHU Nantes at the age of 20 years for isolated PRS (Fig. 1a). Physical examination, ECG, and cardiac echography were all normal. Family history revealed that her 25-year-old sister (patient II-1) presented with a CHD in addition to PRS. ECG showed an incomplete right bundle branch block but no long QT syndrome. Cardiac echography showed an aortic valve insufficiency (grade 2), a mitral valve ballooning without

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Cédric Le Caignec, Service de Génétique Médicale, CHU, 9, quai Moncousu 44093 Nantes, France. E-mail: cedric.lecaignec@chu-nantes.fr

Contract grant sponsors: PHRC Inter regional (2008); Société Française de Cardiologie/Fédération Française de Cardiologie (2009); Translational Research of Région des Pays de la Loire (2009); École nationale supérieure des mines de Nantes; National Human Genome Research Institute (R01HG003988, U54HG006997); NIDCR (U01-DE020060); NIH/NIGMS F32 Fellowship (GM105202); Agence Nationale de la Recherche (EvoDevoMut 2010).

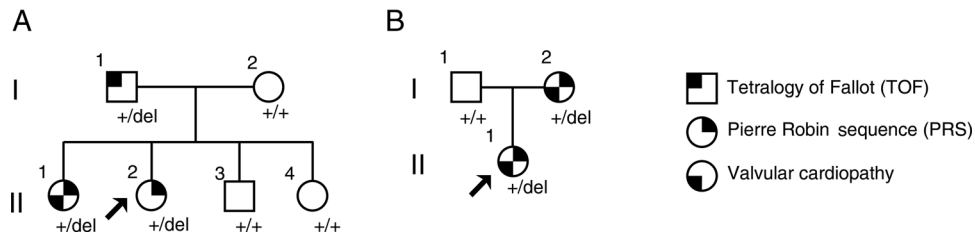


Figure 1. Family 1 (A) and 2 (B) pedigrees showing the phenotype and genotype of each family member. +/del indicates individuals carrying a ~1 Mb heterozygous deletion upstream of *SOX9*; +/+ indicates individuals with a normal 17q24 region.

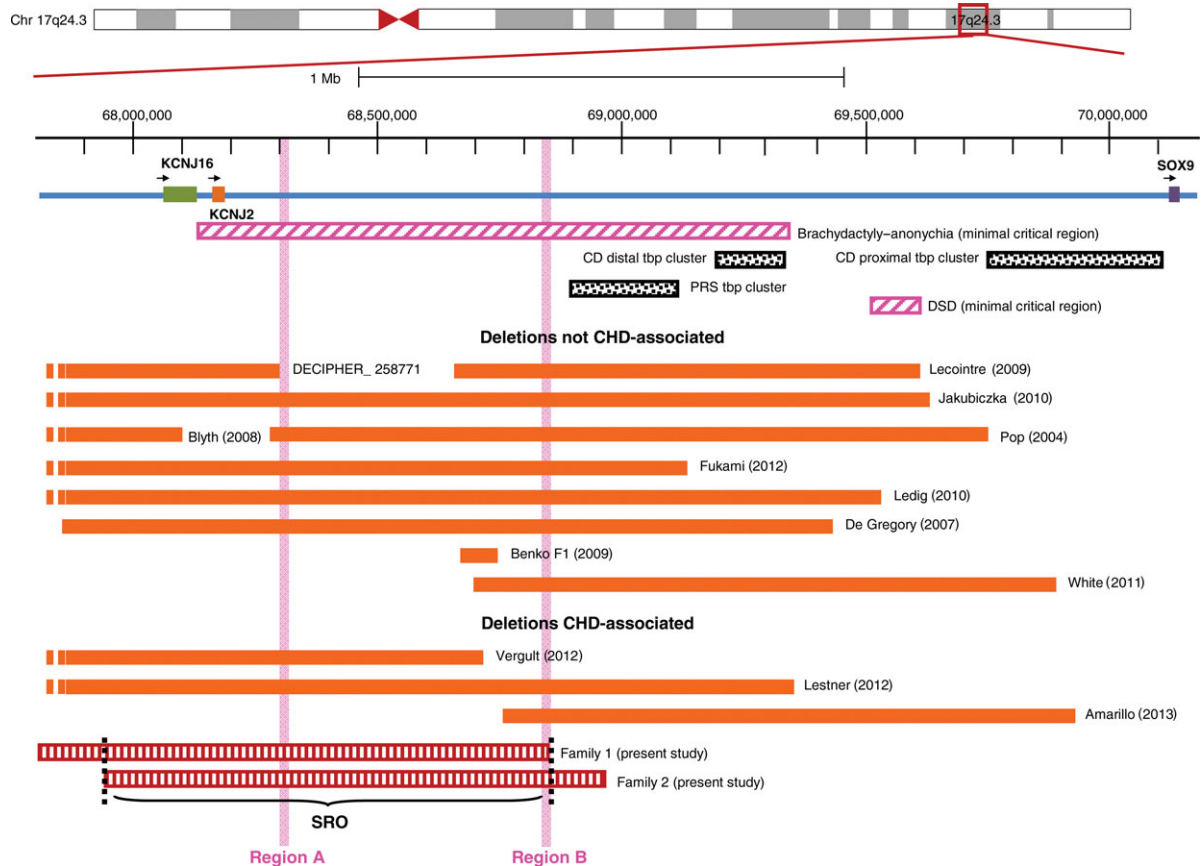


Figure 2. Overview of the *SOX9* upstream region. A 2 Mb region upstream of *SOX9* is shown by a horizontal line, with tick marks at every 100 kb. The black dotted bars indicate the PRS translocation breakpoint (tbp) clusters [Benko et al., 2009; Jakobsen et al., 2007] and the proximal and distal CD tbp clusters summarized in Leipoldt et al. (2007). The diagonal hatched (pink) bars indicate the minimal critical regions for disorders of sex development (DSD) (Supp. Reference List S3) and brachydactyly-anonychia [Kurth et al., 2009]. The vertical hatched (intense red) bars indicate the deletions identified in the two present families showing the shortest region of overlap (SRO); and the dark gray (orange) bars indicate other reported deletions overlapping the SRO that are CHD associated [Amarillo et al., 2013; Lestner et al., 2012; Vergult et al., 2012] or not (Supp. Reference List S6). The light gray (pink) vertical shaded areas indicate the candidate cardiac regulatory regions A and B.

any prolapse, a mild tricuspid valve insufficiency and a pulmonary artery stenosis. Their 52-year-old father (patient I-1) presented an isolated tetralogy of Fallot (ToF). He did not exhibit any feature of PRS. A careful examination of both sisters and their father excluded any skeletal anomalies suggestive of CD. The proband of family 2 (patient II-1), referred to the clinical genetics service of CHRU Lille, presented with PRS and a bicuspid aortic valve (Fig. 2b). Craniofacial features were hypertelorism, thin lips, a posteriorly rotated and low set left ear and a tubular nose; the latter was also present in her father. Radiographs of the spine, thorax, and pelvis were normal.

Her mother (patient I-2) presented with PRS, an atrial septal defect (ASD) and valvular pulmonary stenosis. ECG was normal in the mother with no long QT syndrome.

Informed consent for genetic analyses was obtained from all individuals participating in the study. The protocol was approved by the ethics committee of the University Hospital Center of Nantes (BRD 09/3A). Following normal karyotyping, array CGH analysis performed in the proband of families 1 and 2 using Agilent 180K oligonucleotide arrays showed 17q24 heterozygous deletions located upstream of the *SOX9* gene (Supp. Methods: Genetic analyses). In

patient II-2 of family 1, the 17q24 deleted region was found to be ~1 Mb, extending from -1.25 to -2.3 Mb upstream of *SOX9* [arr 17q24 (67,799,619×2, 67,816,443–68,862,917×1, 68,888,111×2)] (coordinates are based on Human Feb. 2009 (hg19) assembly of International Human Genome Sequencing Consortium) (Supp. Table S1). Array CGH analyses showed identical deletions in her sister (II-1) and her father (I-1) (Supp. Fig. S1a). No other imbalance was identified in the three patients. Subsequently, metaphase fluorescence in situ hybridization analyses with the RP11-693H11 probe located in the deleted region confirmed the deletion in the two affected sisters and their father (data not shown). In patient II-1 of family 2, the 17q24 deleted region was found to be ~1 Mb, extending from -1.15 to -2.15 Mb upstream of *SOX9* [arr 17q24 (67,958,939×2, 67,976,041–68,933,586×1, 68,965,708×2)] (Supp. Table S1 and Supp. Fig. S1b). Subsequently, qPCR analyses using an amplicon located in the deleted region confirmed the deletion in the proband and showed that it was maternally inherited. The shortest region of overlap (SRO) shared by the two families is 887 kb, located between the centromeric boundary 67,976,041 and the telomeric boundary 68,862,917, and includes *KCNJ16*, *KCNJ2*, and 687 kb of the gene desert between *KCNJ2* and *SOX9* (Fig. 2). By Sanger sequencing of patients I-1 of family 1 and I-2 of family 2 (Supp. Table S2), we did not identify any deleterious mutations in genes known to be involved in various CHDs (*TBX5*, *GATA4*, *NKX2.5*, *CFC1*, *FOXH1*, *NODAL*, *TDGF1*, *FOG2/ZFPM2*, *GDF1*, *JAG1*, and *NOTCH1*), which is in favor of a role played by the deletion upstream of *SOX9* in the CHDs observed in these patients.

The deletions that we report here overlap with a region upstream of the PRS translocation breakpoint cluster (Fig. 2) that is presumed to be involved in craniofacial regulation of *SOX9* expression [Benko et al., 2009]. The family 1 and 2 deletions therefore confirm the importance of this region as a PRS locus. XY sex reversal is another feature commonly associated with haploinsufficiency of *SOX9*. Testis specific enhancers have been speculated since some patients with upstream translocations or deletions displayed XY sex reversal. The father from family 1, carrying the deletion, did not present with DSD, which was in agreement with the location of this deletion further upstream than the previously defined minimal critical region for DSDs (Fig. 2).

CHDs have been infrequently reported in association with CD. Houston et al. (1983) reported that CHDs including ToF, ventricular septal defect (VSD) and aortic isthmus stenosis occurred in CD patients. CHDs including VSDs and ASDs were found upon necropsy in 22% of CD patients [Mansour et al., 1995]. A patient was reported with a *SOX9* mutation and CD with a secundum atrioseptal defect [Kim et al., 2011], whereas one patient harboring a 4.7 Mb deletion, including the *SOX9* coding region, presented with CD, coarctation of the aorta, patent ductus arteriosus (PDA), VSD, and persistent foramen ovale [Smyk et al., 2007]. Case 1 in Leipoldt et al. (2007), with a translocation breakpoint 375 kb upstream of *SOX9*, presented with CD including an atrial septal aneurysm and small PDA. Evidence from animal models suggests that *SOX9* is critical for heart development (Supp. Reference List S4). During embryonic development, heart valve leaflets, and supporting structures derive from mesenchymal cells of the ECs in the atrioventricular canal and outflow tract [Lincoln et al., 2004]. ECs elongate as a result of cell proliferation and undergo expansion and remodeling of the extracellular matrix (ECM) to form the mature semilunar valve cusps and AV valve leaflets [Hinton et al., 2006]. *Sox9* is expressed early in valve precursor cells of the EC, during heart septation and remodeling of valve leaflets, and is maintained in the mature valve leaflets

during later stages of development (Supp. Reference List S5). Loss of *Sox9* in mice results in embryonic lethality with hypoplastic ECs [Akiyama et al., 2004]. *Sox9* is required early in valve development for cell proliferation and expansion of the valve progenitor pool and later is required for normal expression and distribution of valve ECM proteins [Lincoln et al., 2007]. The data from mutant mice indicate the potential for dysregulation of *SOX9* in the pathogenesis of valve diseases. Consistent with this, patients from the two families described here with deletions upstream of *SOX9* presented with complex valve and vascular anomalies.

To screen for potential cardiac enhancers located in the SRO, we used ChIP-Seq data generated for H3K27ac, a marker of active regulatory elements, from mouse tissues at across development (Visel, unpublished data) (Supp. Methods: ChIP-Seq and bioinformatic analyses). Via comparison of H3K27ac enrichment, we identified enhancers predicted to be active only in the heart and cardiac enhancers that are also active in other tissues. In the mouse genomic region chr11:110575281–111621449 (mm9), which is orthologous to the region deleted in families 1 and 2, we identified 20 candidate cardiac enhancers, several of which were cardiac specific (Supp. Fig. S2a and b). Given the presence of PRS in the patients, it is likely that these cardiac enhancers are also active in cartilage, despite we do not have data to confirm it. We compared this H3K27ac dataset with two independently generated ChIP-Seq datasets: binding sites for Nkx2.5 and Gata4 (two transcription factors mutated in CHDs) in adult mouse heart [van den Boogaard et al., 2012] and the ENCODE/LICR adult mouse heart p300 ChIP-Seq data (Supp. Fig. S3a). This led to the identification of two major candidate cardiac regulatory regions. Region A is composed of adjacent H3K27ac elements n°10 and n°11. Peaks for Nkx2.5, Gata4, and p300 fell within H3K27ac element n°10, in a region conserved throughout mammals (Supp. Fig. S3b). Region B is composed of a cluster of three H3K27ac elements (n°18–20), with a p300 peak overlapping element n°19, itself conserved to *Xenopus* (Supp. Fig. S3c). Region A is deleted in families 1 and 2 and in two patients with CHDs [Lestner et al., 2012; Vergult et al., 2012], but is not included in the deletion of a patient described in DECIPHER (258771; Fig. 2) who did not present any CHD (personal communication, Michael Speicher). Region B is deleted in families 1 and 2 and in one patient [Lestner et al., 2012] and a family including two individuals presenting with CHDs [Amarillo et al., 2013].

Several hypotheses can be raised to explain the fact that the majority of patients with loss of function coding sequence mutations or translocations located upstream of *SOX9* do not present with CHDs. Variable penetrance of CHDs in these cases may be due to the effect of modifying alleles elsewhere in the genome. It is also possible that the family 1 and 2 deletions lead to a unique reconfiguration of the regulatory domain upstream of *SOX9*, resulting in ectopic upregulation of *SOX9* expression in the heart. A similar outcome could occur if cardiac repressor elements fall within the deletions. It has been suggested that deregulation of *SOX9* can be responsible for phenotypes distinct from those predicted to result from loss of function. For example, some duplications upstream of *SOX9* result in isolated brachydactyly-anonychia, yet deletion within the same interval leads to isolated PRS [Kurth et al., 2009].

It remains to be determined whether *SOX9* is the in vivo target of the putative cardiac regulatory elements described here. Indeed, an alternative target for region A may be *KCNJ2*. *KCNJ2* codes for an inward rectifying potassium channel that has a role in maintaining resting membrane potential in skeletal and cardiac muscle. Coding sequence mutations in *KCNJ2* are responsible for Andersen syndrome (MIM #170390), which involves the core triad of

cardiac arrhythmias, periodic paralysis, and dysmorphic features. Recently, Lestner et al. (2012) reported a patient presenting features of CD and a small PDA with left to right shunting, associated with branch pulmonary artery stenosis, a small inter-atrial connection and mild myxomatous thickening of the mitral valve. The array CGH analysis detected a ~2.3 Mb deletion upstream of *SOX9* extending more proximally and distally compared with our deletions. Also, a 4.16 Mb deletion, covering most of the family 1 and 2 SRO, was reported in a patient with dysmorphic features, pulmonary stenosis with a dysplastic pulmonary valve, PDA and drug-induced long QT syndrome [Vergult et al., 2012]. The deletions described in the present study, by Lestner et al. (2012) and by Vergult et al. (2012) involved the *KCNJ2* and *KCNJ16* genes, which therefore are alternative candidates (Fig. 2). Despite a report of valve defects in several members of a large family harboring a *KCNJ2* mutation [Andelfinger et al., 2002], valvulopathies do not appear to normally be part of the Andersen syndrome phenotype. *KCNJ2* mutations associated with Andersen syndrome typically act in a dominant negative manner [Preisig-Müller et al., 2002]. It is therefore unclear whether genomic lesions influencing *KCNJ2* expression levels could result in CHDs or electrophysiological defects, although this possibility is supported by the phenotype of drug-induced long QT in the patient harboring a deletion of *KCNJ2* in Vergult et al. (2012) (prolonged QT is one of the major features of *KCNJ2* coding mutations). Also, the possibility of long-range transcriptional regulation of *KCNJ2* is supported by SNPs centered on region A associating with susceptibility to thyrotoxic periodic paralysis and SNPs greater than 300 kb downstream of *KCNJ2* (i.e., within the *KCNJ2*-*SOX9* gene desert) that associate with certain ECG profiles and influencing predisposition to arrhythmias in GWASs (Supp. Reference List S7). *KCNJ16* (*Kir5.1*) codes for an inward rectifying potassium channel that coassembles selectively with either *Kir4.1* (*KCNJ10*) or *Kir4.2* (*KCNJ15*) to form heteromeric channels. The functional importance of *Kir4.1/Kir5.1* in renal ion transport has recently been highlighted by mutations in the human *Kir4.1* gene (*KCNJ10*) that result in seizures, sensorineural deafness, ataxia, mental retardation, and electrolyte imbalance/epilepsy, ataxia, sensorineural deafness, and renal tubulopathy syndrome, a complex disorder that includes salt wasting and hypokalemic alkalosis [Paulais et al., 2011]. Therefore, it is unlikely that deregulation of *KCNJ16* would cause CHDs. No CHDs were reported in a patient (258771) described in DECIPHER carrying a de novo microdeletion including *KCNJ2* and *KCNJ16*, whereas several deletions that include these genes (but which leave the majority of the family 1 and 2 SRO intact) have been deposited in the Database of Genomic Variants, suggesting that haploinsufficiency of these genes alone is unlikely to lead to strongly penetrant CHDs. We cannot exclude the possibility that the combination of the loss of *KCNJ2* and regulatory elements required for cardiac expression of *SOX9* contribute to the CHDs observed in families 1 and 2.

In summary, we have identified overlapping deletions of ~1 Mb in two families presenting with PRS and/or CHDs, of which the latter are predominantly valvulopathies. Analysis of ChIP-Seq data allowed us to identify major candidate cardiac regulatory elements in the deleted region. Future work will involve testing the in vivo cardiac activity of these regions in transgenic reporter mice, and investigation of the interactions of these regions with the surrounding candidate target genes.

Acknowledgments

Parts of this research were conducted at the E.O. Lawrence Berkeley National Laboratory and performed under Department of Energy Contract DE-AC02-05CH11231, University of California. We are most grateful to the Genomics platform of Nantes (Biogenouest Genomics) core facility for its technical support.

References

- Akiyama H, Chaboissier MC, Behringer RR, Rowitch DH, Schedl A, Epstein JA, de Crombrughe B. 2004. Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. *Proc Natl Acad Sci USA* 101:6502–6507.
- Amarillo IE, Dipple KM, Quintero-Rivera F. 2013. Familial microdeletion of 17q24.3 upstream of *SOX9* is associated with isolated Pierre Robin sequence due to position effect. *Am J Med Genet* 161A:1167–1172.
- Andelfinger G, Tapper AR, Welch RC, Vanoye CG, George AL, Benson DW. 2002. *KCNJ2* mutation results in Andersen syndrome with sex-specific cardiac and skeletal muscle phenotypes. *Am J Hum Genet* 71:663–668.
- Benko S, Fantes JA, Amiel J, Kleinjan DJ, Thomas S, Ramsay J, Jamshidi N, Essafi A, Heaney S, Gordon CT, McBride D, Golzio C, et al. 2009. Highly conserved non-coding elements on either side of *SOX9* associated with Pierre Robin sequence. *Nat Genet* 41:359–364.
- Van den Boogaard M, Wong LY, Tessadori F, Bakker ML, Dreizehnter LK, Wakker V, Bezzina CR, 't Hoen PA, Bakkers J, Barnett P, Christoffels VM. 2012. Genetic variation in T-box binding element functionally affects *SCN5A/SCN10A* enhancer. *J Clin Invest* 122:2519–2530.
- Hinton RB, Lincoln J, Deutsch GH, Osinska H, Manning PB, Benson DW, Yutzey KE. 2006. Extracellular matrix remodeling and organization in developing and diseased aortic valves. *Circ Res* 98:1431–1438.
- Houston CS, Opitz JM, Spranger JW, Macpherson RI, Reed MH, Gilbert EF, Herrmann J, Schinzel A. 1983. The campomelic syndrome: review, report of 17 cases, and follow-up on the currently 17-year-old boy first reported by Maroteaux et al in 1971. *Am J Med Genet* 15:3–28.
- Jakobsen LP, Ullmann R, Christensen SB, Jensen KE, Mølsted K, Henriksen KF, Hansen C, Knudsen MA, Larsen LA, Tommerup N, Tümer Z. 2007. Pierre Robin sequence may be caused by dysregulation of *SOX9* and *KCNJ2*. *J Med Genet* 44:381–386.
- Kim HY, Yoon CH, Kim GH, Yoo HW, Lee BS, Kim KS, Kim EA. 2011. A case of campomelic dysplasia without sex reversal. *J Korean Med Sci* 26:143–145.
- Kurth I, Klopocki E, Stricker S, van Oosterwijk J, Vanek S, Altmann J, Santos HG, van Harssel JTT, de Ravel T, Wilkie AOM, Gal A, Mundlos S. 2009. Duplications of non-coding elements 5[prime] of *SOX9* are associated with brachydactyly-anonychia. *Nat Genet* 41:862–863.
- Leipoldt M, Erdel M, Bien-Willner GA, Smyk M, Theurl M, Yatsenko SA, Lupski JR, Lane AH, Shanske AL, Stankiewicz P, Scherer G. 2007. Two novel translocation breakpoints upstream of *SOX9* define borders of the proximal and distal breakpoint cluster region in campomelic dysplasia. *Clin Genet* 71:67–75.
- Lestner JM, Ellis R, Canham N. 2012. Delineating the 17q24.2-q24.3 microdeletion syndrome phenotype. *Eur J Med Genet* 55:700–704.
- Lincoln J, Alfieri CM, Yutzey KE. 2004. Development of heart valve leaflets and supporting apparatus in chicken and mouse embryos. *Dev Dyn* 230:239–250.
- Lincoln J, Kist R, Scherer G, Yutzey KE. 2007. Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev Biol* 305:120–132.
- Mansour S, Hall CM, Pembrey ME, Young ID. 1995. A clinical and genetic study of campomelic dysplasia. *J Med Genet* 32:415–420.
- Paulais M, Bloch-Faure M, Picard N, Jacques T, Ramakrishnan SK, Keck M, Sohet F, Eladari D, Houillier P, Lourd S, Teulon J, Tucker SJ. 2011. Renal phenotype in mice lacking the *Kir5.1* (*Kcnj16*) K⁺ channel subunit contrasts with that observed in SeSAME/EAST syndrome. *Proc Natl Acad Sci USA* 108:10361–10366.
- Preisig-Müller R, Schlichthörl G, Goerge T, Heinen S, Brüggemann A, Rajan S, Derst C, Veh RW, Daut J. 2002. Heteromerization of *Kir2.x* potassium channels contributes to the phenotype of Andersen's syndrome. *Proc Natl Acad Sci USA* 99:7774–7779.
- Smyk M, Obersztyn E, Nowakowska B, Bocian E, Cheung SW, Mazurczak T, Stankiewicz P. 2007. Recurrent *SOX9* deletion campomelic dysplasia due to somatic mosaicism in the father. *Am J Med Genet* 143A:866–870.
- Vergult S, Dauber A, Delle Chiaie B, Van Oudenhove E, Simon M, Rihani A, Loeys B, Hirschhorn J, Pfotenhauer J, Phillips JA, Mohammed S, Ogilvie C, et al. 2012. 17q24.2 microdeletions: a new syndromal entity with intellectual disability, truncal obesity, mood swings and hallucinations. *Eur J Hum Genet* 20:534–539.

Supplementary data

Supplementary Methods

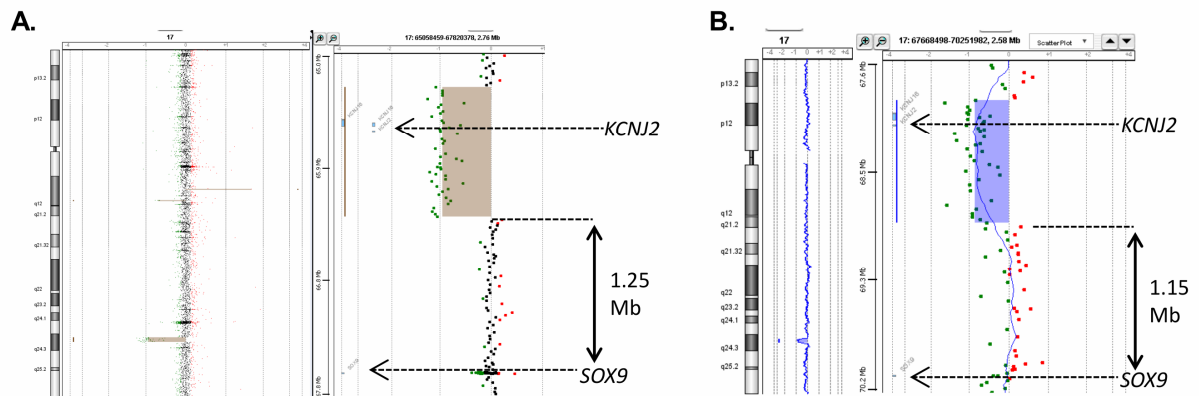
Genetic analyses

Array CGH experiments were performed using Agilent Human Genome CGH 180K oligonucleotide arrays (Agilent, Santa Clara, CA; www.agilent.com) with the ISCA design (www.iscaconsortium.org) in family 1 and 180K oligonucleotide arrays with the catalog design in family 2 following the protocols provided by Agilent (Supplementary Figure 1). The arrays were analyzed with the Agilent scanner and the Feature Extraction software (v. 9.1.3). Graphical overview was obtained using the CGH analytics software (v.3.5.14).

ChIP-Seq and bioinformatic analyses

Generation of H3K27ac ChIP-Seq data at different stages of mouse heart development was performed using methods described elsewhere (Visel et al., 2013) and reanalysed for this study (Nord, Pennacchio, Visel, unpublished work). Candidate cardiac enhancers from the H3K27ac datasets were cross-referenced with the Nkx2.5 and Gata4 peaks obtained from adult mouse heart by ChIP-Seq (van den Boogaard et al., 2012) and the ENCODE/LICR adult mouse heart p300 ChIP-Seq data (<http://genome.ucsc.edu/>), and converted to human genome co-ordinates (Hg19). For visualisation of the mouse ENCODE/LICR data in the UCSC browser, the “Heart Adult 8 weeks p300 TFBS ChIP-seq Peaks from ENCODE/LICR” track was selected, with Minimum Signal Value set to 5.

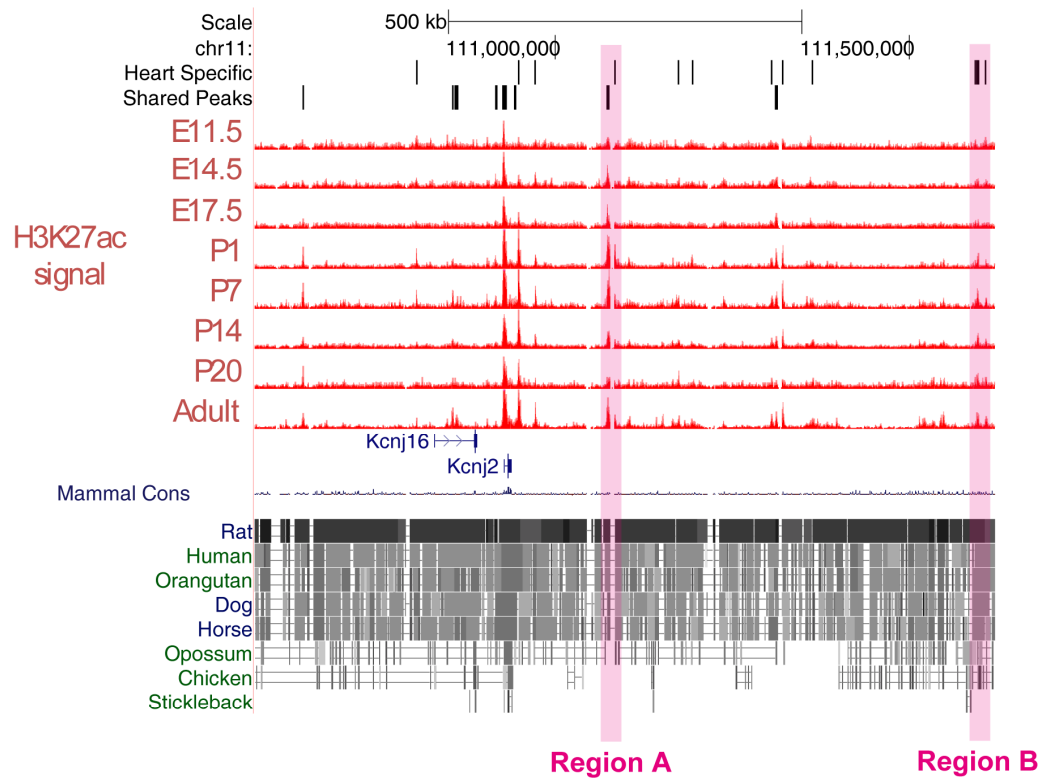
Supplementary Figure 1. The Agilent 180K array-derived profile of chromosome 17 of the affected members from family 1 (A.) and family 2 (B.). For each panel, the ideogram and the oligomers of the corresponding chromosome and a detailed view of the region of interest are shown. The oligomers are arranged along the y-axis according to their position. The x-axis represents the log 2 transformed intensity ratios. Intensity ratios corresponding to deleted probes are colored in green. The abnormal genomic region detected by the CGH Analytics algorithm for 180K arrays is colored in brown in family 1 (A), and in blue in family 2 (B). Location of *SOX9* and *KCNJ2* genes on 17q24.3 is indicated.



Supplementary Figure 2. a) Developmental time course of H3K27ac signal and peaks within a ~1 Mb mouse genomic region orthologous to the deleted region upstream of *SOX9* identified in families 1 and 2. The boundaries of the region investigated in mouse are chr11:110575281-111621449 (mm9). Three different mouse tissues were analysed: heart, forebrain and liver. Heart Specific Peaks refers to H3K27ac peaks found only in heart tissue, while Shared Peaks refers to H3K27ac peaks found in heart as well as in another of the studied tissues. b) Co-ordinates of individual H3K27ac peaks. “Other tissue” refers to comparison of heart H3K27ac signal with that in forebrain and liver.

The shaded pink areas indicate the candidate cardiac regulatory regions A and B.

a. Mm9 Microdeletion region (chr11:110575281-111621449)



b. Putative distal enhancers (H3K27ac-enriched regions)

n°	chr	start	end	Predicted Heart Activity in mouse	Genomic Feature	Other tissue	
1	chr11	110642295	110644942	E14.5 to Adult		forebrain	
2	chr11	110803045	110805187	E11.5 to Adult		heart specific	
3	chr11	110853924	110856494	E11.5 to Adult	Intronic (Kcnj16)	forebrain and liver	
4	chr11	110857456	110862695	E11.5 to Adult	Intronic (Kcnj16)	forebrain	
5	chr11	110914819	110918433	E11.5 to P7		forebrain	
6	chr11	110929645	110931763	E14.5 to Adult	Intronic (Kcnj2)	forebrain	
7	chr11	110942227	110945239	P7 only		forebrain	
8	chr11	110947288	110949697	E14.5 to Adult		heart specific	
9	chr11	110970840	110973133	E11.5 to Adult		heart specific	
10	chr11	111072376	111077006	E11.5 to adult		forebrain	Region A
11	chr11	111083486	111085487	P1 to P7		heart specific	
12	chr11	111172799	111175821	P7 to P20		heart specific	
13	chr11	111192770	111194771	E11.5 only		heart specific	
14	chr11	111304844	111307041	E14.5 to Adult		heart specific	
15	chr11	111310726	111314334	E14.5 to Adult		forebrain	
16	chr11	111320497	111322879	P1 to Adult		heart specific	
17	chr11	111362528	111364529	P20 only		heart specific	
18	chr11	111592616	111594617	E14.5 to P20		heart specific	Region B
19	chr11	111595385	111598970	E14.5 to Adult		heart specific	
20	chr11	111607574	111609808	E14.5 to P20		heart specific	

Supplementary Figure 3. a) Mouse H3K27ac peaks converted into hg19 and placed under the deleted regions present in families 1 and 2. In the red bars representing deletions, the thin ends indicate regions of uncertainty of the precise deletion boundaries. b) zoomed view of region A, depicting Nkx2.5 and Gata4 peaks found in adult mouse heart by Chip-Seq (van den Boogaard et al., 2012), the ENCODE/LICR adult mouse heart p300 peaks and a SNP associating with susceptibility to thyrotoxic periodic paralysis in a genome-wide association study (Cheung et al., 2012) c) zoomed view of region B.



Supplementary Table 1. 180K array CGH results for the deletions upstream of SOX9.

		Family 1 Patient II-2	Family 2 Patient II-1
Last normal copy-number probe (centromeric side)	Agilent probe reference	A_16_P0328244 2	A_16_P0328268 1
	Physical position of the probe (start-end)(hg19)	67,799,560-67,799,619	67958880-67958939
First deleted probe (centromeric side)	Agilent probe reference	A_16_P4088924 1	A_16_P2071569 5
	Physical position of the probe (start-end)(hg19)	67,816,443-67,816,502	67976041-67976100
Last deleted probe (telomeric side)	Agilent probe reference	A_16_P0328422 6	A_16_P2071832 9
	Physical position of the probe (start-end)(hg19)	68,862,917-68,862,976	68933527-68933586
First normal copy-number probe (telomeric side)	Agilent probe reference	A_16_P2071821 3	A_16_P2071838 2
	Physical position of the probe (start-end)(hg19)	68,888,111-68,888,170	68965708-68965767
Minimum size of the deletion (in Kb)		1,046 Kb	958 Kb
Maximum size of the deletion (in Kb)		1,088 Kb	1,007 Kb

Supplementary Table 2. Primers for *NKX2.5*, *CFC1*, *FOXH1*, *NODAL*, *TDGF1*, *FOG2/ZFPM2*, *GDF1*, *JAG1* and *NOTCH1* were designed with Primer3-PCR primer design tool v.0.4.0. Primers sequences for *TBX5* and *GATA4* were taken from literature (Basson et al., 1999; Tomita-Mitchell et al., 2007)

GENE	EXON	FORWARD (5' to 3')	REVERSE (3' to 5')
<i>NKX2.5</i>	Exon 1	AGTTTCTTGGGGACGAAAGC	CTTGTGCTCAGCGCTACCT
	Exon 2.1	GCGCTCCGTAGGTCAAGC	TAGGGATTGAGGCCACG
	Exon 2.2	CAGACTCTGGAGCTGGTGG	CCGAGAGTCAGGGAGCTGT
<i>CFC1</i>	Exon 1	ACAGTGGCTCATGCACTACG	GGGCCTGGAGTAAAGACACC
	Exon 3	TGAGTCCGCAGACTGAGATG	TGCATGAGCCACTGTTCATT
	Exon 4-5	ATGCACTGTGGATCGGTATG	TGGCTGGAGGCTGATTTTAC
	Exon 6	GTCCCTCTCCCAAGGATCTG	ACCGAGCCGGAACACTTTAT
<i>FOXH1</i>	Exon 1	GGCCCTGCCTTCTACACTG	CGAGTCTGGGGAAAGAGAGA
	Exon 2	AGTGGGACCAGTCCTGAGAG	TAGCAGGGACTTGATGCTGA
	Exon 3	GGAAGTCCTGAGCATCCAAG	CTTGGCTCCCTGTCAACAA
<i>NODAL</i>	Exon 1	TGCTGAGGCCCAGGATATAA	CACCCTGAACGCTAGAATCC
	Exon 2	CCCTCCTTCCTTCTTGTTT	AGCAAAGCTAGAGCCCTGTC
	Exon 3	CTATTCTGACCTGCCCATCA	CAGAGTGCAGGCAAATCCA
<i>TDGF1</i>	Exon 1	CGATGCTAACGCCTCTTTTC	GGTCTTTCAGGAAAAACAGCA
	Exon 2-3	TGTGGTCTTGTCTTGTGATG	CAGGTGCCCTTTTAGGTGAA
	Exon 4-5	TTCCATCCCTACACCCTCAG	CCAGGCTCTAGCAAGACAAGA
	Exon 6	AGCAGGATGAACTGCCAGAG	AGCAGCAGCCTTTACTGGTC

<i>FOG2/ZFPM2</i>	Exon 1	GCTCATCTCCGAACGTGAAT	GCAATAATCCCACCAACTCC
	Exon 2	TCTTTTGTGAGTCCCCTTGG	TCAGACGGCTGAATTATGTGT
	Exon 3	CTTTCTCTCCTTGTGTTGGTG	CAAGCAGAGGTAGCACTTTGG
	Exon 4	GCATGTGAAAAGGTGAAAGGA	TTCATTGGAAAGCAAGCAGA
	Exon 5	TTAATCAACGGAAGCAAATGG	TTCCATTCCAGCTCTTCTGG
	Exon 6	TTGATCAAACACTCCACTCCA	TCCCATAAATTCCCATCACC
	Exon 7	AACCTGAGAGCGGAGTCTGA	TCACACTGGTTGACCCAGAA
	Exon 8.1	TCCCTGTCATTCAAAAACCA	TCTGGGGACTTAGCCATCTG
	Exon 8.2	CCGTCATTTACAGCCCTTTG	CTGGAACACACTTGTTGATCG
	Exon 8.3	GCCTACCTGAGCAGGAACAA	AGCACACCCATTTGAGGAAG
	Exon 8.4	TGAAGCTGCTCAGCTCATTG	TTTTTGAAAATGGCACCAATA

<i>GDF1</i>	Exon 7	CTGCGCACTCTCTGGTCAT	GGTGGCCGAAGTTGCTAGT
	Exon 8	CTTCCTCCCTGTCCCTAACC	CACGCGGCATTTATTGTTG

<i>JAG1</i>	Exon 1-2a	GAAAGGCAGCACGATGCG	GTGCATGCTCCAATCGGC
	Exon 2b	TGTCCATGCAGAACGTGAAC	CGCCTAGTTTCAAGCCAAAG
	Exon 3	GGAGGCCAAGAATGTACGAG	CAAGAGGTGGCAGAAATGCT
	Exon 4	CAGGTTGTCTCTGGGTAGGG	CTAAAAGCAACAGGCACACG
	Exon 5	CCTCCCATCTTCCTCTCCTT	GTGTGCATTCCACCAGACAG
	Exon 6-7	AAAAAGCCTGGACTTTGCAG	TGGGATAAGGCCTATCTTTGG
	Exon 8	TATCACACGCTTGCTTTTGG	CACCGAGACATTCACTGG
	Exon 9	CAATGACAGCCCCTCATGTA	CCCAGGCCTGCTAGACTCTT
	Exon 10-11	AGGGAGAGGGATCAGATGCT	CACAGGGACAGAGCTCTCCTA
	Exon 12	AAGCCCTGTGTTTGTGGAAT	CAGCGGCTCTGCTCTAAAAA
	Exon 13	CTCTGATCCCTCCCCCTTT	AGTCTGGAGACAGCCAGGTC

Exon 14	AGGCCCTTGGCTTAGGAAT	CAGTCTCCCCAGCCTTACTG
Exon 15	GCCATGAAAAGCTGCTGAACA	CCCTCCAACATGACCCATAC
Exon 16-17-18	TTCCTGGAAAAGCCTGATTG	CAAGGGTGTGAGGATCTGCT
Exon 19	GCCTGGTTCCAATTTAGCAA	ATGGAATGAAGCGGTAAAGC
Exon 20	ATTGGCAGCCAGAGTAATGG	TAAAGGGAATGGTGGCTTTG
Exon 21	TCAAAACATTGCCACACACC	GACTGCAAAGCTCGCTCAC
Exon 22-23	AAAGAATTGAACCCCGATCC	TAGCTCATGGCATTCTCCT
Exon 24-25	AGCCAGCCTCAAAGAGAACA	GGTAGGGCACTGCCAGATAA
Exon 26	GTTCTGCGATGATGCTTTT	GCCAGTGTAAGCCAGCTTGT

NOTCH1	Exon 1	AGCCAGCCGGGGAAGAGA	GAAGGTCGGTCCTCCCTGAT
	Exon 2	CCCCAGCAACCCATGATACT	GGGCCATCTCCAGAAGACAA
	Exon 3	ACAGGTTCCCGGTGCTGAG	GATCCCGCCAAGTACCTCAA
	Exon 4	TGGGAGCTGTGGTCCTTCAT	CATCCCGCCTTCCCAACT
	Exon 5	GTGGAGGCCAGGCTCTTGT	CCTTCTCGGCCAACCCCTAGT
	Exon 6	GCCTGCAGAGGAAGGAGTGA	GGGTGAGGTCACACAGCTCAG
	Exon 7-8	GCACCCAGGATAACCTGAAA	AAGCAACCCACAGATGTTCC
	Exon 9	GGGTTCGTTTCTGTCCCAAGT	ACAACCCACGGCCCTCAC
	Exon 10	GGTCCCTGAAGGGCCATAG	CCAGACCAAGGTGTCCATGA
	Exon 11-12	TCCTGAGTAGGTGGGAATGC	GAGTGGCTGACCTTGACCTC
	Exon 13	TGGAAGGATGTGGCCAGAAG	TGGCCCATCTCAAGCTCTGT
	Exon 14	GCCCATAGGGCATTGCAGAC	TGTCTCCAGCTCCCCAGACT
	Exon 15	GGCCTCTGGCTCTTGGAGA	CCCCAAGAAAGCCACCACTT
	Exon 16	GGGAGATCCCTCTGGGTGAG	AACGGTGCTCCAGGTCAAT
	Exon 17	TTCCCAACTCCAGTTCCTGT	GAGGCCTGAGAGCTTCCTG
	Exon 18	CTTTGTCACAGGGCTGCTT	CAGGGACGAGACTGACGTTT
	Exon 19	GTGCTGGCGAGGCAGAATC	GCTGCTCAGATCCCCAGAAA

Exon 20-21	TGCAGCCTGGGTTGGAGTAG	TTTCCACAGGCCACCTCTTC
Exon 22	GGGTCTCTCTCCAGGTCTGA	GAGTTTCTGGCTGGTTCCTG
Exon 23-24	ACCCTTGACACCTGCCTCTC	CGTGGACCTCTCCAGGTGTCT
Exon 25	GCCTCCCAGGTTAGAGGAGA G	AGCCAGGCCACATCCAAGT
Exon 26	GAGGCCAGCATGCAGTTCTAA	CCTCTCGGAACCTCCGTCTC
Exon 27	TGTCCTGACTGTGGCGTCAT	CACACCCGTGGGTAGCAACT
Exon 28	ATGGAGGAGAGTGGGTGAG	GGTGAGGGGTCGAGAAAGTGA
Exon 29-30	ATCCTCACCGGGAGTCTG	GGAAACACCCCAAGGATGAA
Exon 31	AAGCTGGATTGGGCTCTGT	CCTGACTTCCCTGGAGAACA
Exon 32-33	GCCGATCTCAGGAGGGTCTC	GGGTTTGGCCCTCACTTCTC
Exon 34.1	TTCCTCTGGTGATGGAACCT	CGCATTGACCATTCAAAGT
Exon 34.2	CTTTGAGACTGGCCACCT	AAGGCTTGGGAAAGGAAGC

TBX5	Exon 2	GCTTCTTGTCTCAGAGCAGA ACCT	CAAGAGAAGCCGAGCAGGAAA GCCA
	Exon 3	TTCTCCTCGTCCCTCTCTCTAC ACA	AGTTTGGGGAAGGAATGCCAC TAC
	Exon 4	AACGGGGCTAGTTTCCGCTTC CACG	CTTTTTGGGAGAAGGTTCCACTT TT
	Exon 5	CTGGTGCGTGAAGTGAAGCAC GC	GAGGGAGACAAGGCGGGGAAT CCAG
	Exon 6	ACGGCCCCAGGCACTGGTTC CTGGG	CAGGGTTTTATCTGGAGACAAA GGG
	Exon 7	ATTAGCTCATGTCCTGAGGTG GTCT	GTGGGGAGGAGAAAGTTGAGG AATC
	Exon 8	CTTTTCTGGTGGATTCTCTCAC ACC	GGGTAGGAACATGTCAAGGGAA CT
	Exon 9	TACTTTGGCCAAATAACTGTCT CC	CGACCTTGAGTGCAGATGTGAA C

GATA4	Exon 1.2	TGTTGCCGTCGTTTTCTCTC	GTCCCCGGAAGGAGAAG
	Exon 1.2	CGACGGAGCCGCTTACAC	GCGTTGGTGGAAAAACAAGA
	Exon 2	TGAGAGCTGGGCATAAACAA	AGAGGATGTCCCACCAAGG
	Exon 3	GGGCAGTGCACACCTTTTAC	GAGAGATGGGCATCAGAAGG
	Exon 4	TGTCTTTCAATGCTGTAGCAG A	TGCCTAACCCGGAAGATATG
	Exon 5	CATTAGCTTGACCCCATCC	AGTACTAGGCTGGCCTCTGG
	Exon 6	CCTAGACCTCCCAAGC	GGCCTCCTTCTTTGCTATCC

Supplementary Reference List 1

Case reports of patients presenting with CD of varying severity found to carry translocations, inversions and deletions that are predicted to disrupt regulatory elements located upstream or downstream of SOX9:

(Foster et al., 1994; Hill-Harfe et al., 2005; Lecointre et al., 2009; Leipoldt et al., 2007; Ninomiya et al., 1996; Pfeifer et al., 1999; Pop et al., 2004; Tommerup et al., 1993; Velagaleti et al., 2005; Wagner et al., 1994; Wirth et al., 1996; Wunderle et al., 1998; Young et al., 1992)

Supplementary Reference List 2

Reported evidence suggesting that some regulatory elements in the region surrounding SOX9 drive tissue specific expression:

(Bagheri-Fam et al., 2006; Bishop et al., 2000; Mead et al., 2013; Qin et al., 2004; Sekido and Lovell-Badge, 2008; Velagaleti et al., 2005; Wunderle et al., 1998)

Supplementary Reference List 3

Reported patients carrying deletions and duplications upstream of SOX9 presenting with DSD without any other abnormalities:

(Benko et al., 2011; Cox et al., 2011; Vetro et al., 2011)

Supplementary Reference List 4

Animal model evidence suggesting that SOX9 is critical for heart development:

(Akiyama et al., 2004; Lincoln et al., 2007; Montero et al., 2002; Rahkonen et al., 2003)

Supplementary Reference List 5

Studies showing that Sox9 is expressed early in valve precursor cells of the EC, during heart septation and remodeling of valve leaflets, and is maintained in the mature valve leaflets during later stages of development:

(Akiyama et al., 2004; Hinton et al., 2006; Lee and Saint-Jeannet, 2009; Lincoln et al., 2006, 2007; Montero et al., 2002; Rahkonen et al., 2003)

Supplementary Reference List 6

Reported patients carrying deletions overlapping the SRO region that are not CHD-associated:

(Benko et al., 2009; Blyth et al., 2008; Fukami et al., 2012; De Gregori et al., 2007; Jakubiczka et al., 2010; Lecointre et al., 2009; Ledig et al., 2010; Pop et al., 2004; White et al., 2011) and DECIPHER patient 258771.

Supplementary Reference List 7

(Cheung et al., 2012; Jongjaroenprasert et al., 2012; Marjamaa et al., 2012; Pfeufer et al., 2009)

References

- Akiyama, H., Chaboissier, M.C., Behringer, R.R., Rowitch, D.H., Schedl, A., Epstein, J.A., and de Crombrughe, B. (2004). Essential role of Sox9 in the pathway that controls formation of cardiac valves and septa. *Proc Natl Acad Sci U S A* 101, 6502–6507.
- Bagheri-Fam, S., Barrionuevo, F., Dohrmann, U., Günther, T., Schüle, R., Kemler, R., Mallo, M., Kanzler, B., and Scherer, G. (2006). Long-range upstream and downstream enhancers control distinct subsets of the complex spatiotemporal Sox9 expression pattern. *Dev. Biol.* 291, 382–397.
- Basson, C.T., Huang, T., Lin, R.C., Bachinsky, D.R., Weremowicz, S., Vaglio, A., Bruzzone, R., Quadrelli, R., Lerone, M., Romeo, G., et al. (1999). Different TBX5 interactions in heart and limb defined by Holt-Oram syndrome mutations. *Proc Natl Acad Sci U S A* 96, 2919–2924.
- Benko, S., Fantes, J.A., Amiel, J., Kleinjan, D.J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C.T., et al. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. *Nat Genet* 41, 359–364.
- Benko, S., Gordon, C.T., Mallet, D., Sreenivasan, R., Thauvin-Robinet, C., Brendehaug, A., Thomas, S., Bruland, O., David, M., Nicolino, M., et al. (2011). Disruption of a long distance regulatory region upstream of SOX9 in isolated disorders of sex development. *J. Med. Genet.* 48, 825–830.
- Bishop, C.E., Whitworth, D.J., Qin, Y., AgoulNIK, A.I., AgoulNIK, I.U., Harrison, W.R., Behringer, R.R., and Overbeek, P.A. (2000). A transgenic insertion upstream of sox9 is associated with dominant XX sex reversal in the mouse. *Nat Genet* 26, 490–494.

- Blyth, M., Huang, S., Maloney, V., Crolla, J.A., and Karen Temple, I. (2008). A 2.3Mb deletion of 17q24.2-q24.3 associated with “Carney Complex plus.” *Eur J Med Genet* 51, 672–678.
- Van den Boogaard, M., Wong, L.Y., Tessadori, F., Bakker, M.L., Dreizehnter, L.K., Wakker, V., Bezzina, C.R., 't Hoen, P.A., Bakkers, J., Barnett, P., et al. (2012). Genetic variation in T-box binding element functionally affects SCN5A/SCN10A enhancer. *J Clin Invest* 122, 2519–2530.
- Cheung, C.L., Lau, K.S., Ho, A.Y., Lee, K.K., Tiu, S.C., Lau, E.Y., Leung, J., Tsang, M.W., Chan, K.W., Yeung, C.Y., et al. (2012). Genome-wide association study identifies a susceptibility locus for thyrotoxic periodic paralysis at 17q24.3. *Nat Genet* 44, 1026–1029.
- Cox, J.J., Willatt, L., Homfray, T., and Woods, C.G. (2011). A SOX9 Duplication and Familial 46,XX Developmental Testicular Disorder. *N. Engl. J. Med.* 364, 91–93.
- Foster, J.W., Dominguez-Steglich, M.A., Guioli, S., Kwok, C., Weller, P.A., Stevanović, M., Weissenbach, J., Mansour, S., Young, I.D., and Goodfellow, P.N. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. *Nature* 372, 525–530.
- Fukami, M., Tsuchiya, T., Takada, S., Kanbara, A., Asahara, H., Igarashi, A., Kamiyama, Y., Nishimura, G., and Ogata, T. (2012). Complex genomic rearrangement in the SOX9 5' region in a patient with Pierre Robin sequence and hypoplastic left scapula. *Am J Med Genet* 158A, 1529–1534.
- De Gregori, M., Ciccone, R., Magini, P., Pramparo, T., Gimelli, S., Messa, J., Novara, F., Vetro, A., Rossi, E., Maraschio, P., et al. (2007). Cryptic deletions are a common finding in “balanced” reciprocal and complex chromosome rearrangements: a study of 59 patients. *J Med Genet* 44, 750–762.
- Hill-Harfe, K.L., Kaplan, L., Stalker, H.J., Zori, R.T., Pop, R., Scherer, G., and Wallace, M.R. (2005). Fine mapping of chromosome 17 translocation breakpoints > or = 900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. *Am J Hum Genet* 76, 663–671.

Hinton, R.B., Lincoln, J., Deutsch, G.H., Osinska, H., Manning, P.B., Benson, D.W., and Yutzey, K.E. (2006). Extracellular matrix remodeling and organization in developing and diseased aortic valves. *Circ Res* 98, 1431–1438.

Jakubiczka, S., Schröder, C., Ullmann, R., Volleth, M., Ledig, S., Gilberg, E., Kroisel, P., and Wieacker, P. (2010). Translocation and deletion around SOX9 in a patient with acampomelic campomelic dysplasia and sex reversal. *Sex Dev* 4, 143–149.

Jongjaroenprasert, W., Phusantisampan, T., Mahasirimongkol, S., Mushiroda, T., Hirankarn, N., Snabboon, T., Chanprasertyotin, S., Tantiwong, P., Soonthornpun, S., Rattanapichart, P., et al. (2012). A genome-wide association study identifies novel susceptibility genetic variation for thyrotoxic hypokalemic periodic paralysis. *J Hum Genet* 57, 301–304.

Lecointre, C., Pichon, O., Hamel, A., Heloury, Y., Michel-Calemard, L., Morel, Y., David, A., and Le Caignec, C. (2009). Familial acampomelic form of campomelic dysplasia caused by a 960 kb deletion upstream of SOX9. *Am J Med Genet* 149A, 1183–1189.

Ledig, S., Hiort, O., Scherer, G., Hoffmann, M., Wolff, G., Morlot, S., Kuechler, A., and Wieacker, P. (2010). Array-CGH analysis in patients with syndromic and non-syndromic XY gonadal dysgenesis: evaluation of array CGH as diagnostic tool and search for new candidate loci. *Hum Reprod* 25, 2637–2646.

Lee, Y.H., and Saint-Jeannet, J.P. (2009). Characterization of molecular markers to assess cardiac cushions formation in *Xenopus*. *Dev Dyn* 238, 3257–3265.

Leipoldt, M., Erdel, M., Bien-Willner, G.A., Smyk, M., Theurl, M., Yatsenko, S.A., Lupski, J.R., Lane, A.H., Shanske, A.L., Stankiewicz, P., et al. (2007). Two novel translocation breakpoints upstream of SOX9 define borders of the proximal and distal breakpoint cluster region in campomelic dysplasia. *Clin Genet* 71, 67–75.

Lincoln, J., Lange, A.W., and Yutzey, K.E. (2006). Hearts and bones: shared regulatory mechanisms in heart valve, cartilage, tendon, and bone development. *Dev Biol* 294, 292–302.

- Lincoln, J., Kist, R., Scherer, G., and Yutzey, K.E. (2007). Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev Biol* 305, 120–132.
- Marjamaa, A., Oikarinen, L., Porthan, K., Ripatti, S., Peloso, G., Noseworthy, P.A., Viitasalo, M., Nieminen, M.S., Toivonen, L., Kontula, K., et al. (2012). A common variant near the KCNJ2 gene is associated with T-peak to T-end interval. *Heart Rhythm* 9, 1099–1103.
- Mead, T.J., Wang, Q., Bhattaram, P., Dy, P., Afelik, S., Jensen, J., and Lefebvre, V. (2013). A far-upstream (-70 kb) enhancer mediates Sox9 auto-regulation in somatic tissues during development and adult regeneration. *Nucleic Acids Res.* 41, 4459–4469.
- Montero, J.A., Giron, B., Arrechedera, H., Cheng, Y.C., Scotting, P., Chimal-Monroy, J., Garcia-Porrero, J.A., and Hurle, J.M. (2002). Expression of Sox8, Sox9 and Sox10 in the developing valves and autonomic nerves of the embryonic heart. *Mech Dev* 118, 199–202.
- Ninomiya, S., Isomura, M., Narahara, K., Seino, Y., and Nakamura, Y. (1996). Isolation of a testis-specific cDNA on chromosome 17q from a region adjacent to the breakpoint of t(12;17) observed in a patient with acampomelic campomelic dysplasia and sex reversal. *Hum Mol Genet* 5, 69–72.
- Pfeifer, D., Kist, R., Dewar, K., Devon, K., Lander, E.S., Birren, B., Korniszewski, L., Back, E., and Scherer, G. (1999). Campomelic dysplasia translocation breakpoints are scattered over 1 Mb proximal to SOX9: evidence for an extended control region. *Am J Hum Genet* 65, 111–124.
- Pfeufer, A., Sanna, S., Arking, D.E., Müller, M., Gateva, V., Fuchsberger, C., Ehret, G.B., Orrú, M., Pattaro, C., Köttgen, A., et al. (2009). Common variants at ten loci modulate the QT interval duration in the QTSCD Study. *Nat Genet* 41, 407–414.
- Pop, R., Conz, C., Lindenberg, K.S., Blesson, S., Schmalenberger, B., Briault, S., Pfeifer, D., and Scherer, G. (2004). Screening of the 1 Mb SOX9 5' control region by array CGH identifies a large deletion in a case of campomelic dysplasia with XY sex reversal. *J Med Genet* 41, e47.

Qin, Y., Kong, L., Poirier, C., Truong, C., Overbeek, P.A., and Bishop, C.E. (2004). Long-range activation of Sox9 in Odd Sex (Ods) mice. *Hum. Mol. Genet.* **13**, 1213–1218.

Rahkonen, O., Savontaus, M., Abdelwahid, E., Vuorio, E., and Jokinen, E. (2003). Expression patterns of cartilage collagens and Sox9 during mouse heart development. *Histochem Cell Biol* **120**, 103–110.

Sekido, R., and Lovell-Badge, R. (2008). Sex determination involves synergistic action of SRY and SF1 on a specific Sox9 enhancer. *Nature* **453**, 930–934.

Tomita-Mitchell, A., Maslen, C.L., Morris, C.D., Garg, V., and Goldmuntz, E. (2007). GATA4 sequence variants in patients with congenital heart disease. *J Med Genet* **44**, 779–783.

Tommerup, N., Schempp, W., Meinecke, P., Pedersen, S., Bolund, L., Brandt, C., Goodpasture, C., Guldberg, P., Held, K.R., and Reinwein, H. (1993). Assignment of an autosomal sex reversal locus (SRA1) and campomelic dysplasia (CMPD1) to 17q24.3-q25.1. *Nat Genet* **4**, 170–174.

Velagaleti, G.V., Bien-Willner, G.A., Northup, J.K., Lockhart, L.H., Hawkins, J.C., Jalal, S.M., Withers, M., Lupski, J.R., and Stankiewicz, P. (2005). Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. *Am J Hum Genet* **76**, 652–662.

Vetro, A., Ciccone, R., Giorda, R., Patricelli, M.G., Della Mina, E., Forlino, A., and Zuffardi, O. (2011). XX males SRY negative: a confirmed cause of infertility. *J Med Genet* **48**, 710–712.

Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R.V., McKinsey, G.L., Pattabiraman, K., Silberberg, S.N., Blow, M.J., et al. (2013). A high-resolution enhancer atlas of the developing telencephalon. *Cell* **152**, 895–908.

Wagner, T., Wirth, J., Meyer, J., Zabel, B., Held, M., Zimmer, J., Pasantes, J., Bricarelli, F.D., Keutel, J., Hustert, E., et al. (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the SRY-related gene SOX9. *Cell* **79**, 1111–1120.

White, S., Ohnesorg, T., Notini, A., Roeszler, K., Hewitt, J., Daggag, H., Smith, C., Turbitt, E., Gustin, S., van den Bergen, J., et al. (2011). Copy number variation in patients with disorders of sex development due to 46,XY gonadal dysgenesis. *PLoS One* 6, e17793.

Wirth, J., Wagner, T., Meyer, J., Pfeiffer, R.A., Tietze, H.U., Schempp, W., and Scherer, G. (1996). Translocation breakpoints in three patients with campomelic dysplasia and autosomal sex reversal map more than 130 kb from SOX9. *Hum Genet* 97, 186–193.

Wunderle, V.M., Critcher, R., Hastie, N., Goodfellow, P.N., and Schedl, A. (1998). Deletion of long-range regulatory elements upstream of SOX9 causes campomelic dysplasia. *Proc Natl Acad Sci U S A* 95, 10649–10654.

Young, I.D., Zuccollo, J.M., Maltby, E.L., and Broderick, N.J. (1992). Campomelic dysplasia associated with a de novo 2q;17q reciprocal translocation. *J Med Genet* 29, 251–252.

PROJET 3

Disruption of the SEMA3D gene in a patient with congenital heart defects

Marta Sanchez-Castro^{1,2}, Olivier Pichon³, Annaig Briand³, Damien Poulain³, Véronique Gournay⁴, Albert David³, Cedric Le Caignec^{1,2,3*}

Soumis dans le journal : Human Mutation

¹INSERM, UMR1087, l'institut du thorax, Nantes, France

²Université de Nantes, Nantes, France

³CHU Nantes, Service de Génétique Médicale, Nantes, France

⁴CHU Nantes, Service de Cardiologie Pédiatrique, Nantes, France

*Correspondence to: Cédric Le Caignec, MD, PhD, Service de Génétique Médicale, CHU, 9, quai Moncousu 44093 Nantes, France, Tél.: +33 2 40 08 42 84, Fax: +33 2 40 08 39 43, E-mail: cedric.lecaignec@chu-nantes.fr

Abstract

Congenital heart disease (CHD) is the leading malformation among all newborns. Although genetic variation contributes to CHD, the genetic basis for the disease remains unknown in the majority of the patients. We report a child with transposition of the great arteries, ventricular septal defect and coarctation of the aortic isthmus. By array comparative genomic hybridization (aCGH), we identified a duplication of the 5' half of the semaphorin 3D (SEMA3D) gene. Breakpoint sequencing and fiber fluorescent in situ hybridization showed a tandem duplication. SEMA3D expression studies, performed on the patient's immortalized lymphoblasts, showed a higher level of expression of SEMA3D mRNA compared to the controls. Moreover, we demonstrated the presence of a truncated SEMA3D poly-A tailed mRNA, resulting from an abnormal transcription of the partial duplication of SEMA3D. Sema3D is an axon guidance protein essential for the correct migration of cardiac neural crest cells into the outflow tract during cardiogenesis. Sema3D null mice present with anomalous pulmonary venous connections and atrial septal defects but the role of SEMA3D in humans remains unclear. The results suggest that a truncated SEMA3D may have hampered the migration of cardiac neural crest cells during heart development, and consequently contributed to CHD in our patient.

Key Words

Congenital heart defects, Sema3D, transposition of the great arteries, coarctation of the aorta, outflow tract defect, CGH, CNV, gene duplication

Introduction

Congenital heart defects (CHD) are the most common congenital malformations with an incidence of 0.5-1% of live births (Øyen et al. 2009), and are among the leading causes of morbidity and mortality during the first year of newborns' life in developed countries (Gillum 1994; Petrini et al. 1997). Transposition of the great arteries (TGA), which accounts for 5-7% of all CHD (Digilio et al. 2001), is the most common cyanotic disorder, and the most frequent CHD diagnosed in the neonatal period, with a prevalence of 0.2 per 1000 live births. The commonly observed form of TGA is the dextro-looped type, which consists of a discordant ventriculo-arterial connection so that the aorta incorrectly arises anterior and right-sided from the right ventricle, whereas the pulmonary artery incorrectly arises posterior and left-sided from the left ventricle. Normally in the healthy heart, both outflow tracts and great vessels show a dextral (right-handed) spiralisation, but on the contrary, in TGA the great vessels present a parallel course without normal spiralisation. Apart from TGA, coarctation of the aorta (CoA), which is an outflow tract defect whereby the aorta narrows in the area where the ductus arteriosus inserts, is a relatively common defect that accounts for around 7% of all CHD (Rosenthal 2005). While TGA and CoA often occur as isolated defects, some patients present both anomalies.

Although most of the patients are successfully operated, the risk of cardiac malformation in their offspring is 10-fold higher than in the general population (Burn et al. 1998). Despite the high incidence of CHD, the aetiology of these malformations remains largely unknown. A multifactorial origin, associating both environmental and genetic factors, seems to be the usual mode of inheritance (Nora 1968; Gillum 1994; Petrini et al. 1997; Blue et al. 2012). Currently, about 20% of CHD can be attributed to the known causes, such as, chromosomal abnormalities, single gene disorders or teratogens. However, no aetiology is identified in about 80% of the patients presenting with CHD (Blue et al. 2012).

Cardiac neural crest cells play a key role in heart development. The cells migrate ventrally and form third, fourth, and sixth pharyngeal arch arteries to the cardiac outflow tract and then fuse to form the aortopulmonary septum, distally, and the conotruncal septum, proximally, to divide the outflow tract into aortic and pulmonary outflow vessels,

for simultaneously, forming the outflow valves and the subvalvular component of the mature outflow tracts (Jiang et al. 2000). One of the known causes of CHD is a disruption of neural crest cells' migration into the outflow tract (Feiner et al. 2001; Zhou et al. 2012). It was observed, in quail-chick chimera embryos, that neural crest cells ablation leads to a spectrum of pathologies involving the cardiac outflow and the great arteries among them persistent truncus arteriosus and TGA (Kirby et al. 1983; Waldo et al. 1998).

The array comparative genomic hybridization (aCGH) is a method able to detect genomic copy number variations (CNVs) (i.e. deletions and duplications), that can be used in sporadic patients and therefore able to pinpoint candidate genes. The emergence of aCGH technique has facilitated the study of sporadic cases, presenting an isolated CHD, for the identification of new candidate genes implicated in these pathologies. The recent studies have evidenced rare CNVs in patients with CoA and/or TGA using the aCGH method (van Bon et al. 2008; Mefford et al. 2008; Fakhro et al. 2011; Hitz et al. 2012; Soemedi et al. 2012; Lalani et al. 2013).

Here, we report a patient with TGA, ventricular septal defect and CoA carrying a duplication of the 5' half of SEMA3D. Genomic and mRNA expression studies suggest that the truncated SEMA3D allele hampered cardiac neural crest cell migration during embryonic heart development and thus contributed to congenital heart defects.

Methods

Patients

The proband (individual II-1) is a male, born to healthy unrelated parents (individuals I-1 and I-2) (Figure 1a). During pregnancy, the ultrasound examination, performed at 22 weeks of gestation, revealed a TGA, without any other abnormalities. At birth, the echocardiography confirmed the TGA with ventricular septal defect and a coarctation of the aorta with mild arch hypoplasia. Echocardiography also revealed that the pulmonary valve was bicuspid, slightly dysplastic with moderate leakage. Surgical repair (arterial switch, VSD closure, and repair of coarctation) of the cardiopathy was performed with success. The patient was referred to our medical genetic department at the age of 39/12 years for mild psychomotor delay and at the time of referral the patient had a mild

language delay and a mild form of attention deficit hyperactivity disorder. He acquired walking alone at 21 months. The patient did not present any dysmorphic features, extra cardiac malformations, feeding or sleeping difficulties. At physical examination at 43/12 years old, his growth measurements were normal (height: 104.5 cm; weight: 16 kg and head circumference: 50.5 cm).

SEMA3D copy number variant (CNV) detection

Informed consent for genetic analyses was obtained from all individuals participating in the study. The protocol was approved by the ethics committee of the University Hospital Center of Nantes (BRD 09/3A). Karyotyping based on RTG-banding was performed using standard methods on metaphase spreads from peripheral blood of the patient. Genomic DNA was extracted from peripheral blood using standard protocols. aCGH experiments were performed using Agilent Human Genome CGH 60K oligonucleotide arrays (Agilent, Santa Clara, CA; www.agilent.com) with the ISCA design (www.iscaconsortium.org), following the protocols provided by Agilent. The arrays were analyzed with the Agilent scanner and the Feature Extraction software (v. 9.1.3). Graphical overview was obtained using the CGH analytics software (v.3.5.14). Fluorescence in situ hybridization (FISH) was performed using RP11-349O05 probe, located in the 7q21.11 duplicated region and the CTB-164D18, located in the 7p subtelomeric region, was used as a control probe (<http://bacpac.chori.org/home.htm>) (Supp. Figure S1a).

Breakpoint sequencing

To narrow down the genomic region harbouring the breakpoints of the duplication, we designed primers for quantitative PCR (qPCR) assays of every 2-3 Kb in the candidate regions (Supp. Table S1). qPCR reactions were performed with a LightCycler®480 System (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany) and SYBR®Premix Ex Taq™ (Takara Bio Inc., Shiga, Japan). $\Delta\Delta C_t$ method was used to quantify the number of DNA copies (Livak and Schmittgen 2001). $\Delta\Delta C_t$ were calculated with the ALB gene for normalization.

To sequence the breakpoint junction fragment of the duplication and to characterize its orientation, we designed two primers at both ends of the duplication (Supp. Figure S1b). Long range PCR (LR-PCR) reactions were performed with different primer combinations using TaKaRa LA Taq™ system (Takara Bio Inc., Shiga, Japan). The LR-PCR products were sequenced with different primers, located within the amplicon, using an ABI 3730 automated sequencer.

FiberFISH analysis

Chromatin fiber stretching was performed on patient's immortalized lymphoblasts (Heiskanen et al. 1996). LR-PCR reactions were performed on a pool of control DNAs using TaKaRa LA Taq™ system (Takara Bio Inc., Shiga, Japan) to amplify two specific fiberFISH probes located at both ends of the duplication (Figure 2a). The size of the fiberFISH probes were 8 Kb and 9.5 Kb (Supp. Table S2). The LR-PCR products were labelled in Green-dUTP (ENZO®, Life Sciences Inc, NY, USA) for the probe located at the 5' end of the duplication and in Orange-dUTP (ENZO®, Life Sciences Inc, NY, USA) for the probe located at the 3' end with the Nick Translation Reagent Kit (Abbot Molecular Inc, IL, US). FISH was performed with the two probes using a protocol reported elsewhere (Le Caignec et al. 2003).

mRNA expression analyses

RNAs were extracted from immortalized lymphoblasts of patient II-1 and two normal individuals, used as controls, with RNeasy® Mini Kit (Qiagen Group, Hilden, Germany) and an RNase-free DNase I (Qiagen Group, Hilden, Germany) treatment in accordance with the manufacturer's instructions. cDNA synthesis was performed using OligodT, RNaseOUT™ Recombinant Ribonuclease Inhibitor (Life Technologies, Carlsbad, California, US) and M-MLV Reverse Transcriptase (Life Technologies, Carlsbad, California, US) following the manufacturer's instructions.

RT-qPCR reactions were performed with a LightCycler®480 System (Roche Diagnostics GmbH, Roche Applied Science, Mannheim, Germany) with 25 ng of cDNA and SYBR®Premix Ex Taq™ (Takara Bio Inc., Shiga, Japan). $\Delta\Delta C_t$ method was used

to determine the level of expression of SEMA3D transcripts (Livak and Schmittgen 2001). $\Delta\Delta C_t$ were obtained by comparing patient's cDNA versus control cDNA and the GAPDH house keeping gene for normalization.

For mRNA sequencing, primers were designed to bridge an exon-exon junction with the aim to avoid genomic DNA amplification (Supp. Table S3). Several pairs of primers were designed within the non-duplicated and duplicated regions of SEMA3D.

SEMA3D Western blot analysis

Cultured lymphoblasts of patient II-1 and two normal individuals were used for SEMA3D western blotting. Briefly, the membrane was first incubated with primary antibodies (SEMA3D Antibody, Novus Biologicals, 1:10000 dilution and GAPDH (6C5), Santa Cruz Biotechnology, 1:1000 dilution), and later with HRP-conjugated goat anti-rabbit/mouse secondary antibody (Anti-Rabbit IgG, Rockland, 1:40000 dilution; Anti-Mouse IgG, Rockland, 1:10000 dilution).

Results

Following normal karyotyping, aCGH analysis performed in patient II-1 showed a ~ 350 Kb duplication in 7q21.11 region (Figure 1b and Supp. Table S4). The duplicated region encompassed the promoter region and the exons 1 to 6 of SEMA3D (arr 7q21.11 (84,521,699x2, 84,523,431-84,866,480x3, 84,904,037x2) (coordinates are based on Human Feb. 2009 (hg19) assembly of International Human Genome Sequencing Consortium) (Supp. Table S4). The last 11 exons of SEMA3D are not included in the duplication. The duplication does not include any other genes and no other imbalance located elsewhere in the genome was identified in this patient. SEMA3D is not located in a copy-number variable region (<http://dgv.tcag.ca/dgv/app/home>).

qPCR analyses using primers located in SEMA3D confirmed the duplication in the proband. Parental qPCR analysis showed that the duplication was absent in the mother while the father was not available for analysis. Metaphase FISH analyses with RP11-349O05 BAC probe located in the duplicated region only showed two signals at 7q21.11 in both homologous chromosomes, confirming that the duplicated fragment was inserted

at the same locus as the normal one and not translocated elsewhere in the genome (Supp. Figure S1a).

To characterize the breakpoints of the duplication, a qPCR primer walking strategy was used to narrow down the breakpoint regions, followed by LR-PCR reactions. A ~3,500 bp band obtained by LR-PCR showed that the duplication was in tandem orientation (Supp. Figure S1b). Subsequent, Sanger sequencing of the LR-PCR product showed the exact location of the proximal and the distal breakpoints at 84,685,273 bp and 85,037,934 bp, respectively (Figure 2b) and its size (352,662 bp). In addition, fiberFISH experiments led us to confirm the tandem orientation (Figure 2c).

Expression studies showed that the level of expression of SEMA3D mRNA was similar to controls when using primers located in the non-duplicated region of the gene (Figure 3a). In contrast, a higher level of expression of SEMA3D mRNA compared to the controls was observed when using primers located in the duplicated part of the gene (Figure 3b). These results suggest that the 5' half of SEMA3D of the duplicated fragment was transcribed leading to a truncated mRNA of SEMA3D, in addition to normal mRNA copies transcribed from the two wild alleles. Next, a series of oligo(dT) based RT-PCR sequencing evidenced the presence of an aberrant mRNA of SEMA3D in the patient but not in controls (Supp. Figure S2). Therefore, that aberrant mRNA resulting from the SEMA3D partial duplication is likely constituted by a normal 5'UTR region, exons 1 to 6 and an abnormal 3'UTR region containing part of intron 6 of SEMA3D. These results suggest the presence of a truncated poly A tail mRNA, which escapes non-sense mediated decay (NMD), transcribed from the 5' half of SEMA3D of the duplicated allele. We were not able to detect the presence of the truncated protein by Western blot because of the lack of specificity of the antibody.

Discussion

In this study, we characterized a partial duplication of SEMA3D which was detected in a patient with complex CHD. The mRNA expression study showed that, in addition to the SEMA3D normal alleles, the partially duplicated SEMA3D gene also was transcribed to a truncated mRNA.

Class 3 semaphorins comprise 7 different secreted proteins (designated by the letters A-G). These proteins contain a Sema domain involved in the formation of a complex with neuropilin (Nrp) and plexin transmembrane co-receptors. Sema3-Nrp-Plexin complex activates the R-ras signalling pathway and thus decreases the Integrin attachment to the extracellular matrix leading cells to move (Kruger et al. 2005; Yazdani and Terman 2006). Semaphorins are best characterized during axon guidance cues in development (Pasterkamp and Kolodkin 2003) and have been implicated in multiple biological processes, such as angiogenesis (Crivellato 2011), cardiovascular patterning and development of the heart (Behar et al. 1996; Jin et al. 2006; Toyofuku et al. 2008). For example, (i) Sema3C null mice exhibit cardiac outflow tract and aortic arch abnormalities (Feiner et al. 2001), and (ii) Sema3A null mice show vascular defects (Serini et al. 2003).

The complex Sema3-Nrp-Plexin plays a major role in cardiac neural crest cell development. Sema3C through Plexin-D1 is required for the precise navigation of migrating cardiac neural crest cells (Toyofuku et al. 2008). Sema3D, in particular, is expressed in a subset of neural crest cells that contribute to cardiac development in zebrafish (Sato et al. 2006). The knockdown of Sema3D in zebrafish prevented the migration of cardiac neural crest cells which resulted in the defects in myocardial development within the primary heart field. Zebrafish mutants present with complex congenital heart defects, including hypertrophic cardiomyocytes, decreased ventricular size and defects in trabeculation and in atrioventricular valve development (Sato et al. 2006). Sema3D null mice showed an abnormal pulmonary venous connections associated with atrial septal defects (Degenhardt et al. 2013). In addition, it has been found that the disruption of the Sema3D receptors, Nrp1 or PlexinD1, causes a failure to septate the outflow tract into the ascending aorta and pulmonary trunk in null Nrp1 mice (Gu et al. 2003), or causes persistent truncus arteriosus and other patterning defects of the aortic arch arteries in null PlexinD1 mice (Gitler et al. 2004; Zhang et al. 2009). While Sema3D and receptors have been extensively studied in different animal models, their roles in humans remain unclear. A deletion of three exons of SEMA3D and two intragenic deletions in PLEXNA2 have been identified in three unrelated patients with tetralogy of Fallot (Silversides et al. 2012). Based on these data, the presence of abnormal SEMA3D in our patient may have hampered the interaction SEMA3D-NRP1-Plexin complex resulting in an activation of R-Ras signaling. Therefore, cardiac neural

crest cells may have failed to migrate properly to the primary heart during embryo heart development and contribute to the CHD of the patient.

References

- Behar O, Golden JA, Mashimo H, Schoen FJ, Fishman MC. 1996. Semaphorin III is needed for normal patterning and growth of nerves, bones and heart. *Nature* 383: 525–528.
- Blue GM, Kirk EP, Sholler GF, Harvey RP, Winlaw DS. 2012. Congenital heart disease: current knowledge about causes and inheritance. *Med. J. Aust.* 197: 155–159.
- Burn J, Brennan P, Little J, Holloway S, Coffey R, Somerville J, Dennis NR, Allan L, Arnold R, Deanfield JE, Godman M, Houston A, et al. 1998. Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* 351: 311–316.
- Crivellato E. 2011. The role of angiogenic growth factors in organogenesis. *Int. J. Dev. Biol.* 55: 365–375.
- Degenhardt K, Singh MK, Aghajanian H, Massera D, Wang Q, Li J, Li L, Choi C, Yzaguirre AD, Francey LJ, Gallant E, Krantz ID, et al. 2013. Semaphorin 3d signaling defects are associated with anomalous pulmonary venous connections. *Nat. Med.* 19: 760–765.
- Digilio MC, Casey B, Toscano A, Calabrò R, Pacileo G, Marasini M, Banaudi E, Giannotti A, Dallapiccola B, Marino B. 2001. Complete transposition of the great arteries: patterns of congenital heart disease in familial precurrence. *Circulation* 104: 2809–2814.
- Fakhro KA, Choi M, Ware SM, Belmont JW, Towbin JA, Lifton RP, Khokha MK, Brueckner M. 2011. Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc. Natl. Acad. Sci. U. S. A.* 108: 2915–2920.
- Feiner L, Webber AL, Brown CB, Lu MM, Jia L, Feinstein P, Mombaerts P, Epstein JA, Raper JA. 2001. Targeted disruption of semaphorin 3C leads to persistent truncus arteriosus and aortic arch interruption. *Dev. Camb. Engl.* 128: 3061–3070.
- Gillum RF. 1994. Epidemiology of congenital heart disease in the United States. *Am. Heart J.* 127: 919–927.

- Gitler AD, Lu MM, Epstein JA. 2004. PlexinD1 and semaphorin signaling are required in endothelial cells for cardiovascular development. *Dev. Cell* 7: 107–116.
- Gu C, Rodriguez ER, Reimert DV, Shu T, Frittsch B, Richards LJ, Kolodkin AL, Ginty DD. 2003. Neuropilin-1 conveys semaphorin and VEGF signaling during neural and cardiovascular development. *Dev. Cell* 5: 45–57.
- Heiskanen M, Kallioniemi O, Palotie A. 1996. Fiber-FISH: experiences and a refined protocol. *Genet. Anal. Biomol. Eng.* 12: 179–184.
- Hitz M-P, Lemieux-Perreault L-P, Marshall C, Feroz-Zada Y, Davies R, Yang SW, Lionel AC, D'Amours G, Lemyre E, Cullum R, Bigras J-L, Thibeault M, et al. 2012. Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet.* 8: e1002903.
- Jiang X, Rowitch DH, Soriano P, McMahon AP, Sucov HM. 2000. Fate of the mammalian cardiac neural crest. *Dev. Camb. Engl.* 127: 1607–1616.
- Jin Z, Chau MD, Bao Z-Z. 2006. Sema3D, Sema3F, and Sema5A are expressed in overlapping and distinct patterns in chick embryonic heart. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 235: 163–169.
- Kirby ML, Gale TF, Stewart DE. 1983. Neural crest cells contribute to normal aorticopulmonary septation. *Science* 220: 1059–1061.
- Kruger RP, Aurandt J, Guan K-L. 2005. Semaphorins command cells to move. *Nat. Rev. Mol. Cell Biol.* 6: 789–800.
- Lalani SR, Ware SM, Wang X, Zapata G, Tian Q, Franco LM, Jiang Z, Bucasas K, Scott DA, Campeau PM, Hanchard N, Umaña L, et al. 2013. MCTP2 is a dosage-sensitive gene required for cardiac outflow tract development. *Hum. Mol. Genet.* 22: 4339–4348.
- Le Caignec C, Winer N, Boceno M, Delnatte C, Podevin G, Liet JM, Quere MP, Joubert M, Rival JM. 2003. Prenatal diagnosis of sacrococcygeal teratoma with constitutional partial monosomy 7q/trisomy 2p. *Prenat. Diagn.* 23: 981–984.
- Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods San Diego Calif* 25: 402–408.

Mefford HC, Sharp AJ, Baker C, Itsara A, Jiang Z, Buysse K, Huang S, Maloney VK, Crolla JA, Baralle D, Collins A, Mercer C, et al. 2008. Recurrent rearrangements of chromosome 1q21.1 and variable pediatric phenotypes. *N. Engl. J. Med.* 359: 1685–1699.

Nora JJ. 1968. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* 38: 604–617.

Øyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PKA, Melbye M. 2009. Recurrence of congenital heart defects in families. *Circulation* 120: 295–301.

Pasterkamp RJ, Kolodkin AL. 2003. Semaphorin junction: making tracks toward neural connectivity. *Curr. Opin. Neurobiol.* 13: 79–89.

Petrini J, Damus K, Johnston RB Jr. 1997. An overview of infant mortality and birth defects in the United States. *Teratology* 56: 8–10.

Rosenthal E. 2005. Coarctation of the aorta from fetus to adult: curable condition or life long disease process? *Heart Br. Card. Soc.* 91: 1495–1502.

Sato M, Tsai H-J, Yost HJ. 2006. Semaphorin3D regulates invasion of cardiac neural crest cells into the primary heart field. *Dev. Biol.* 298: 12–21.

Serini G, Valdembri D, Zanivan S, Morterra G, Burkhardt C, Caccavari F, Zammataro L, Primo L, Tamagnone L, Logan M, Tessier-Lavigne M, Taniguchi M, et al. 2003. Class 3 semaphorins control vascular morphogenesis by inhibiting integrin function. *Nature* 424: 391–397.

Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, Yuen T, Rickaby J, Thiruvahindrapuram B, Marshall CR, Scherer SW, Bassett AS. 2012. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* 8: e1002843.

Soemedi R, Wilson IJ, Bentham J, Darlay R, Töpf A, Zelenika D, Cosgrove C, Setchfield K, Thornborough C, Granados-Riveron J, Blue GM, Breckpot J, et al. 2012. Contribution of global rare copy-number variants to the risk of sporadic congenital heart disease. *Am. J. Hum. Genet.* 91: 489–501.

Toyofuku T, Yoshida J, Sugimoto T, Yamamoto M, Makino N, Takamatsu H, Takegahara N, Suto F, Hori M, Fujisawa H, Kumanogoh A, Kikutani H. 2008. Repulsive and attractive semaphorins cooperate to direct the navigation of cardiac neural crest cells. *Dev. Biol.* 321: 251–262.

van Bon BWM, Koolen DA, Pfundt R, Burgt I van der, Leeuw N de, Vries BBA de. 2008. Transposition of the great vessels in a patient with a 2.9 Mb interstitial deletion of 9q31.1 encompassing the *inversin* gene: clinical report and review. *Am. J. Med. Genet. A.* 146A: 1225–1229.

Waldo K, Miyagawa-Tomita S, Kumiski D, Kirby ML. 1998. Cardiac neural crest cells provide new insight into septation of the cardiac outflow tract: aortic sac to ventricular septal closure. *Dev. Biol.* 196: 129–144.

Yazdani U, Terman JR. 2006. The semaphorins. *Genome Biol.* 7: 211.

Zhang Y, Singh MK, Degenhardt KR, Lu MM, Bennett J, Yoshida Y, Epstein JA. 2009. Tie2Cre-mediated inactivation of *plexinD1* results in congenital heart, vascular and skeletal defects. *Dev. Biol.* 325: 82–93.

Zhou J, Pashmforoush M, Sucov HM. 2012. Endothelial neuropilin disruption in mice causes DiGeorge syndrome-like malformations via mechanisms distinct to those caused by loss of *Tbx1*. *PloS One* 7: e32429.

Acknowledgments

C.L.C and M.S.C. were supported by grants from PHRC Inter regional (2008), Société Française de Cardiologie / Fédération Française de Cardiologie (2009) and Translational Research of Région des Pays de la Loire (2009). M.S.C. was also supported by École nationale supérieure des mines de Nantes and Genavie enterprise foundation. We are most grateful to the Genomics platform of Nantes (Biogenouest Genomics) core facility for its technical support.

Figures

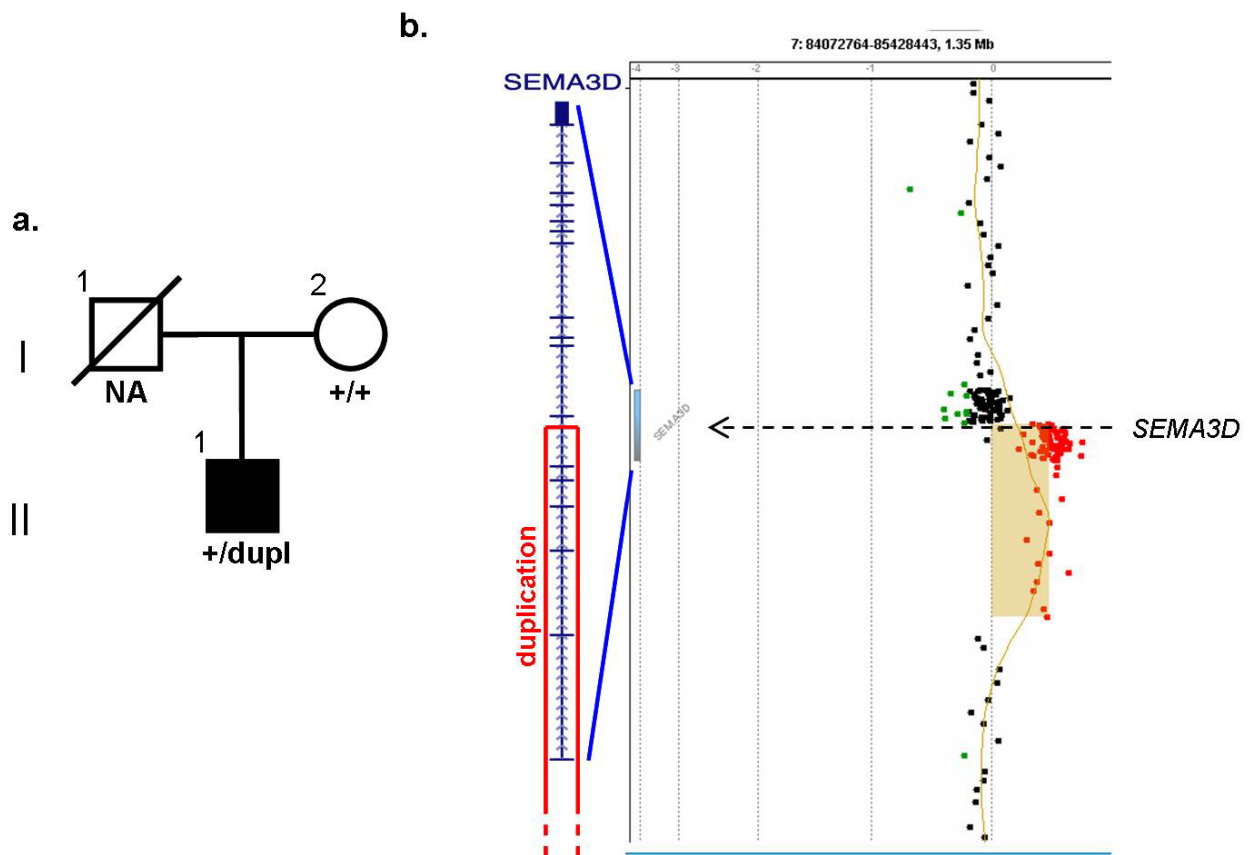


Figure 1. a. Pedigree of the family. The black square indicates the patient with congenital heart defects; the white symbols indicate the unaffected individuals; +/dupl. indicates the patient carrying the partial duplication of SEMA3D; ++ indicates the individual with a normal 7q21.11 region; NA: not available for DNA analysis. **b.** Detection of CNV in the affected patient. Left panel, ideogram representing the duplication with the first 6 exons of SEMA3D. Right panel, detailed view of the array-derived profile of chromosome 7 of patient II-1. The oligomers are arranged along the y-axis according to their position. The x-axis represents the log transformed intensity ratio. Intensity ratio corresponding to duplicated probes is colored in red. The abnormal genomic region detected by CGH Analytics algorithm for 60K arrays is colored in brown. The arrow indicates the breakpoint position between exon 6 and exon 7 of SEMA3D.

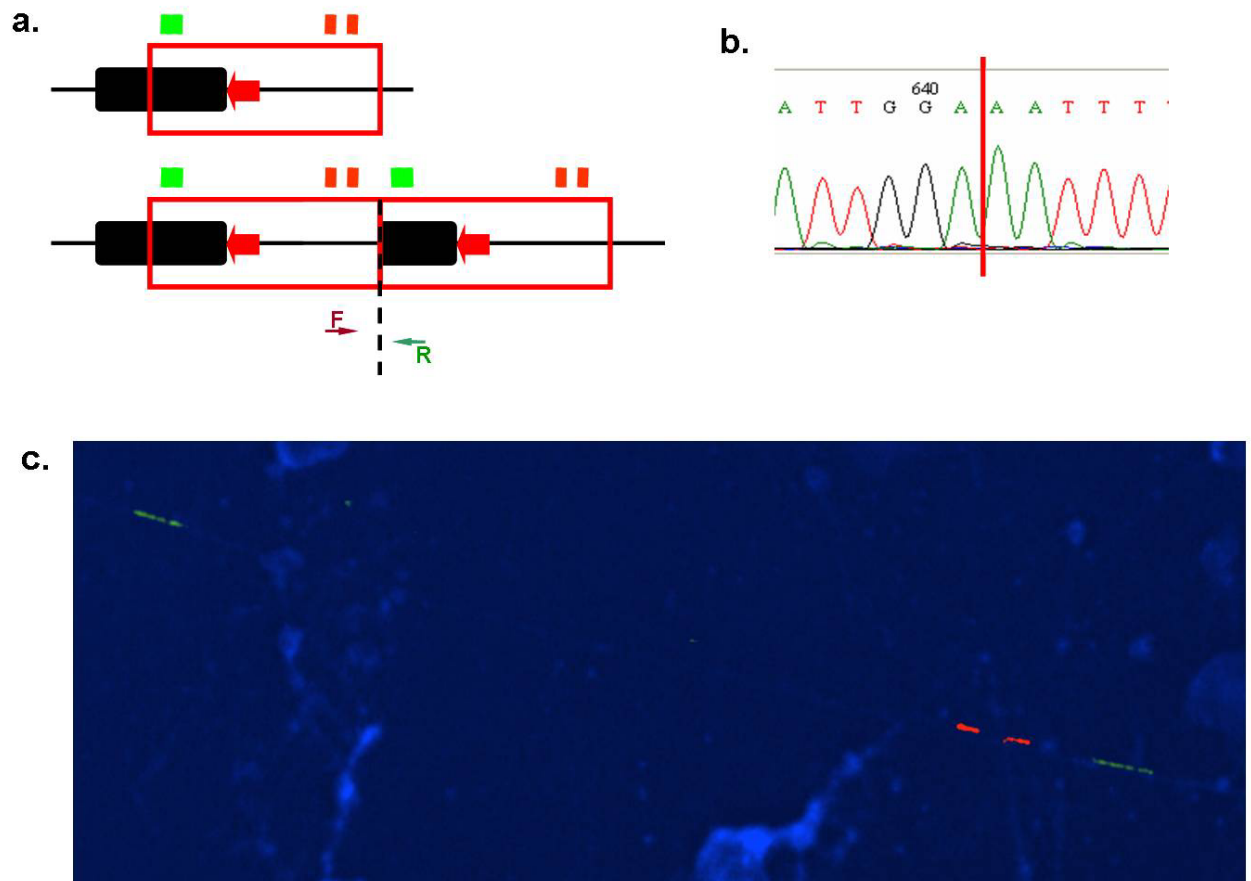


Figure 2. Breakpoint sequencing and fiberFISH analysis. **a.** Schematic representation of normal (upper panel) and duplicated (lower panel) SEMA3D loci observed in patient II-1. The green and the two red squares indicate the three fiberFISH probes. The red arrow represents the promoter region; the red rectangle indicates the duplicated region. F and R indicate the forward and reverse primers, used for breakpoint sequencing. **b.** Electropherogram of the breakpoint (red vertical arrow) of the duplication using primers F and R at 7q21.11. **c.** Fiber-FISH image of the duplicated allele in the same chromatin fiber. The green probe located on the left side of the image and the two red probes on the right side belong to the normal fragment and are closely followed by a green probe belonging to the duplicated fragment. The two red probes of the duplicated fragment could not be observed because of the long distance.

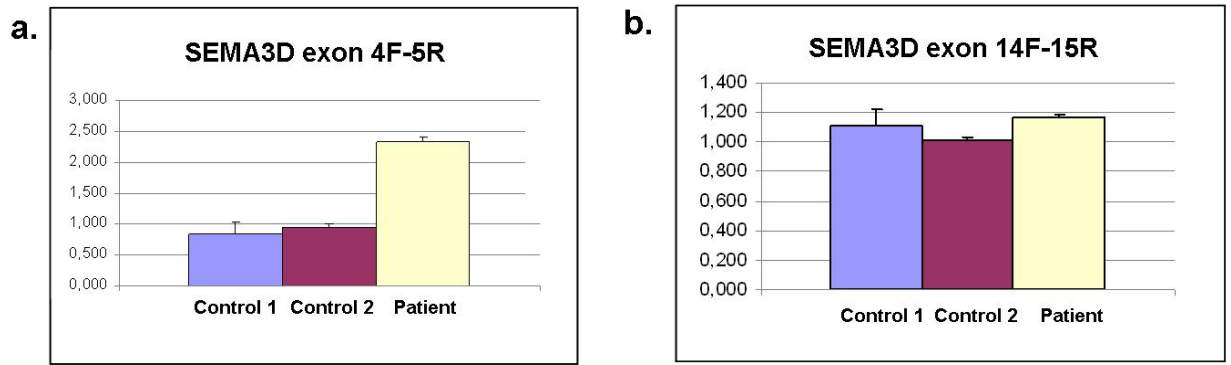
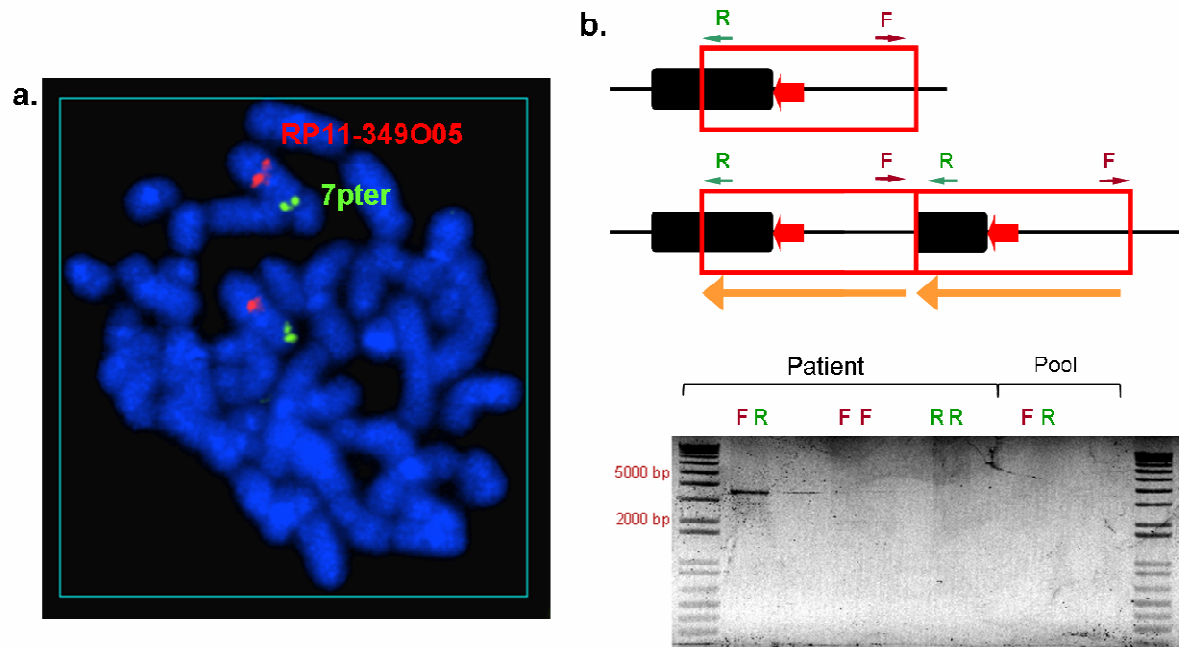
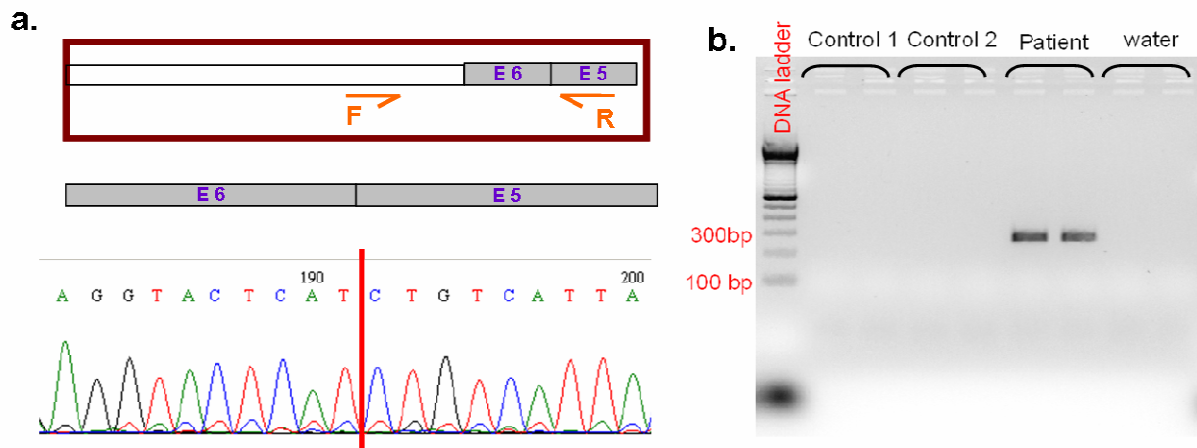


Figure 3. SEMA3D mRNA expression analyses by RT-qPCR. **a.** An increased level of expression of SEMA3D mRNA is observed in patient II-1 compared to two different controls when using primers located in the duplicated region of SEMA3D (exons 1 to 6). **b.** In contrast, comparable level of expression of SEMA3D mRNA is observed in patient II-1 compared to two different controls when using primers located in the non-duplicated region of SEMA3D (exons 7 to 17). These results suggest that the duplicated portion of SEMA3D is transcribed. Scale bars represent standard errors.

Supplementary data



Supp. Figure S1. Characterization of the duplication identified in patient II-1. **a.** Metaphase fluorescence in situ hybridization (FISH) performed on patient's immortalized lymphoblasts using the RP11-349O05 probe (colored in red) located in the chr7q21.11 duplicated region and the 7pter probe (colored in green) located in the subtelomeric region of chromosome 7 used as a control. Two red signals were observed showing that the duplication is located at the same locus as the normal one. **b.** Schematic representation of the SEMA3D alleles of the affected patient. The two primers designed to detect the possible orientations of the duplication are represented as F (in red) and R (in green) located at both ends of the duplication and orientated to the outside of the duplication. A single PCR product was obtained with the combination of primers F and R showing that the patient carries tandem duplication. LR-PCR products run in a 0.7% agarose electrophoresis gel showing a ~3500bp band only when using F-R primers.



Supp. Figure S2. mRNA SEMA3D expression. **a.** RT-qPCR using a forward primer located in the last duplicated intron and a reverse primer located in the penultimate exon of the duplicated fragment. A single PCR product of ~267 bp was observed in patient II-1 but was absent in the controls. The sequence obtained by Sanger sequencing of the amplicon matched to partial intron6 + exon6 + partial exon5 evidencing the presence of an aberrant mRNA in the affected patient. **b.** PCR products run in a 2% agarose electrophoresis gel showing a ~267bp band in the affected patient and not in the controls, evidencing the presence of an aberrant mRNA resulting from the SEMA3D gene partial duplication present in the affected patient.

Supp. Table S1. Primers designed for qPCR every 2-3 Kb intervals to narrow down the location of the duplication breakpoints. All primers were designed with Primer3-PCR primer design tool v.0.4.0.

GENE	FORWARD (5' to 3')	REVERSE (3' to 5')
SEMA3D_1	ATGAAGCCAATGCAAACCTC	ATGAAGCCAATGCAAACCTC
SEMA3D_2	GTGTCATCATGTGCCAATGC	GTGTCATCATGTGCCAATGC
SEMA3D_3	TGCCTGGTTCCAAGGACTAC	TGCCTGGTTCCAAGGACTAC
SEMA3D_4	CCATATATTTCCCCCTCCTCA	CCATATATTTCCCCCTCCTCA
SEMA3D_5	TGCTGTGGGCTTAACATCTG	TGCTGTGGGCTTAACATCTG
SEMA3D_6	CACCCCAATGGTTTCCTATG	CACCCCAATGGTTTCCTATG
SEMA3D_7	CAGAACTTCAACACCGCAAG	CAGAACTTCAACACCGCAAG
SEMA3D_8	TGACATGGAATTTGCCAGAA	TGACATGGAATTTGCCAGAA
SEMA3D_9	TAGGGCTAGCAGGCAAAGAA	TAGGGCTAGCAGGCAAAGAA
SEMA3D_10	TGCTGTGCTGAAATTCCTTC	TGCTGTGCTGAAATTCCTTC
SEMA3D_11	CACTCTTTGGCCTCCATTGT	CACTCTTTGGCCTCCATTGT
SEMA3D_12	ACCTTGCAGGGGATCTTTCT	ACCTTGCAGGGGATCTTTCT
SEMA3D_13	AGGCCACTCTGCATCTCAAT	AGGCCACTCTGCATCTCAAT

Supp. Table S2. Primers designed to amplify probes for the fiberFISH protocol. All primers were designed with Primer3-PCR primer design tool v.0.4.0.

GENE	FORWARD (5' to 3')	REVERSE (3' to 5')
SEMA3D.1C.DUP	GTGAATTAACCAGGTAACAGAT GCGTAATAGGA	AAGTCTAAATGCTCTGTAGGGA TTTGCCATT
SEMA3D.2C.DUP	AAATGTCATGGTGCTTCACTCT TATCGTCTATT	ATAAAGCTGTTTTGGATTTTATT GGACCCTACA
SEMA3D.1T.DUP	CCAGCATGGTTTCTTTTAACCC CTATTTTAAG	AACTCTGAGCAGGCTTGTGGTG TATGTAGTTTA
SEMA3D.2T.DUP	AATATTCTTACAGTATTTGCCTA AGCATTTTGCAC	GAGGTAGTAATTGCTCCCACTG TATGCAGTATC
SEMA3D.1C	TTCTGATCTATTCTGAACCAATA GTGCATTTC	TAGAAATAACCCACTGAAGATT GAGACAACACA
SEMA3D.2C	ATCTATAGTGTTTCCTTGAGGGT AATGCTGAACA	GGAAATAAGTGTTGAACCTTCTCA GGTGCATAAAC
SEMA3D.1T	ACCATGTAACAGGTGTGTTGTT TCTGATTTTAC	TATGATGTTGCAGAAATGGAAT GAAGACTTTTT
SEMA3D.2T	TACGATCTTCTAAAATGTTTCAGC AACAGTGAAA	GCTTAAAGTTGCCATACTTCCC ATTTATTGTCT

Supp. Table S3. Primers designed for RT-PCR to bridge an exon-exon junction. All primers were designed with Primer3-PCR primer design tool v.0.4.0.

GENE	FORWARD (5' to 3')	REVERSE (3' to 5')
SEMA3D_ex1-2	GTCACTGGCACTTTGAAGCA	GCCTGCCTCTTTCCTCATCT
SEMA3D_ex2-3	TCCCTTTTTGGGTTCATCAG	TTTGCAGCAGGCCAATAAAT
SEMA3D_ex3-4	ATTTATTGGCCTGCTGCAAA	TGAAATGCTCCAGTTCCACA
SEMA3D_ex4-5	TGTGTGTGGAAGTGGAGCAT	CTGCTGAGGATCGAAAGGAC
SEMA3D_ex5-6	GTCCTTTCGATCCTCAGCAG	TGGTCATGAGTAGGCCCAAG
SEMA3D_ex6-7	GATCCCTTGGGCCTACTCAT	TCGGAGGTACTGCCTTCTTG
SEMA3D_ex7-8	ACCATCCTTTCTCGAGTTGG	GTATCTGCCCCATCACTTCC
SEMA3D_ex8-9	CAACGCAGCCTGATAAACAA	TCTTTCATCTCTTGTGGGGAGT
SEMA3D_ex9-10	TCCCCACAAGAGATGAAAGAA	AACGATGGTCTGCACTTTCC
SEMA3D_ex10-11	GGAAAGTGCAGACCATCGTT	CATCACAGAGTGCCGCTTTA
SEMA3D_ex11-12	GGTGGATCATGTCATTGCAG	GCAACTCCTCCAGCACTACC
SEMA3D_ex12-14	GGTAGTGCTGGAGGAGTTGC	TCTGTGCAAGGAGAGCTGAA
SEMA3D_ex14-15	CTTGACAGATGCGACACTT	TGATTGGGTCGCCATATTTT
SEMA3D_ex15-16	AAAATATGGCGACCCAATCA	CATCCCCTGACCTCTGGATA
SEMA3D_ex16-17	TATCCAGAGGTCAGGGGATG	TGGATGAAAGTGTGCTCCTG

Supp. Table S4. CGH 60K oligonucleotide array result for the SEMA3D partial duplication in the affected patient.

First normal copy-number probe (centromeric side)	Agilent probe reference	A_18_P16158724
	Physical position of the probe (start-end)(hg19)	84683705-84683763
First duplicated probe (centromeric side)	Agilent probe reference	A_18_P16158492
	Physical position of the probe (start-end)(hg19)	84685495-84685554
Last duplicated probe (telomeric side)	Agilent probe reference	A_16_P18009686
	Physical position of the probe (start-end)(hg19)	85028485-85028544
Last normal copy-number probe (telomeric side)	Agilent probe reference	A_16_P18009758
	Physical position of the probe (start-end)(hg19)	85066101-85066160
Minimum size of the duplication (in Kb)		343,049 Kb
Maximum size of the duplication (in Kb)		382,338 Kb

DISCUSSION GÉNÉRALE ET CONCLUSION

La fréquence élevée et la grande variabilité des malformations cardiaques congénitales sont le reflet de la complexité de l'embryogenèse du cœur. L'identification de formes monogéniques de cardiopathies complexes a modifié notre vision de ces anomalies que l'on pensait au début surtout liées à des facteurs environnementaux. La mise à jour des premières interactions moléculaires initie une compréhension plus globale du développement cardiaque normal, mais également pathologique conduisant aux MCC.

Les observations issues de notre travail suggèrent un rôle important des CNVs dans la genèse des MCC. Les résultats de notre première étude basée sur la détection de CNVs chez 316 enfants atteints d'une coarctation de l'aorte (CoA), d'une transposition des gros vaisseaux (TGA) ou d'une tétralogie de Fallot (ToF) indiquent un nombre important de CNVs rares chez ces patients. La différence significative ($p < 0.01$) de CNVs *de novo* identifiés chez les patients avec CoA (4.1%) et ToF (6.6%) en comparaison des événements *de novo* identifiés chez les patients présentant une TGA (0%) pourrait être expliquée par le fait que les événements génétiques forts chez les patients avec TGA seraient moins fréquents que pour les CoA et les ToF. A l'inverse, les TGA pourraient être le résultat d'une prédisposition génétique liée à de multiples variants à effet faible associée à des facteurs environnementaux. Par ailleurs nous avons montré qu'une grande partie des CNVs rares observés dans notre étude (54/69, 77%) incluent des gènes possédant des sites de fixation à un même facteur de transcription : *FOXC1* qui est connu pour jouer un rôle essentiel dans la régulation du développement embryonnaire dans différents modèles animaux (Kume et al. 2000; Wilm et al. 2004; Siegenthaler et al. 2013) ainsi que dans le développement cardiaque du conotruncus (Kume et al. 2001; Seo and Kume 2006; Kume 2009). Ainsi, l'existence de microduplications / microdéletions incluant des sites de fixation à *FOXC1* peut altérer l'expression de gènes régulés par *FOXC1*, ce qui en fait un excellent gène candidat aux MCC.

Dans une deuxième partie de ce travail, nous avons mis en évidence des CNVs affectant des régions régulatrices intergéniques pouvant également être responsables

de MCC. En effet, nous avons identifié une délétion d'environ 1Mb en amont du gène *SOX9* chez des individus de deux familles différentes au sein desquelles plusieurs personnes présentaient une séquence de Pierre Robin (PRS) associée ou non à une MCC. Des analyses de ChIP-Seq chez la souris ont permis d'identifier, dans la région délétée, des éléments régulateurs cardiaques qui pourraient altérer l'expression du gène *SOX9* au cours du développement du cœur.

Dans la troisième partie de ce travail, nous avons décrit une duplication en tandem comprenant la région 5' du gène *SEMA3D* à l'origine d'un ARNm tronqué, susceptible de perturber la migration des cellules de la crête neurale au cours du développement du cœur.

En conclusion, le développement cardiaque fait intervenir un grand nombre de gènes interagissant de façon complexe. Des CNV incluant plusieurs gènes, un seul gène, une partie d'un gène ou des régions régulatrices de gènes impliqués dans le développement du cœur peuvent être liés aux MCC. Les CNVs rares doivent être considérés comme un des facteurs impliqués dans ces MCC en association avec d'autres variations génétiques et des facteurs environnementaux.

Dans l'avenir, des études *in vivo* réalisées à l'aide de modèles animaux permettront de mieux caractériser le rôle de chacun de ces CNV dans l'apparition de ces MCC. Par ailleurs, le séquençage des gènes candidats par des techniques de haut débit sur une grande cohorte de patients atteints de MCC serait intéressant à réaliser de façon à mieux caractériser leur rôle.

ANNEXES

I. Annexe 1. Tableau complet de correspondance entre le code IPCCC et la CIM-10.

CONGENITAL HEART DEFECTS	IPCCC	ICD-10
1. Anomalies of position of the heart		
- Heterotaxy syndromes	03.01.02	Q89.3
- Dextrocardia	02.01.02	Q24.0
- Mesocardia	02.01.04	Q24.8
- Situs inversus totalis	03.01.03	
2. Anomalies of the outflow tract and arterial valves (VA connections)		
- Transposition of the great arteries (TGA)		
- isolated TGA	01.01.02	Q20.3
- TGA with associated defect (to be added)	01.05.01	Q20.3
- Other malpositions of the great arteries		
- Double outlet right ventricle (DORV)	01.01.04	Q20.1
- with subaortic VSD	01.01.40	Q20.1
- with subpulmonary VSD	01.01.18	Q20.1
- with non-committed VSD	01.01.19	Q20.1
- with doubly committed VSD	01.01.23	Q20.1
- without VSD	01.01.24	Q20.1
- Fallot type (subaortic VSD + subpulm stenosis)	01.01.17	Q20.1
- Double outlet left ventricle	01.05.03	Q20.2
- Anatomically corrected malposition of the great arteries	01.05.10	Q25.9
- Conotruncal anomalies		
- Tetralogy of Fallot	01.01.01	Q20.3
- Pulmonary atresia with VSD	01.01.26	Q22.0+Q21.0
- with MAPCA	01.01.25	Q22.0+Q21.0
- Absent pulmonary valve syndrome		
- with VSD	09.05.28	Q22.3
- with VSD and pulmonary stenosis	09.05.25	Q22.3
- Common arterial trunk	09.01.01	Q20.0
- type I	09.01.02	Q20.0
- type II	09.01.03	Q20.0
- type III	09.01.04	Q20.0
- type IV	09.01.12	Q20.0
- Interruption of the aortic arch	09.29.31	Q25.4
- type B	09.29.33	Q25.4
- type C	09.29.34	Q25.4
- Aorto-pulmonary window	09.04.01	Q21.4
- Pulmonary artery from ascending aorta (hemitruncus)	09.09.08	Q25.7
- Right PA from ascending aorta	09.09.03	Q25.7
- Left PA from ascending aorta	09.09.05	Q25.7
- Left ventricular outflow tract anomalies		
- valvular aortic stenosis	09.15.01	Q23.0
- bicuspid aortic valve	09.15.22	Q23.1
- dysplastic aortic valve	09.15.09	Q23.8
- aortic valve atresia	09.15.06	Q23.0
- subvalvular aortic stenosis (muscular)	07.09.16	Q24.4
- subaortic fibrous shelf	07.09.03	Q24.4
- congenital aortic insufficiency	09.15.07	Q23.1
- Right ventricular outflow tract anomalies		

- pulmonary valvular stenosis	09.05.04	Q22.1
- dysplastic pulmonary valve	09.05.24	Q22.3
- bicuspid pulmonary valve	09.05.32	
- pulmonary atresia associated with another CHD	09.05.11	Q22.0
- subpulmonary stenosis (RVOT obstruction)	07.05.01	Q24.3
- stenosis of ostium infundibuli	07.05.02	Q24.3
- deviation of conal septum	07.05.04	Q24.3
- double chambered right ventricle	07.03.01	Q24.8
3. Anomalies of the atrioventricular valves (AV connections)		
- Congenital anomalies of the tricuspid valve (TV)		
- Ebstein's anomaly	06.01.34	Q22.5
- Associated with congenitally corrected TGA	06.01.75	Q22.5+Q20.5
- Tricuspid valve agenesis (unguarded tricuspid orifice)	06.01.32	Q22.9
- Dysplastic tricuspid valve	06.01.39	Q22.8
- Congenital tricuspid insufficiency	06.01.25	Q22.8
- Cleft of the tricuspid valve	06.01.36	Q22.8
- Congenital tricuspid stenosis	06.01.92	Q22.4
- Tricuspid annulus hypoplasia	06.01.04	Q22.4
- Straddling tricuspid valve	06.01.09	Q22.8
- Overriding tricuspid valve	06.01.05	Q22.8
- Anomalies of the subvalvular tensor apparatus of the tricuspid valve	06.01.51	Q22.8
- Congenital anomalies of the mitral valve (MV)		
- Dysplastic mitral valve	06.02.40	Q23.8
- Congenital mitral insufficiency	06.02.25	Q23.3
- Isolated cleft of the mitral valve (not AVSD type)	06.02.36	Q23.9
- Mitral valve prolapse	06.02.35	I34.1
- Congenital mitral stenosis	06.02.16	Q23.2
- Supravalvular mitral ring	05.02.02	Q20.8
- Mitral annulus hypoplasia	06.02.04	Q23.2
- Parachute mitral valve	06.02.56	Q23.8
- Double mitral orifice	06.02.33	Q23.9
- Straddling mitral valve	06.02.09	Q23.8
- Overriding mitral valve	06.02.05	Q23.8
- Anomalies of the subvalvular tensor apparatus of the mitral valve	06.02.51	Q23.8
- Accessory mitral tissue	06.02.39	Q23.9
- Atrioventricular septal defects (AVSD)		
- Complete AVSD	06.06.09	Q21.2
- with LV hypoplasia	06.07.05	Q21.2+Q20.8
- with RV hypoplasia	06.07.06	Q21.2+Q20.8
- with Tetralogy of Fallot	01.01.20	Q21.2+Q20.3
- with LVOT obstruction	07.09.17	Q21.2+Q24.4
- Partial AVSD "typical" (ostium primum type)	06.06.01	Q21.2
- intermediate type	06.06.10	Q21.2
- cleft of the mitral valve, AVSD type	06.05.69	Q23.9
4. Functionally univentricular hearts		
- Double-inlet ventricles (univentricular hearts)		

- with 2 atrioventricular valves	01.01.14	Q20.4
- Double-inlet right ventricle	01.04.03	Q20.4
- Double-inlet left ventricle	01.04.04	Q20.4
- with 1 atrioventricular valve		
- Mitral atresia	01.04.13	Q20.4+Q23.2
- Tricuspid atresia	01.04.12	Q20.4+Q22.4
- Common AV valve	01.06.02	Q20.4+Q21.2
- Left ventricular hypoplasia		
- Hypoplastic left heart syndrome	01.01.09	Q23.4
- Mitral atresia	06.02.01	Q23.2
- LV hypoplasia, others	07.08.42	Q20.8
- Right ventricular hypoplasia		
- Pulmonary atresia with intact ventricular septum	01.01.07	Q22.0
- with severe RV hypoplasia	07.02.01	Q22.0
- with sinusoids	07.01.13	Q22.0
- Tricuspid atresia	06.01.01	Q22.4
- Hypoplastic right heart syndrome	07.02.01	Q22.6
- RV hypoplasia, others	07.08.41	Q20.8
- Uhl's anomaly	07.01.06	Q24.8
5. Anomalies of the great arteries (aorta and pulmonary artery)		
- Anomalies of the ascending aorta		
- Hypoplasia of the aortic arch (tubular)	09.29.11	Q25.4
- Atresia of the ascending aorta	09.16.03	Q25.2
- supravalvular aortic stenosis	09.16.00	Q25.3
- Aortic sinus of Valsalva aneurysm	09.18.01	Q25.4
- aorto-left ventricular tunnel	09.17.02	Q25.4
- Coarctation of the aorta		
- coarctation of the aorta, isthmus	09.29.01	Q25.1
- coarctation of the aorta, abdominal	09.29.06	Q25.1
- interruption of the aortic arch	09.29.31	Q25.4
- interruption of the aortic arch, type A	09.29.32	Q25.4
- Anomalies of the aortic arches		
- right aortic arch	09.30.26	Q25.4
- cervical aortic arch	09.28.06	Q25.4
- anomalous subclavian artery	09.30.01	Q27.8
- aberrant origin of right subclavian artery	09.30.02	Q27.8
- aberrant origin of left subclavian artery	09.30.04	Q27.8
- arterial duct from left innominate artery	09.27.47	Q27.8
- double aortic arch	09.28.09	Q25.4
- persistent fifth aortic arch	09.29.08	Q25.4
- Anomalies of the pulmonary artery and its branches		
- Supravalvular pulmonary trunk stenosis	09.07.13	Q25.6
- Pulmonary arterial stenosis	09.10.01	Q25.6
- Right PA stenosis	09.10.02	Q25.6
- Left PA stenosis	09.10.04	Q25.7
- Pulmonary arterial hypoplasia	09.10.14	Q25.7
- Right PA hypoplasia	09.10.12	Q25.7
- Left PA hypoplasia	09.10.13	Q25.7
- Pulmonary artery absent	09.10.21	Q25.7

- Right PA absent	09.10.22	Q25.7
- Left PA absent	09.10.23	Q25.7
- Pulmonary artery from patent arterial duct	09.09.11	Q25.7
- Right PA from patent arterial duct	09.09.02	Q25.7
- Left PA from patent arterial duct	09.09.04	Q25.7
- Pulmonary arterial sling	09.09.06	Q25.7
6. Congenital anomalies of the coronary arteries	09.46.00	Q24.5
- Anomalous connection of left coronary artery to pulmonary artery	09.41.03	Q24.5
- Single coronary artery	09.43.14	Q24.5
- Coronary orifice stenosis	09.44.05	Q24.5
- Coronary orifice atresia	09.44.06	Q24.5
- Left anterior descending from right coronary artery (in TOF)	09.43.04	Q24.5
- Coronary artery fistula	09.45.16	Q24.5
- within the pulmonary artery	09.45.06	Q24.5
7. Ventricular septal defects (VSD)	07.10.00	Q21.0
- Membranous VSD	07.10.01	Q21.0
- Membranous VSD, small	07.15.03	Q21.0
- Conoventricular (malalignment) VSD	07.10.12	Q21.0
- Doubly committed (subarterial) VSD	07.12.01	Q21.0
- Muscular VSD	07.11.01	Q21.0
- Muscular VSD, small	07.15.02	Q21.0
- Inlet VSD	06.06.08	Q21.0
- Multiple VSD	07.15.04	Q21.0
- Associated with aortic insufficiency	07.15.15	Q21.0
8. Anomalies of the atria and atrial septal defects		
- Cor triatriatum	05.02.01	Q24.2
- Cor triatriatum dexter	05.01.21	Q24.2
- Atrial septal defects (excepted ostium primum type)	05.04.01	Q21.1
- patent oval foramen	05.03.01	Q21.1
- ostium secundum type	05.04.02	Q21.1
- sinus venosus type (superior)	05.05.01	Q21.1
- coronary sinus type	05.05.03	Q21.1
- common atrium	05.06.01	Q21.2
- Juxtaposition of the atrial appendages	05.00.01	Q20.8
- Left JAA	05.01.06	Q20.8
- Right JAA	05.02.04	Q20.8
- Isomerism of the atrial appendages	01.03.08	Q20.6
- left isomerism	01.03.02	Q20.6
- Right isomerism	01.03.03	Q20.6
9. Anomalies of the venous return		
- Anomalies of the systemic venous return		
- Anomalies of the superior vena cava (SVC)	04.01.00	Q26.9
- Absent right SVC	04.01.05	Q26.8
- Retro-aortic innominate vein	04.00.04	Q26.9
- Left SVC persisting to coronary sinus	04.01.01	Q26.1
- Left SVC persisting to left atrium	04.04.02	Q26.1

- Anomalies of the coronary sinus	04.04.00	Q21.1
- Partially unroofed coronary sinus	04.04.01	Q21.1
- Totally unroofed coronary sinus	04.04.02	Q21.1
- Coronary sinus orifice atretic	04.04.03	Q21.1
- Anomalies of the inferior vena cava (IVC)	04.03.00	Q26.9
- Azygos continuation of the IVC	04.03.10	Q26.8
- Right-sided azygos continuation of the IVC	04.03.02	Q26.8
- Left-sided azygos continuation of the IVC	04.03.01	Q26.8
- Anomalies of the supra-hepatic veins	04.02.00	Q26.8
- Separate hepatic and IVC venous return	04.02.06	Q26.8
- Anomalies of the pulmonary venous connections		
- Anomalous pulmonary venous return	04.08.07	Q26.2
- Totally	04.08.05	Q26.2
- Supracardiac	04.06.00	Q26.2
- Intracardiac	04.08.10	Q26.2
- Infracardiac	04.08.20	Q26.2
- Miscellaneous	04.08.30	Q26.2
- Partially	04.07.01	Q26.3
- Congenital pulmonary vein stenosis	04.08.01	Q26.8
- Congenital pulmonary vein hypoplasia	04.08.08	Q26.8
- Pulmonary vein atresia	04.08.02	Q26.8
- Scimitar syndrome	01.01.16	Q26.8
10. Complex anomalies of atrioventricular connections		
- Congenitally corrected TGA (double discordance)	01.01.03	Q20.5
- Criss-cross atrioventricular connections	02.03.03	Q24.8
- Supero-inferior ventricles	02.04.00	Q24.8

II. Annexe 2. Facteurs de risque non génétiques et MCC (Patel and Burns 2013).

II. Annexe 2. Table 1

Table 1 Risk factors associated with congenital heart defects: characteristics and conditions

Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Age				
Advanced maternal age	MACDP (1968–2005) [112]	739	1,301,143	1.2 (1.1–1.3)
	PaRCM (1998–2002) [109]	4,337	8,683	1.1 (1.0–1.1)
	PHHS (1988–1994) [78]	3,757	102,728	4.0 (1.7–9.2)
	NSW/ACT (1981–1984) [86]	1,479	343,521	1.3 (1.1–1.4)
	BWIS (1981–1989) [59]	3,377	3,572	1.3 (1.1–1.5)
Advanced paternal age	FRCM (1982–1983) [147]	408	755	1.3 (1.0–1.6)
	PRCM (1998–2002) [109]	3,933	8,683	1.1 (1.0–1.3)
	NCHS (1999–2000) [167]	9,767	77,514	1.2 (1.1–1.4)
Young paternal age	EHS (1995–1997) [13]	894	894	1.5 (1.2–2.0)
	MBRN (1967–1998) [85]	3,656	1,869,388	1.5 (1.0–2.0)
Diabetes mellitus (pre-gestational)	Bei/Heb provinces (1988) [167]	497	6,222	2.3 (1.9–2.8)
	EUROCAT (1990–2005) [62]	323	92,976	2.2 (1.9–2.6)
	HCCSCA (1980–1996) [11]	4,480	38,151	2.5 (1.6–3.9)
	NBDPS (1997–2003) [38]	3,519	4,689	4.6 (2.9–7.5)
	Milwaukee (1997–1999) [169]	245	3,780	4.1 (1.5–11.2)
	BWIS (1981–1989) [101]	3,377	3,572	3.0 (1.9–4.8)
	WA State (1984–1991) [79]	1,511	8,934	4.0 (3.1–5.1)
	SMBR (1981–1986) [121]	1,324	2,648	2.7 (1.4–5.0)
Febrile illness	BWIS (1981–1989) [117]	2,361	3,435	1.1 (0.9–1.5)
	HCCSCA (1980–1996) ^a [2]	4,480	38,151	2.6 (1.2–5.4)
	Shan province (2004–2005) ^a [98]	164	328	5.9 (2.7–13.1)
	NBDPS (1997–2003) ^b [34]	3,690	4,760	1.7 (1.0–3.0)
	ABDCCS (1968–1980) [20]	829	3,029	1.8 (1.4–2.4)
	SBDMP (1986–1987) ^c [171]	986	990	1.4 (0.7–2.9)
	FRCM (1982–1983) ^c [147]	583	756	1.5 (1.1–2.1)
Hypercholesterolemia	HAVEN (2003–2010) [142]	229	320	1.5 (1.1–2.1)
Hyperhomocysteinemia	ARHMS (1998–2008) [74]	417	250	1.5 (1.2–1.8)
	HAVEN (2003–2005) [156]	151	183	2.9 (1.4–6.0)
Hypertension	APGC (1988–1998) [160]	26	116	3.5 (1.2–10.2)
	KPNCA (1995–2008) [94]	6,873	453,078	1.4 (1.3–1.5)
	HCCSCA (1980–1996) [44]	4,480	38,151	1.3 (1.0–1.5)
	NBDPS (1997–2003) [30]	5,021	4,796	1.8 (1.1–2.7)
	Milwaukee (1997–1999) [169]	245	3,780	2.8 (1.2–6.7)
Influenza	FRCM (1982–1983) [147]	583	756	2.6 (1.1–5.2)
	BWIS (1981–1989) [117]	2,361	3,435	1.1 (0.9–1.4)
	HCCSCA (1980–1996) [1]	4,480	38,151	1.7 (1.3–2.3)
	ABDCCS (1968–1980) [20]	829	3,029	2.1 (0.8–5.5)
Pregpregnancy weight				
Overweight (BMI 25–30)	NBDPS (1997–2004) [63]	6,440	5,673	1.2 (1.1–1.3)
	SMBR (1995–2007) [16]	11,163	1,235,877	1.1 (1.0–1.1)
	ABDRFSS (1993–1997) [158]	195	330	2.0 (1.2–3.1)
Overweight + obesity (BMI > 25)	ENN (1997–2008) [9]	797	322	1.1 (0.8–1.5)
	NBDPS (1997–2004) [63]	6,440	5,673	1.2 (1.1–1.3)
	ABDCCS (1968–1980) [157]	851	2,767	1.4 (1.0–1.9)

Table 1 continued

Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Obese (BMI > 30)	WCHARS (1992–2007) [103]	14,412	141,420	1.2 (1.2–1.3)
	NBDPS (1997–2004) [63]	6,440	5,673	1.3 (1.2–1.4)
	NYSCMR (1993–2003) [113]	7,392	56,304	1.2 (1.1–1.2)
	SMBR (1995–2007) [16]	11,163	1,235,877	1.2 (1.1–1.2)
	WABDR (1997–2000) [116]	111	418	1.3 (0.6–2.7)
	ABDRFSS (1993–1997) [158]	195	330	2.0 (1.2–3.4)
	ABDCCS (1968–1980) [157]	851	2,767	1.3 (0.8–2.1)
	MBR (1990–1994) [124]	1,451	8,088	0.8 (0.3–1.9)
Severely obese (BMI > 40)	SMBR (1995–2007) [16]	11,163	1,235,877	1.5 (1.2–1.8)
Reproductive history				
Infertility/ART	PaRCM (1987–2006) [146]	5,493	3,104	1.4 (1.1–1.7)
Parity				
Multiparous	NBDPS (1997–2007) [56]	7,575	7,954	1.0 (0.9–1.1)
	PRCM (2005–2006) [110]	1,673	4,017	1.2 (1.1–1.4)
	Shan province (2004–2005) [98]	164	328	2.5 (1.5–4.0)
	BWIS (1981–1989) [59]	3,377	3,572	1.3 (1.1–1.5)
Nulliparous	NBDPS (1997–2007) [56]	7,575	7,954	1.1 (1.0–1.2)
Socioeconomic status (low)	Kaunas (1999–2005) [90]	187	643	3.4 (1.5–7.6)
	DNBC (1997–2002) [155]	659	81,453	1.6 (1.3–2.0)
Stress	Shan province (2004–2005) [98]	164	328	2.7 (1.5–4.8)

All exposures are maternal unless otherwise specified

Significant values are in bold

OR odds ratio, RR relative risk, CI confidence interval, BMI body mass index, ART assisted reproductive technique, ABDCCS Atlanta Birth Defects Case-Control Study, ABDRFSS Atlanta Birth Defects Risk Factor Surveillance Study, ARHMS Arkansas Reproductive Health Monitoring System, APGC University of Alabama Prenatal Genetics Clinic, Bei/Heb provinces Beijing and Hebei provinces, BWIS Baltimore-Washington Infant Study, DNBC Danish National Birth Cohort, EHS Egyptian Health Service, ENN Eurocat Northern Netherlands Birth Registry, EUROCAT all 18 registries, FRM Finnish Register of Congenital Malformations, HAVEN Heart Defects, Vascular Status, Genetic Factors and Nutrition Study, HCCSCA Hungarian Case-Control Surveillance of Congenital Abnormalities, KPNCa Kaiser Permanente Northern California, MACDP Metropolitan Atlanta Congenital Defects Program, MBR Mainz (Germany) Birth Registry, NBDPS National Birth Defect Prevention Study, NCHS National Center for Health Statistics, MBRN Medical Birth Registry of Norway, NSW/ACT New South Wales/Australia Capitol Territory, NYSCMR New York State Congenital Malformations Registry, PaRCM Paris Registry of Congenital Malformations, PHHS Parkland Health and Hospital System, PRCM Polish Registry of Congenital Malformations, SBDMP Shanghai Birth Defects Monitoring Program, Shan province Shandong province, SMBR Swedish Medical Birth Registry, WABDR Western Australia Birth Defect Registry, WCHARS Washington Comprehensive Hospital Abstract Reporting System

^a Pelvic inflammatory disease

^b Urinary tract infection

^c URI upper respiratory infection

II. Annexe 2. Table 2

Table 2 Risk factors associated with congenital heart defects: therapeutic drug exposures

Exposure	Source	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Anti-asthmatic medication	Lin et al. [96]	NYSCMR (1988–1991)	502	1,066	2.4 (1.2–4.8)
	Kallen et al. [83]	SMBR (1995–2004)	11,367	585,372	1.1 (1.0–1.3)
Bronchodilators	Lin et al. [96]	NYSCMR (1988–1991)	502	1,066	2.2 (1.1–4.6)
	Kallen et al. [83]	SMBR (1995–2004)	11,367	585,372	1.4 (1.1–1.7)
Antibiotics (any)	Crider et al. [41]	NBDPS (1997–2003)	5,269	4,941	1.1 (1.0–1.3)
	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.1 (0.9–1.3)
Macrolides	Crider et al. [41]	NBDPS (1997–2003)	5,269	4,941	1.0 (0.7–1.3)
Penicillins	Crider et al. [41]	NBDPS (1997–2003)	5,269	4,941	1.0 (0.8–1.1)
	Denker et al. [51]	NJC (1991–1998)	4,055	9,263	1.7 (0.8–3.7)
Sulfonamides	Ferencz et al. [59]	BWIS (1981–1989)	3,377	3,572	1.3 (1.1–1.6)
	Crider et al. [41]	NBDPS (1997–2003)	5,269	4,941	1.5 (1.0–2.2)
	Matok et al. [111]	Clalit HMO (1998–2007)	571	85,250	1.8 (1.1–3.0)
	Czeizal et al. [47]	HCCSCA (1980–1996)	4,467	38,151	2.1 (1.4–3.3)
Sulfonamide with folic acid	Hernandez-Diaz et al. [72]	SEUBDS (1976–1988)	3,870	8,387	3.4 (1.8–6.4)
	Czeizal et al. [47]	HCCSCA (1980–1996)	4,467	38,151	1.2 (0.9–1.6)
	Hernandez-Diaz et al. [72]	SEUBDS (1976–1988)	3,870	8,387	1.5 (0.6–3.8)
Anticonvulsants	Werler et al. [162]	NBDPS (1997–2005)	7,093	6,052	1.4 (1.0–2.1)
	Matok et al. [111]	Clalit HMO (1998–2007)	571	85,250	0.7 (0.2–2.9)
	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.6 (1.0–2.5)
	Hernandez-Diaz et al. [72]	SEUBDS (1976–1988)	3,870	8,387	2.2 (1.4–3.5)
With folic acid	Hernandez-Diaz et al. [72]	SEUBDS (1976–1988)	3,870	8,387	2.3 (1.1–4.7)
Antidepressants (any)	Ferencz et al. [59]	BWIS (1981–1989)	3,377	3,572	3.0 (1.2–7.6)
Bupropion	Alwan et al. [6]	NBDPS (1997–2004)	6,853	5,869	1.1 (0.7–1.9)
SSRI (any)	Malm et al. [106]	FRCM (1996–2003)	8,253	635,583	1.3 (1.1–1.6)
	Reis et al. [128]	SMBR (1995–2007)	14,821	1,062,190	1.0 (0.8–1.2)
	Alwan et al. [5]	NBDPS (1997–2002)	4,268	4,092	0.9 (0.7–1.2)
Paroxetine	Louik et al. [102]	SEUBDS (1993–2004)	3,724	5,860	1.2 (0.9–1.6)
	Malm et al. [106]	FRCM (1996–2003)	8,253	635,583	1.3 (0.8–2.1)
	Bakker et al. [10]	ENN (1997–2006)	678	615	1.5 (0.6–4.2)
	Reis et al. [128]	SMBR (1995–2007)	1,208	1,062,190	1.7 (1.1–2.5)
	Louik et al. [102]	SEUBDS (1993–2004)	3,724	5,860	1.4 (0.8–2.5)
Fluoxetine	Louik et al. [106]	FRCM (1996–2003)	8,253	635,583	1.6 (1.1–2.2)
	Malm et al. [102]	SEUBDS (1993–2004)	3,724	5,860	0.9 (0.6–1.5)
SNRI (Venlafaxine)	Polen et al. [120]	NBDPS (1997–2007)	8,069	8,002	2.7 (1.5–5.0)
TCA	Reis et al. [128]	SMBR (1995–2007)	1,662	1,062,190	1.6 (1.1–2.4)
Antifungals (metronidazole)	Ferencz et al. [59]	BWIS (1981–1989)	3,377	3,572	2.5 (1.1–5.8)
Antihypertensives	Banhidy et al. [12]	HCCSCA (1980–1996)	4,480	38,151	1.3 (1.0–1.5)
	Li et al. [94]	KPNCA (1995–2008)	6,873	453,078	1.5 (1.0–2.2)
	Caton et al. [30]	NBDPS (1997–2003)	5,021	4,796	1.8 (1.1–2.7)
	Lennestall et al. [93]	SMBR (1982–2006)	1,418	1,045,425	2.6 (1.9–3.5)
β-Blockers	Caton et al. [30]	NBDPS (1997–2003)	5,021	4,796	2.6 (1.2–5.3)
	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.9 (1.2–2.8)
ACE inhibitors	Li et al. [94]	KPNCA (1995–2008)	6,873	453,078	1.5 (0.9–2.6)
	Caton et al. [30]	NBDPS (1997–2003)	5,021	4,796	1.9 (0.5–7.2)
	Lennestall et al. [93]	SMBR (1982–2006)	1,418	1,045,425	2.9 (0.9–6.8)
	Cooper et al. [36]	TN Medicaid (1985–2000)	209	29,096	3.7 (1.9–7.3)

Table 2 continued

Exposure	Source	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Folic acid (multivitamin)	Correa et al. [39]	NBDPS (1997–2004)	5,206	4,737	0.9 (0.8–1.0)
	Van Beynum et al. [153]	EUROCAT (1996–2005)	611	3,343	0.7 (0.6–0.9)
	Smedts et al. [141]	HAVEN (2003–2006)	276	324	1.3 (0.9–1.9)
	Czeizel et al. [48]	HPS (1993–1996)	31	50	0.6 (0.4–1.0)
	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.2 (1.0–1.6)
	Botto et al. [18]	ABDCCS (1968–1980)	958	3,029	0.8 (0.6–1.0)
	Czeizel et al. [45]	HCCSCA (1980–1991)	2,976	30,663	0.9 (0.8–1.0)
NSAIDs (any)	van Gelder et al. [154]	MoBa (1999–2007)	435	66,662	1.0 (0.7–1.6)
	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.2 (1.0–1.6)
Aspirin	van Gelder et al. [154]	MoBa (1999–2007)	435	66,662	1.6 (0.5–5.2)
	Czeizel et al. [46]	HCCSCA (1980–1996)	4,056	38,151	1.4 (0.9–2.1)
Ibuprofen	van Gelder et al. [154]	MoBa (1999–2007)	435	66,662	0.9 (0.5–1.5)
	Ferencz et al. [59]	BWIS (1981–1989)	3,377	3,572	1.4 (1.1–1.8)
Naproxen	Kallen et al. [82]	SMBR (1995–2001)	5,565	577,730	1.7 (1.1–2.5)
Thyroid replacement	Winker et al. [163]	SMBR (1995–2004)	11,028	848,468	1.3 (1.1–1.5)

All exposures are maternal unless otherwise specified

Significant values are in bold

OR odds ratio, RR relative risk, CI confidence interval, SSRI selective serotonin receptor inhibitors, SNRI selective serotonin receptor inhibitors, TCA tricyclic antidepressants, ACE angiotensin-converting enzyme, NSAIDs nonsteroidal anti-inflammatory drugs, ABDCCS Atlanta Birth Defects Case-Control Study, WIS Baltimore-Washington Infant Study, ENN Eurocat Northern Netherlands Birth Registry, EUROCAT all 18 registries, FRCM Finnish Register of Congenital Malformations, HAVEN Heart Defects, Vascular Status, Genetic Factors and Nutrition Study, HCCSCA Hungarian Case-Control Surveillance of Congenital Abnormalities, HPS Hungarian Periconceptional Service, KPNCa Kaiser Permanente Northern California, MoBa Norwegian Mother and Child Cohort Study, NBDPS National Birth Defect Prevention Study, NJC North Jutland County, Denmark, NYSCMR New York State Congenital Malformations Registry, SEUBDS Sloane Epidemiology Unit Birth Defects Study, SMBR Swedish Medical Birth Registry, TN Medicaid Tennessee Medicaid

II. Annexe 2. Table 3

Table 3 Risk factors associated with congenital heart defects: nontherapeutic drug exposures

Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Alcohol	PRAMS (1996–2005) [108]	237	948	3.0 (1.2–7.5)
Cigarette smoking				
Maternal use	BWIS (1981–1989) [4]	2,525	3,435	1.1 (1.0–1.2)
	Patras (2006–2009) [84]	157	208	2.8 (1.8–4.6)
	NBDPS (1997–2002) [105]	3,067	3,947	1.2 (1.1–1.4)
	ARHMS 1998–2004 [73]	275	118	1.7 (1.0–3.1)
	SMBR (1983–1996) [81]	3,384	1,413,811	1.1 (1.0–1.2)
Paternal use	Italy (2008–2010) [40]	360	360	1.7 (1.1–2.6)
Cocaine				
Maternal use	BWIS (1981–1989) [59]	3,377	3,572	1.6 (1.1–2.3)
	Boston City (1988–1990) [97]	49	505	3.7 (1.4–9.4)
Paternal use	BWIS (1981–1989) [59]	3,377	3,572	1.7 (1.3–2.2)
Marijuana (paternal)	BWIS (1981–1989) [59]	3,377	3,572	1.2 (1.1–1.4)
Vitamin E	HAVEN (2003–2006) [141]	276	324	1.7 (1.0–2.6)

All exposures are maternal unless otherwise specified

Significant values are in bold

OR odds ratio, RR relative risk, CI confidence interval, ARHMS Arkansas Reproductive Health Monitoring System, BWIS Baltimore-Washington Infant Study, HAVEN Heart Defects, Vascular Status, Genetic Factors and Nutrition Study, NBDPS National Birth Defect Prevention Study, Patras University of Patras, Greece, PRAMS Pregnancy Risk Assessment Monitoring System, SMBR Swedish Medical Birth Registry

II. Annexe 2. Table 4

Table 4 Risk factors associated with congenital heart defects: environmental exposures

Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Air pollution				
Carbon monoxide	NCAS (1993–2003) [49]	2,140	14,256	1.2 (0.9–1.7)
Chemical exposures				
Maternal cyanide exposure	San Francisco (1983–1985) [136]	5,046	20,882	2.2 (1.3–3.9)
Maternal heavy metal exposure	HAVEN (2003–2010) [143]	424	480	0.9 (0.2–3.8)
	San Francisco (1983–1985) [136]	5,046	20,882	1.5 (1.1–2.3)
Maternal pesticide exposure	HAVEN (2003–2010) [143]	424	480	0.8 (0.3–2.1)
	San Francisco (1983–1985) [136]	5,046	20,882	1.6 (1.0–2.5)
Maternal phthalate exposure	HAVEN (2003–2010) [143]	424	480	2.2 (0.9–5.2)
Paternal phthalate exposure	HAVEN (2003–2010) [143]	424	480	1.7 (1.1–2.5)
Paternal alkylphenolic compound exposure	HAVEN (2003–2010) [143]	424	480	1.8 (1.1–3.0)
Drinking water contaminants				
Dichloroethylene	New Jersey (1985–1988) [21]	108	52,334	2.8 (1.3–5.9)*
Trichloroethylene	Milwaukee (1997–1999) [169]	245	3,780	6.2 (2.6–14.5)
Hazardous waste site	Dallas County (1979–1984) [104]	1,283	2,292	1.2 (1.1–1.4)

All exposures are maternal unless otherwise specified

Significant values are in bold

OR odds ratio, RR relative risk, CI confidence interval, HAVEN Heart Defects, Vascular Status, Genetic Factors and Nutrition Study, NCAS Northern Congenital Abnormality Survey

* 90 % CI

II. Annexe 2. Table 5

Pediatr Cardiol (2013) 34:1535–1555

1543

Table 5 Risk Factors Associated with Specific Congenital Heart Defects

Defect	Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
ASD	Advanced maternal age	NBDPS (1997–2007) [65]	468	8,169	2.5 (1.5–4.1)
		MACDP (1968–2005) [112]	398	1,301,143	1.4 (1.1–1.8)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	101	26,228	2.3 (1.8–2.8)
		NBDPS (1997–2003) [38]	419	4,689	8.5 (4.4–16.4)
	Febrile illness	BWIS (1981–1989) ^a [60]	187	3,572	1.7 (1.2–2.5)
	Infertility/ART	NBDPS (1997–2003) [126]	1,080	5,008	3.4 (1.8–6.2)
	Nulliparous	NBDPS (1997–2007) [56]	1,422	7,954	1.3 (1.1–1.5)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	621	5,673	1.3 (1.1–1.6)
	Obese	WCHARS (1992–2007) [103]	1,690	141,420	1.2 (1.0–1.4)
		NYSCMR (1993–2003) [113]	2,075	56,304	1.2 (1.1–1.4)
		SMBR (1992–2001) [31]	639	812,457	1.4 (1.1–1.7)
	Cephalosporins	NBDPS (1997–2003) [41]	1,225	4,941	1.9 (1.1–3.2)
	SSRI	ENN (1997–2006) [10]	3	53	5.7 (1.4–23.7)
		NBDPS (1997–2002) [5]	768	4,092	1.1 (0.6–1.9)
	Paroxetine	ENN (1997–2006) [10]	56	615	5.7 (1.4–23.5)
	SNRI (venlafaxine)	NBDPS (1997–2007) [120]	2,181	8,002	2.9 (1.2–6.9)
	Antihypertensives	NBDPS (1997–2003) [30]	1,137	4,796	2.4 (1.3–4.4)
	Alcohol use	FRCM (1982–1983) [149]	50	756	2.0 (1.1–3.4)
	Cigarette use	BWIS (1981–1989) [4]	186	3,435	1.4 (1.0–1.8)
		NBDPS (1997–2002) [105]	338	3,947	2.0 (1.5–2.6)
		SMBR (1983–1996) [81]	100	1,413,811	1.6 (1.0–2.6)
	Cocaine use (paternal)	BWIS (1981–1989) [60]	65,187	3,572	2.3 (1.3–4.2)
	Air pollution				
	Nitrogen dioxide	MACDP (1986–2003) [145]	379	715,500	1.6 (1.2–2.1)
	Particulate matter	Brisbane (1997–2004) [69]	127	2,860	1.1 (1.0–1.3)
		TBDR (1997–2000) [64]	977	3,431	1.3 (1.0–1.6)
	Sulphur dioxide	Brisbane (1997–2004) [69]	127	2,860	1.3 (1.0–1.7)
AVS	Pre-gestational diabetes mellitus	NBDPS (1997–2003) [38]	113	4,689	5.0 (1.1–22.9)
	Infertility/ART	NBDPS (1997–2005) [127]	243	6,500	2.6 (1.2–5.3)
	Prepregnancy WEIGHT				
AVSD	Overweight + obesity	NBDPS (1997–2004) [63]	154	5,673	0.8 (0.6–1.2)
	Obesity	NYSCMR (1993–2003) [113]	288	56,304	2.0 (1.4–2.9)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	11	26,228	2.2 (1.2–4.0)
AVSD		NBDPS (1997–2003) [38]	66	4,689	12.4 (3.7–41.5)
		BWIS (1981–1989) [101]	31	3,572	22.8 (7.4–70.5)
	Febrile illness	BWIS (1981–1989) [117]	244	3,435	1.7 (1.0–2.8)
		NBDPS (1997–2003) ^a [34]	98	4,760	2.3 (1.1–4.7)
	Injuries	NBDPS (1997–2005) [152]	166	6,328	2.4 (1.2–6.6)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	81	5,673	0.9 (0.6–1.5)
	Obesity	WCHARS (1992–2007) [103]	120	141,420	0.5 (0.3–1.1)
		NYSCMR (1993–2003) [113]	125	56,304	1.0 (0.6–1.7)
		SMBR (1992–2001) [31]	905	812,457	1.2 (1.0–1.5)
	Antibacterials	NBDPS (1997–2003) [41]	128	4,941	1.7 (1.1–2.6)
	Cough medications	BWIS (1981–1989) [60]	76	3,572	8.9 (2.6–30.6)
	Cigarette use	NBDPS (1997–2007) [118]	187	6,703	1.5 (1.1–2.1)
		BWIS (1981–1989) [4]	57	3,435	1.5 (1.0–2.3)
	Passive smoke exposure	NBDPS (1997–2007) [118]	187	6,703	1.5 (1.1–2.1)
	Cocaine use	BWIS (1981–1989) [60]	76	3,572	3.5 (1.1–11.4)
	Paternal polychlorinated compound exposure	HAVEN (2003–2010) [143]	44	480	4.2 (1.2–14.4)
DS-AVSD	Febrile illness	BWIS (1981–1989) [117]	187	3,435	1.9 (1.1–3.4)
	Ibuprofen	BWIS (1981–1989) ^a [60]	190	3,572	2.4 (1.1–4.2)

Table 5 continued

Defect	Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
CoA	Advanced maternal age	MACDP (1968–2005) [112]	256	1,301,143	1.5 (1.1–2.2)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	18	26,228	1.9 (1.2–3.1)
		NBDPS (1997–2003) [38]	196	4,689	2.1 (0.5–9.5)
	Epilepsy	BWIS (1981–1989) [60]	120	3,572	6.5 (1.8–23.0)
	Febrile illness	ABDCCS (1968–1980) [20]	60	3,029	2.7 (1.2–6.0)
	Infertility/ART	NBDPS (1997–2005) [127]	603	6,500	2.3 (1.4–3.8)
		BWIS (1981–1989) [60]	120	3,572	6.1 (2.1–18.2)
	Prepregnancy weight				
	Overweight	ABDRFSS (1993–1997) [158]	12	330	3.9 (1.1–13.8)
	Overweight + obesity	NBDPS (1997–2004) [63]	257	5,673	1.2 (0.9–1.5)
	Obesity	NYSCMR (1993–2003) [113]	433	56,304	1.3 (1.0–1.7)
		SMBR (1992–2001) [31]	117	812,457	1.5 (0.9–2.5)
	Influenza	ABDCCS (1968–1980) [20]	60	3,029	3.8 (1.6–8.8)
	Sulfonamides	NBDPS (1997–2003) [41]	431	4,941	2.7 (1.3–5.6)
	Bupropion	NBDPS (1997–2004) [6]	546	5,869	2.6 (1.0–6.9)
	SNRI (venlafaxine)	NBDPS (1997–2007) [120]	768	8,002	4.5 (1.4–12.5)
	Antihypertensive	NBDPS (1997–2003) [30]	406	4,796	3.0 (1.3–6.6)
	Mineral oil exposure	FRCM (1982–1983) [151]	50	756	5.9 (1.8–19.2)
	Organic solvents	BWIS (1981–1989) [60]	120	3,572	3.2 (1.3–7.9)
	Paternal alkylphenolic compound exposure	HAVEN (2003–2010) [143]	44	480	3.9 (1.2–12.7)
DORV	Pre-gestational diabetes mellitus	BWIS (1981–1989) [60]	27	3,572	12.3 (2.8–55.2)
	Febrile illness	BWIS (1981–1989) ^b [60]	27	3,572	2.8 (1.2–6.4)
	Prepregnancy weight				
	Severe obesity	NYSCMR (1993–2003) [113]	117	56,304	2.5 (1.1–5.8)
Ebstein's	Ibuprofen	BWIS (1981–1989) [60]	27	3,572	3.6 (1.1–12.2)
	Advanced maternal age	BWIS (1981–1989) [60]	44	3,572	2.6 (1.4–4.8)
	Advanced paternal age	BWIS (1981–1989) [60]	44	3,572	1.8 (1.0–3.5)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	56	5,673	1.8 (1.0–1.3)
	Benzodiazepines	BWIS (1981–1989) [60]	44	3,572	5.3 (1.5–18.5)
	Antihypertensives	NBDPS (1997–2003) [30]	65	4,796	11.4 (2.8–34.1)
	Marijuana use	BWIS (1981–1989) [60]	34	3,572	3.6 (1.6–8.5)
HLHS	Febrile illness	FRCM (1982–1983) ^b [147]	34	756	2.5 (1.2–5.4)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	5	26,228	0.7 (0.3–1.8)
		NBDPS (1997–2003) [38]	203	4,689	2.5 (0.7–8.8)
		BWIS (1981–1989) [60]	138	3,572	3.4 (1.0–11.5)
	Injuries	NBDPS (1997–2005) [152]	338	6,328	1.7 (1.0–3.0)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	268	5,673	1.3 (1.0–1.7)
	Obese	WCHARS (1992–2007) [103]	174	141,420	1.9 (1.1–1.3)
		NYSCMR (1993–2003) [113]	241	56,304	1.7 (1.2–2.5)
		SMBR (1992–2001) [31]	166	812,457	1.4 (0.9–2.2)
	Sulfonamides	NBDPS (1997–2003) [41]	6	4,941	3.2 (1.3–7.6)
	Metronidazole	NBDPS (1997–2003) [29]	176	4,581	2.3 (1.0–5.1)
	Organic solvents	BWIS (1981–1989) [60]	138	3,572	3.4 (1.6–6.9)
IAA	Injuries	NBDPS (1997–2005) [152]	22	6,328	5.9 (1.7–20.0)
	Aspirin	BWIS (1981–1989) [100]	46	3,572	2.1 (1.1–4.0)
PA	Pre-gestational diabetes mellitus	BWIS (1981–1989) [60]	45	3,572	7.2 (1.6–31.4)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	83	5,673	1.6 (1.0–2.5)
	Influenza	BWIS (1981–1989) [117]	39	3,435	2.7 (1.2–6.3)
	Injuries	NBDPS (1997–2005) [152]	140	6,328	2.9 (1.5–5.5)
	Air pollution				
	Particulate matter	TBDR (1997–2000) [64]	96	3,448	2.0 (1.1–3.6)

Table 5 continued

Defect	Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
PVS	Young maternal age	MACDP (1968–2005) [112]	262	1,301,143	0.7 (0.4–1.0)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	495	5,673	1.4 (1.1–1.7)
	Obesity	NYSCMR (1993–2003) [113]	957	56,304	1.3 (1.1–1.6)
	Febrile illness	NBDPS (1997–2003) ^a [34]	448	4,760	3.0 (1.2–7.6)
		BWIS (1981–1989) [60]	112	3,572	2.9 (1.5–5.4)
	Antihypertensives	NBDPS (1997–2003) [30]	534	4,796	2.6 (1.3–5.4)
	Naproxen	NBDPS (1997–2004) [71]	519	5,546	2.4 (1.3–4.5)
	Cigarette use	BWIS (1981–1989) [4]	205	3,435	1.4 (1.1–1.7)
		NBDPS (1997–2002) [105]	386	3,947	2.3 (1.1–4.8)
	Air pollution				
	Carbon monoxide	NCAS (1993–2003) [49]	259	14,256	2.7 (1.3–5.5)
	Organic solvents	BWIS (1981–1989) [60]	112	3,572	5.0 (1.3–8.7)
TA	Young maternal age	EUROCAT (2000–2004) [99]	167	1,177	2.6 (1.4–5.1)
		BWIS (1981–1989) [60]	30	3,572	2.8 (1.3–6.4)
	Febrile illness	BWIS (1981–1989) [117]	26	3,435	7.5 (2.6–21.8)
		ABDCCS (1968–1980) [20]	12	3,029	5.2 (1.3–20.2)
	Influenza	BWIS (1981–1989) [117]	26	3,435	6.0 (2.4–15.4)
	Injuries	NBDPS (1997–2005) [152]	2296	6,328	2.5 (1.1–5.7)
	Paternal cocaine use	BWIS (1981–1989) [60]	30	3,572	4.8 (1.6–14.0)
	Paternal marijuana use	BWIS (1981–1989) [60]	30	3,572	2.7 (1.3–5.8)
TAPVR	Young maternal age	NBDPS (1997–2007) [65]	190	8,169	2.3 (1.3–4.0)
	Pre-gestational diabetes mellitus	NBDPS (1997–2003) [38]	102	4,689	7.1 (2.0–25.4)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	119	5,673	1.5 (1.0–2.3)
TGA	Maternal pesticide exposure	BWIS (1981–1989) [60]	56	3,572	6.8 (1.5–31.5)
	Advanced maternal age	MACDP (1968–2005) [112]	177	1,301,143	1.7 (1.1–2.5)
		BWIS (1981–1989) [60]	214	3,572	1.7 (1.1–2.3)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	17	26,228	2.0 (1.2–3.2)
		NBDPS (1997–2003) [38]	254	4,689	3.3 (1.1–10.1)
	Influenza	BWIS (1981–1989) [60]	106	3,572	2.2 (1.2–4.1)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	314	5,673	1.0 (0.8–1.3)
	Obesity	WCHARS (1992–2007) [103]	177	141,420	1.1 (0.7–1.8)
		NYSCMR (1993–2003) [113]	331	56,304	0.9 (0.6–1.2)
		SMBR (1992–2001) [31]	164	812,457	1.5 (1.0–2.3)
	Stress	CBDMP (1987–1988) [25]	77	464	1.9 (1.1–3.1)
	Benzodiazepines	BWIS (1981–1989) [60]	189	3,572	4.1 (1.9–8.6)
	Folic acid	NBDPS (1997–2004) [39]	320	4,737	0.7 (0.5–1.0)
		BWIS (1981–1989) [134]	53	679	1.0 (0.5–2.2)
		ABDCCS (1968–1980) [17]	79	1,610	0.4 (0.2–0.9)
	Ibuprofen	BWIS (1981–1989) [60]	106	3,572	2.5 (1.2–4.9)
	Alcohol use	LA/SF/SC (1999–2004) [67]	106	425	1.9 (1.1–3.2)
	Cigarette use	SMBR (1983–1996) [81]	307	1,413,811	1.3 (1.0–1.71)
	Vitamin A	BWIS (1987–1989) [20]	47	679	2.4 (1.1–5.1)
	Organic solvents	BWIS (1981–1989) [60]	189	3,572	3.4 (1.5–7.5)
TOF	Advanced maternal age	Gill et al. [65]	750	8,169	2.2 (1.4–3.3)
	Pre-gestational diabetes mellitus	NBDPS (1997–2003) [38]	351	4,689	4.9 (2.2–11.0)
		BWIS (1987–1989) [101]	294	3,572	6.6 (3.2–13.3)
	Infertility/ART	BWIS (1981–1989) [60]	204	3,572	3.6 (1.9–6.9)
	Nulliparous	NBDPS (1997–2007) [56]	711	7,954	1.3 (1.1–1.6)
	Prepregnancy weight				
	Overweight	NYSCMR (1993–2003) [113]	487	56,304	1.3 (1.0–1.7)
	Overweight + obesity	NBDPS (1997–2004) [63]	447	5,673	1.2 (1.0–1.5)

Table 5 continued

Defect	Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
Truncus	Obese	WCHARS (1992–2007) [103]	377	141,420	1.1 (0.8–1.5)
		SMBR (1992–2001) [31]	223	812,457	1.1 (0.7–1.6)
	Stress	CBDMP (1987–1988) [25]	82	464	1.3 (0.8–2.1)
	Anti-asthmatic medication	SMBR (1995–2004) [83]	405	585,372	1.7 (1.1–2.6)
	Folic acid (multivitamin)	CBDMP (1987–1988) [137]	90	481	0.5 (0.3–1.0)
	Metronidazole	BWIS (1981–1989) [60]	341	3,572	6.0 (1.8–20.7)
	Air pollution				
	Carbon monoxide	TBDR (1997–2000) [64]	136	3,343	2.0 (1.3–3.3)
	Sulphur dioxide	England (1991–1999) [54]	146	759,993	1.4 (1.1–1.8)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	5	229	2.8 (1.2–6.9)
VSD		BWIS (1981–1989) [60]	38	3,572	13.2 (3.8–46.1)
	Prepregnancy weight				
	Obesity	MBR (1990–1994) [124]	NA	8,088	6.3 (1.6–24.8)
	Cigarette use	BWIS (1981–1989) [4]	20	3,435	1.9 (1.0–3.5)
	Hazardous waste site	TBDR (1996–2000) [91]	101	4,965	2.8 (1.2–6.5)
	Advanced maternal age	NBDPS (1997–2007) [65]	1,293	8,169	2.5 (1.8–3.5)
	Advanced paternal age	EHS (1995–1997) [13]	453	894	1.7 (1.3–2.3)
	Prepregnancy weight				
	Overweight + obesity	NBDPS (1997–2004) [63]	734	5,673	1.1 (0.9–1.3)
	Obesity	WCHARS (1992–2007) [103]	2,915	141,420	1.0 (0.9–1.1)
		NYSCMR (1993–2003) [113]	4,081	56,304	1.0 (0.9–1.1)
		SMBR (1992–2001) [31]	2,676	812,457	1.1 (1.0–1.3)
	Pre-gestational diabetes mellitus	EUROCAT (1990–2005) [62]	134	26,228	1.5 (1.2–1.8)
		NBDPS (1997–2003) [38]	571	4,689	2.9 (1.3–6.6)
		BWIS (1981–1989) [101]	563	3,572	3.1 (1.5–6.3)
	Febrile illness	NBDPS (1997–1998) [33]	168	692	0.9 (0.6–1.3)
		ABDCCS (1968–1980) [20]	209	3,029	1.8 (1.1–2.9)
	Infertility/ART	PaRCM (1987–2006) [146]	2,248	3,104	1.4 (1.1–1.8)
		NBDPS (1997–2005) [127]	1,333	6,500	1.6 (1.1–2.4)
	Hypertension	HCCSCA (1980–1996) [42]	1,661	38,151	1.3 (1.0–1.6)
	Influenza	HCCSCA (1980–1996) [42]	1,661	38,151	1.3 (1.0–1.6)
		ABDCCS (1968–1980) [20]	209	3,029	2.0 (1.1–3.6)
	Nulliparous	NBDPS (1997–2007) [56]	1,237	7,954	1.4 (1.2–1.6)
	SSRI	NBDPS (1997–2002) [5]	797	4,092	1.1 (0.6–1.9)
	Paroxetine	ENN (1997–2006) [10]	183	615	0.6 (0.1–4.6)
	Fluoxetine	FRCM (1996–2003) [106]	143	1,818	1.7 (1.1–2.4)
	Cough and cold medications	Northern England (1995) [14]	296	296	2.2 (1.4–3.5)
		BWIS (1981–1989) [60]	87	3,572	5.5 (1.6–18.7)
	Folic acid	HCCSCA (1980–1991) [42]	1,661	38,151	0.8 (0.8–0.9)
		HPS (1993–1996) [48]	5	19	0.3 (0.1–0.7)
		SEUBDS (1993–1996) [161]	186	521	1.2 (0.8–1.8)
	Metronidazole	HCCSCA (1980–1996) [42]	1,661	38,151	1.6 (1.2–2.3)
		BWIS (1981–1989) [60]	640	3,572	3.5 (1.1–10.8)
	Ibuprofen	BWIS (1981–1989) [60]	640	3,572	1.5 (1.0–2.3)
	Alcohol use	ABDCCS (1968–1980) [164]	122	3,029	3.1 (1.2–8.2)
		BWIS (1981–1989) [60]	73	3,572	1.8 (1.0–3.3)
	Alcohol use (paternal)	ABDCCS (1968–1980) [164]	122	3,029	4.0 (1.6–9.9)
	Cigarette use	NBDPS (1997–2002) [105]	584	3,947	1.3 (1.1–1.7)
	Cocaine use	BWIS (1981–1989) [60]	640	3,572	2.9 (1.7–4.8)
	Cocaine use (paternal)	BWIS (1981–1989) [60]	640	3,572	2.3 (1.6–3.3)
	Marijuana use	ABDCCS (1968–1980) [164]	122	3,029	2.4 (1.4–3.9)
	Marijuana use (paternal)	ABDCCS (1968–1980) [164]	122	3,029	2.2 (1.1–4.4)
		BWIS (1981–1989) [58]	491	3,549	1.4 (1.1–1.8)

Table 5 Risk Factors Associated with Specific Congenital Heart Defects

Defect	Exposure	Database	Cases (n)	Controls (n)	OR or RR (95 % CI)
	Air pollution				
	Carbon monoxide	NCAS (1993–2003) [49]	1,154	14,256	2.6 (1.9–3.7)
		TBDR (1997–2000) [64]	1,757	3,342	1.2 (1.0–1.4)
		CBDMP (1989–1993) [131]	260	9,106	1.6 (1.1–2.5)
	Nitrogen	MACDP (1986–2003) [145]	1,108	715,500	1.2 (1.0–1.4)
	Particulate matter	Brisbane (1997–2004) [69]	222	2,860	1.2 (1.0–1.3)
	Sulfur dioxide	TBDR (1997–2000) [64]	1,007	1,991	1.3 (1.1–1.6)
	Paternal phthalate exposure	HAVEN (2003–2010) [143]	113	480	2.8 (1.4–5.9)

All exposures are maternal unless otherwise specified

Significant values are in bold

OR odds ratio, RR relative risk, CI confidence interval, NA not available, ASD atrial septal defect, ART assisted reproductive techniques, SSRI selective serotonin receptor inhibitors, SNRI serotonin-norepinephrine reuptake inhibitor, AVS aortic valve stenosis, AVSD atrioventricular septal defect, DS-AVSD Down syndrome atrioventricular septal defect, CoA coarctation of the aorta, DORV double-outlet right ventricle, HLHS hypoplastic left heart syndrome, IAA interrupted aortic arch, PA pulmonary atresia, PVS pulmonary valve stenosis, TA tricuspid atresia, TAPVR total anomalous pulmonary venous return, TGA transposition of the great arteries, TOF tetralogy of Fallot, VSD ventricular septal defect, ABDCCS Atlanta Birth Defects Case-Control Study, ABDRFSS Atlanta Birth Defects Risk Factor Surveillance Study, BWIS Baltimore-Washington Infant Study, CBDMP California Birth Defects Monitoring Program, EHS Egyptian Health Service, ENN Eurocat Northern Netherlands Birth Registry, EUROCAT all 18 registries, FRGM Finnish Register of Congenital Malformations, HAVEN Heart Defects, Vascular Status, Genetic Factors and Nutrition Study, HCCSCA Hungarian Case-Control Surveillance of Congenital Abnormalities, HPS Hungarian Periconceptional Service, LA/SF/SC Los Angeles, San Francisco, and Santa Clara counties, MACDP Metropolitan Atlanta Congenital Defects Program, MBR Mainz (Germany) Birth Registry, NBDPS National Birth Defect Prevention Study, NCAS Northern Congenital Abnormality Survey, NYSCMR New York State Congenital Malformations Registry, PaRCM Paris Registry of Congenital Malformations, SEUBDS Sloane Epidemiology Unit Birth Defects Study, SMBR Swedish Medical Birth Registry, TBDR Texas Birth Defect Registry, WCHARS Washington Comprehensive Hospital Abstract Reporting System

^a Urinary tract infection

^b Upper respiratory infection

RÉFÉRENCES BIBLIOGRAPHIQUES

Adatia I, Coe JY, Harder J. 1987. Transposition of the great arteries in a neonate with Klinefelter's syndrome. *Pediatr. Cardiol.* 8: 285–286.

Ambros V. 2004. The functions of animal microRNAs. *Nature* 431: 350–355.

Anderson RH, Chaudhry B, Mohun TJ, Bamforth SD, Hoyland D, Phillips HM, Webb S, Moorman AFM, Brown NA, Henderson DJ. 2012. Normal and abnormal development of the intrapericardial arterial trunks in humans and mice. *Cardiovasc. Res.* 95: 108–115.

Anderson RH, Jacobs ML. 2008. The anatomy of tetralogy of Fallot with pulmonary stenosis. *Cardiol. Young* 18 Suppl 3: 12–21.

Apitz C, Webb GD, Redington AN. 2009. Tetralogy of Fallot. *The Lancet* 374: 1462–1471.

Arsenian S, Weinhold B, Oelgeschläger M, Rüther U, Nordheim A. 1998. Serum response factor is essential for mesoderm formation during mouse embryogenesis. *EMBO J.* 17: 6289–6299.

Ashworth M, Adnani M Al, Sebire NJ. 2006. Neonatal death due to transposition in association with premature closure of the oval foramen. *Cardiol. Young* 16: 586–589.

Bajolle F, Bonnet D. 2009. Génétique et cardiopathies congénitales. In: *Journées Parisiennes de Pédiatrie (JPP)*,.

Bajolle F, Zaffran S, Kelly RG, Hadchouel J, Bonnet D, Brown NA, Buckingham ME. 2006. Rotation of the myocardial wall of the outflow tract is implicated in the normal positioning of the great arteries. *Circ. Res.* 98: 421–428.

Bao ZZ, Bruneau BG, Seidman JG, Seidman CE, Cepko CL. 1999. Regulation of chamber-specific gene expression in the developing heart by *Irx4*. *Science* 283: 1161–1164.

Bartram U, Molin DG, Wisse LJ, Mohamad A, Sanford LP, Doetschman T, Speer CP, Poelmann RE, Gittenberger-de Groot AC. 2001. Double-outlet right ventricle and overriding tricuspid valve reflect disturbances of looping, myocardialization, endocardial cushion differentiation, and apoptosis in TGF-beta(2)-knockout mice. *Circulation* 103: 2745–2752.

Bianca S, Ettore G. 2001. Sex ratio imbalance in transposition of the great arteries and possible agricultural environmental risk factors. *Images Paediatr. Cardiol.* 3: 10–14.

Blue GM, Kirk EP, Sholler GF, Harvey RP, Winlaw DS. 2012. Congenital heart disease: current knowledge about causes and inheritance. *Med. J. Aust.* 197: 155–159.

Bodmer R. 1993. The gene *tinman* is required for specification of the heart and visceral muscles in *Drosophila*. *Dev. Camb. Engl.* 118: 719–729.

- Bom T van der, Zomer AC, Zwinderman AH, Meijboom FJ, Bouma BJ, Mulder BJM. 2011. The changing epidemiology of congenital heart disease. *Nat. Rev. Cardiol.* 8: 50–60.
- Boneva RS, Botto LD, Moore CA, Yang Q, Correa A, Erickson JD. 2001. Mortality associated with congenital heart defects in the United States: trends and racial disparities, 1979–1997. *Circulation* 103: 2376–2381.
- Botto LD, Correa A. 2003. Decreasing the burden of congenital heart anomalies: an epidemiologic evaluation of risk factors and survival. *Prog. Pediatr. Cardiol.* 18: 111–121.
- Breckpot J, Thienpont B, Arens Y, Tranchevent LC, Vermeesch JR, Moreau Y, Gewillig M, Devriendt K. 2011. Challenges of interpreting copy number variation in syndromic and non-syndromic congenital heart defects. *Cytogenet. Genome Res.* 135: 251–259.
- Breckpot J, Thienpont B, Peeters H, Ravel T de, Singer A, Rayyan M, Allegaert K, Vanhole C, Eyskens B, Vermeesch JR, Gewillig M, Devriendt K. 2010. Array Comparative Genomic Hybridization as a Diagnostic Tool for Syndromic Heart Defects. *J. Pediatr.* 156: 810–817.e4.
- Brown ML, Burkhardt HM, Connolly HM, Dearani JA, Cetta F, Li Z, Oliver WC, Warnes CA, Schaff HV. 2013. Coarctation of the aorta: lifelong surveillance is mandatory following surgical repair. *J. Am. Coll. Cardiol.* 62: 1020–1025.
- Bruneau BG, Bao ZZ, Fatkin D, Xavier-Neto J, Georgakopoulos D, Maguire CT, Berul CI, Kass DA, Kuroski-de Bold ML, Bold AJ de, Conner DA, Rosenthal N, et al. 2001a. Cardiomyopathy in *Irx4*-deficient mice is preceded by abnormal ventricular gene expression. *Mol. Cell. Biol.* 21: 1730–1736.
- Bruneau BG, Nemer G, Schmitt JP, Charron F, Robitaille L, Caron S, Conner DA, Gessler M, Nemer M, Seidman CE, Seidman JG. 2001b. A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor *Tbx5* in cardiogenesis and disease. *Cell* 106: 709–721.
- Buckingham M, Meilhac S, Zaffran S. 2005. Building the mammalian heart from two sources of myocardial cells. *Nat. Rev. Genet.* 6: 826–835.
- Burn J, Brennan P, Little J, Holloway S, Coffey R, Somerville J, Dennis NR, Allan L, Arnold R, Deanfield JE, Godman M, Houston A, et al. 1998. Recurrence risks in offspring of adults with major heart defects: results from first cohort of British collaborative study. *Lancet* 351: 311–316.
- Calcagni G, Digilio MC, Sarkozy A, Dallapiccola B, Marino B. 2007. Familial recurrence of congenital heart disease: an overview and review of the literature. *Eur. J. Pediatr.* 166: 111–116.
- Charron F, Nemer M. 1999. GATA transcription factors and cardiac development. *Semin. Cell Dev. Biol.* 10: 85–91.
- Chelly J, Marlhens F, Marec B Le, Jeanpierre M, Lambert M, Hamard G, Dutrillaux B, Kaplan JC. 1986. De novo DNA microdeletion in a girl with Turner syndrome and Duchenne muscular dystrophy. *Hum. Genet.* 74: 193–196.

- Chen J-F, Wang S, Wu Q, Cao D, Nguyen T, Chen Y, Wang D-Z. 2008. Myocardin marks the earliest cardiac gene expression and plays an important role in heart development. *Anat. Rec. Hoboken NJ* 2007 291: 1200–1211.
- Chen Z, Friedrich GA, Soriano P. 1994. Transcriptional enhancer factor 1 disruption by a retroviral gene trap leads to heart defects and embryonic lethality in mice. *Genes Dev.* 8: 2293–2301.
- Conway SJ, Henderson DJ, Kirby ML, Anderson RH, Copp AJ. 1997. Development of a lethal congenital heart defect in the splotch (Pax3) mutant mouse. *Cardiovasc. Res.* 36: 163–173.
- Cousins L. 1991. Etiology and prevention of congenital anomalies among infants of overt diabetic women. *Clin. Obstet. Gynecol.* 34: 481–493.
- Driel LMJW van, Smedts HPM, Helbing WA, Isaacs A, Lindemans J, Uitterlinden AG, Duijn CM van, Vries JHM de, Steegers EAP, Steegers-Theunissen RPM. 2008. Eight-fold increased risk for congenital heart defects in children carrying the nicotinamide N-methyltransferase polymorphism and exposed to medicines and low nicotinamide. *Eur. Heart J.* 29: 1424–1431.
- Durocher D, Charron F, Warren R, Schwartz RJ, Nemer M. 1997. The cardiac transcription factors Nkx2-5 and GATA-4 are mutual cofactors. *EMBO J* 16: 5687–96.
- Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, Grossfeld P, Fatkin D, Jones O, Hayes P, Feneley M, Harvey RP. 2003. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J. Am. Coll. Cardiol.* 41: 2072–2076.
- Erdogan F, Larsen LA, Zhang L, Tümer Z, Tommerup N, Chen W, Jacobsen JR, Schubert M, Jurkatis J, Tzschach A, Ropers H-H, Ullmann R. 2008. High frequency of submicroscopic genomic aberrations detected by tiling path array comparative genome hybridisation in patients with isolated congenital heart disease. *J. Med. Genet.* 45: 704–709.
- Fakhro KA, Choi M, Ware SM, Belmont JW, Towbin JA, Lifton RP, Khokha MK, Brueckner M. 2011. Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc. Natl. Acad. Sci.* 108: 2915–2920.
- Gajewski K, Zhang Q, Choi CY, Fossett N, Dang A, Kim YH, Kim Y, Schulz RA. 2001. Pannier is a transcriptional target and partner of Tinman during *Drosophila* cardiogenesis. *Dev. Biol.* 233: 425–436.
- Galan-Roosen AE De, Kuijpers JC, Meershoek AP, Velzen D van. 1998. Contribution of congenital malformations to perinatal mortality. A 10 years prospective regional study in The Netherlands. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 80: 55–61.
- Galvin KM, Donovan MJ, Lynch CA, Meyer RI, Paul RJ, Lorenz JN, Fairchild-Huntress V, Dixon KL, Dunmore JH, Gimbrone MA, Falb D, Huszar D. 2000. A role for smad6 in development and homeostasis of the cardiovascular system. *Nat. Genet.* 24: 171–174.

- Garg V, Kathiriyi IS, Barnes R, Schluterman MK, King IN, Butler CA, Rothrock CR, Eapen RS, Hirayama-Yamada K, Joo K, Matsuoka R, Cohen JC, et al. 2003. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424: 443–7.
- Gelb BD, Towbin JA, McCabe ER, Sujansky E. 1991. San Luis Valley recombinant chromosome 8 and tetralogy of Fallot: a review of chromosome 8 anomalies and congenital heart disease. *Am. J. Med. Genet.* 40: 471–476.
- Giglio S, Graw SL, Gimelli G, Pirola B, Varone P, Voullaire L, Lerzo F, Rossi E, Dellavecchia C, Bonaglia MC, Digilio MC, Giannotti A, et al. 2000. Deletion of a 5-cM region at chromosome 8p23 is associated with a spectrum of congenital heart defects. *Circulation* 102: 432–437.
- Gill HK, Splitt M, Sharland GK, Simpson JM. 2003. Patterns of recurrence of congenital heart disease: An analysis of 6,640 consecutive pregnancies evaluated by detailed fetal echocardiography. *J. Am. Coll. Cardiol.* 42: 923–929.
- Gillum RF. 1994. Epidemiology of congenital heart disease in the United States. *Am. Heart J.* 127: 919–927.
- Gioli-Pereira L, Pereira AC, Mesquita SM, Xavier-Neto J, Lopes AA, Krieger JE. 2010. NKX2.5 mutations in patients with non-syndromic congenital heart disease. *Int. J. Cardiol.* 138: 261–265.
- Goldmuntz E, Paluru P, Glessner J, Hakonarson H, Biegel JA, White PS, Gai X, Shaikh TH. 2011. Microdeletions and microduplications in patients with congenital heart disease and multiple congenital anomalies. *Congenit. Heart Dis.* 6: 592–602.
- Gottlieb PD, Pierce SA, Sims RJ, Yamagishi H, Weihe EK, Harriss JV, Maika SD, Kuziel WA, King HL, Olson EN, Nakagawa O, Srivastava D. 2002. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat. Genet.* 31: 25–32.
- Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH, Korn JM, McCarroll SA, Gorham JM, Gabriel S, et al. 2009. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet* 41: 931–935.
- Güçer S, Ince T, Kale G, Akçören Z, Ozkutlu S, Talim B, Çağlar M. 2005. Noncardiac malformations in congenital heart disease: a retrospective analysis of 305 pediatric autopsies. *Turk. J. Pediatr.* 47: 159–166.
- Gurvitz MZ, Inkelas M, Lee M, Stout K, Escarce J, Chang R-K. 2007. Changes in hospitalization patterns among patients with congenital heart disease during the transition from adolescence to adulthood. *J. Am. Coll. Cardiol.* 49: 875–882.
- Habets PEMH, Moorman AFM, Clout DEW, Roon MA van, Lingbeek M, Lohuizen M van, Campione M, Christoffels VM. 2002. Cooperative action of Tbx2 and Nkx2.5 inhibits ANF expression in the atrioventricular canal: implications for cardiac chamber formation. *Genes Dev.* 16: 1234–1246.

- Harvey RP. 1996. NK-2 homeobox genes and heart development. *Dev. Biol.* 178: 203–216.
- Hitz M-P, Lemieux-Perreault L-P, Marshall C, Feroz-Zada Y, Davies R, Yang SW, Lionel AC, D'Amours G, Lemyre E, Cullum R, Bigras J-L, Thibeault M, et al. 2012. Rare copy number variants contribute to congenital left-sided heart disease. *PLoS Genet.* 8: e1002903.
- Hoffman JIE, Kaplan S. 2002. The incidence of congenital heart disease. *J. Am. Coll. Cardiol.* 39: 1890–1900.
- Huehnergarth KV, Gurvitz M, Stout KK, Otto CM. 2008. Repaired tetralogy of Fallot in the adult: monitoring and management. *Heart Br. Card. Soc.* 94: 1663–1669.
- Hutson MR, Kirby ML. 2007. Model systems for the study of heart development and disease. Cardiac neural crest and conotruncal malformations. *Semin. Cell Dev. Biol.* 18: 101–110.
- JARCHO S. 1961. Coarctation of the aorta (Meckel, 1750; Paris, 1791). *Am. J. Cardiol.* 7: 844–852.
- Jenkins KJ, Correa A, Feinstein JA, Botto L, Britt AE, Daniels SR, Elixson M, Warnes CA, Webb CL, American Heart Association Council on Cardiovascular Disease in the Young. 2007. Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115: 2995–3014.
- Johnson MC, Hing A, Wood MK, Watson MS. 1997. Chromosome abnormalities in congenital heart disease. *Am. J. Med. Genet.* 70: 292–298.
- Karnebeek CD van, Hennekam RC. 1999. Associations between chromosomal anomalies and congenital heart defects: a database search. *Am. J. Med. Genet.* 84: 158–166.
- Kenny D, Hijazi ZM. 2011. Coarctation of the aorta: from fetal life to adulthood. *Cardiol. J.* 18: 487–495.
- Ko LJ, Engel JD. 1993. DNA-binding specificities of the GATA transcription factor family. *Mol. Cell. Biol.* 13: 4011–4022.
- Kume T. 2009. The cooperative roles of Foxc1 and Foxc2 in cardiovascular development. *Adv. Exp. Med. Biol.* 665: 63–77.
- Kume T, Deng K, Hogan BL. 2000. Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract. *Dev. Camb. Engl.* 127: 1387–1395.
- Kume T, Jiang H, Topczewska JM, Hogan BL. 2001. The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes Dev.* 15: 2470–2482.

- Kuo CT, Morrissey EE, Anandappa R, Sigrist K, Lu MM, Parmacek MS, Soudais C, Leiden JM. 1997. GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev.* 11: 1048–1060.
- Lalani SR, Shaw C, Wang X, Patel A, Patterson LW, Kolodziejska K, Szafranski P, Ou Z, Tian Q, Kang S-HL, Jinnah A, Ali S, et al. 2013. Rare DNA copy number variants in cardiovascular malformations with extracardiac abnormalities. *Eur. J. Hum. Genet. EJHG* 21: 173–181.
- Laverriere AC, MacNeill C, Mueller C, Poelmann RE, Burch JB, Evans T. 1994. GATA-4/5/6, a subfamily of three transcription factors transcribed in developing heart and gut. *J. Biol. Chem.* 269: 23177–23184.
- Lewin MB, McBride KL, Pignatelli R, Fernbach S, Combes A, Menesses A, Lam W, Bezold LI, Kaplan N, Towbin JA, Belmont JW. 2004. Echocardiographic evaluation of asymptomatic parental and sibling cardiovascular anomalies associated with congenital left ventricular outflow tract lesions. *Pediatrics* 114: 691–696.
- Lickert H, Takeuchi JK, Both I Von, Walls JR, McAuliffe F, Adamson SL, Henkelman RM, Wrana JL, Rossant J, Bruneau BG. 2004. Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature* 432: 107–112.
- Lièvre CS Le, Douarin NM Le. 1975. Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *J. Embryol. Exp. Morphol.* 34: 125–154.
- Lin Q, Schwarz J, Bucana C, Olson EN. 1997. Control of mouse cardiac morphogenesis and myogenesis by transcription factor MEF2C. *Science* 276: 1404–1407.
- Linde D van der, Konings EEM, Slager MA, Witsenburg M, Helbing WA, Takkenberg JJM, Roos-Hesselink JW. 2011. Birth Prevalence of Congenital Heart Disease Worldwide: A Systematic Review and Meta-Analysis. *J. Am. Coll. Cardiol.* 58: 2241–2247.
- Lindsay EA. 2001. Chromosomal microdeletions: dissecting del22q11 syndrome. *Nat. Rev. Genet.* 2: 858–868.
- Loffredo CA, Chokkalingam A, Sill AM, Boughman JA, Clark EB, Scheel J, Brenner JL. 2004. Prevalence of congenital cardiovascular malformations among relatives of infants with hypoplastic left heart, coarctation of the aorta, and d-transposition of the great arteries. *Am. J. Med. Genet. A.* 124A: 225–230.
- Lomonico MP, Bostrom MP, Moore GW, Hutchins GM. 1988. Arrested rotation of the outflow tract may explain tetralogy of Fallot and transposition of the great arteries. *Pediatr. Pathol. Affil. Int. Paediatr. Pathol. Assoc.* 8: 267–281.
- Lyons I, Parsons LM, Hartley L, Li R, Andrews JE, Robb L, Harvey RP. 1995. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. *Genes Dev.* 9: 1654–1666.
- Männer J. 2000. Cardiac looping in the chick embryo: a morphological review with special reference to terminological and biomechanical aspects of the looping process. *Anat. Rec.* 259: 248–262.

- Marino B, Digilio MC, Toscano A, Anaclerio S, Giannotti A, Feltri C, Ioris MA de, Angioni A, Dallapiccola B. 2001. Anatomic patterns of conotruncal defects associated with deletion 22q11. *Genet. Med. Off. J. Am. Coll. Med. Genet.* 3: 45–48.
- Mazen I, El-Ruby M, El-Bassyouni HT. 2010. Variable associations of Klinefelter syndrome in children. *J. Pediatr. Endocrinol. Metab. JPEM* 23: 985–989.
- McElhinney DB, Geiger E, Blinder J, Benson DW, Goldmuntz E. 2003. NKX2.5 mutations in patients with congenital heart disease. *J. Am. Coll. Cardiol.* 42: 1650–1655.
- Merscher S, Funke B, Epstein JA, Heyer J, Puech A, Lu MM, Xavier RJ, Demay MB, Russell RG, Factor S, Tokooya K, Jore BS, et al. 2001. TBX1 is responsible for cardiovascular defects in velo-cardio-facial/DiGeorge syndrome. *Cell* 104: 619–629.
- Moens CB, Stanton BR, Parada LF, Rossant J. 1993. Defects in heart and lung development in compound heterozygotes for two different targeted mutations at the N-myc locus. *Dev. Camb. Engl.* 119: 485–499.
- Molkentin JD, Lin Q, Duncan SA, Olson EN. 1997. Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev.* 11: 1061–1072.
- Molkentin JD, Lu JR, Antos CL, Markham B, Richardson J, Robbins J, Grant SR, Olson EN. 1998. A calcineurin-dependent transcriptional pathway for cardiac hypertrophy. *Cell* 93: 215–228.
- Morris CD, Menashe VD. 1991. 25-year mortality after surgical repair of congenital heart defect in childhood. A population-based cohort study. *JAMA J. Am. Med. Assoc.* 266: 3447–3452.
- Naderi S. 1979. Congenital abnormalities in newborns of consanguineous and nonconsanguineous parents. *Obstet. Gynecol.* 53: 195–199.
- Nemer G, Fadlalah F, Usta J, Nemer M, Dbaibo G, Obeid M, Bitar F. 2006. A novel mutation in the GATA4 gene in patients with Tetralogy of Fallot. *Hum. Mutat.* 27: 293–294.
- Nemer M. 2008. Genetic insights into normal and abnormal heart development. *Cardiovasc. Pathol. Off. J. Soc. Cardiovasc. Pathol.* 17: 48–54.
- Niu Z, Yu W, Zhang SX, Barron M, Belaguli NS, Schneider MD, Parmacek M, Nordheim A, Schwartz RJ. 2005. Conditional mutagenesis of the murine serum response factor gene blocks cardiogenesis and the transcription of downstream gene targets. *J. Biol. Chem.* 280: 32531–32538.
- Nora JJ. 1968. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. *Circulation* 38: 604–617.
- Orkin SH. 1992. GATA-binding transcription factors in hematopoietic cells. *Blood* 80: 575–581.

- Øyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PKA, Melbye M. 2009. Recurrence of congenital heart defects in families. *Circulation* 120: 295–301.
- Patel SS, Burns TL. 2013. Nongenetic risk factors and congenital heart defects. *Pediatr. Cardiol.* 34: 1535–1555.
- Pereira FA, Qiu Y, Zhou G, Tsai MJ, Tsai SY. 1999. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev.* 13: 1037–1049.
- Perry GH, Dominy NJ, Claw KG, Lee AS, Fiegler H, Redon R, Werner J, Villanea FA, Mountain JL, Misra R, Carter NP, Lee C, et al. 2007. Diet and the evolution of human amylase gene copy number variation. *Nat. Genet.* 39: 1256–1260.
- Petrini J, Damus K, Johnston RB Jr. 1997. An overview of infant mortality and birth defects in the United States. *Teratology* 56: 8–10.
- Pierpont ME, Basson CT, Benson DW Jr, Gelb BD, Giglia TM, Goldmuntz E, McGee G, Sable CA, Srivastava D, Webb CL, American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young. 2007. Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115: 3015–3038.
- Pillutla P, Shetty KD, Foster E. 2009. Mortality associated with adult congenital heart disease: Trends in the US population from 1979 to 2005. *Am. Heart J.* 158: 874–879.
- Pompa JL de la, Timmerman LA, Takimoto H, Yoshida H, Elia AJ, Samper E, Potter J, Wakeham A, Marengere L, Langille BL, Crabtree GR, Mak TW. 1998. Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* 392: 182–186.
- Praagh R Van. 2009. The first Stella van Praagh memorial lecture: the history and anatomy of tetralogy of Fallot. *Semin. Thorac. Cardiovasc. Surg. Pediatr. Card. Surg. Annu.* 19–38.
- Prall OWJ, Menon MK, Solloway MJ, Watanabe Y, Zaffran S, Bajolle F, Biben C, McBride JJ, Robertson BR, Chaulet H, Stennard FA, Wise N, et al. 2007. An Nkx2-5/Bmp2/Smad1 negative feedback loop controls heart progenitor specification and proliferation. *Cell* 128: 947–959.
- Prichard R. 1979. Selected items from the history of pathology. Transposition of the great vessels. *Am. J. Pathol.* 97: 562.
- Priest JR, Girirajan S, Vu TH, Olson A, Eichler EE, Portman MA. 2012. Rare copy number variants in isolated sporadic and syndromic atrioventricular septal defects. *Am. J. Med. Genet. A.* 158A: 1279–1284.
- Ramsdell AF. 2005. Left-right asymmetry and congenital cardiac defects: getting to the heart of the matter in vertebrate left-right axis determination. *Dev. Biol.* 288: 1–20.

- Rauch A, Devriendt K, Koch A, Rauch R, Gewillig M, Kraus C, Weyand M, Singer H, Reis A, Hofbeck M. 2004. Assessment of association between variants and haplotypes of the remaining TBX1 gene and manifestations of congenital heart defects in 22q11.2 deletion patients. *J. Med. Genet.* 41: e40.
- Reddy KS. 1999. A paternally inherited terminal deletion, del(8)(p23.1)pat, detected prenatally in an amniotic fluid sample: a review of deletion 8p23.1 cases. *Prenat. Diagn.* 19: 868–872.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, et al. 2006. Global variation in copy number in the human genome. *Nature* 444: 444–454.
- Reiter JF, Alexander J, Rodaway A, Yelon D, Patient R, Holder N, Stainier DY. 1999. Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev.* 13: 2983–2995.
- Rooij E van, Olson EN. 2007. MicroRNAs: powerful new regulators of heart disease and provocative therapeutic targets. *J. Clin. Invest.* 117: 2369–2376.
- Roos-Hesselink JW, Kerstjens-Frederikse WS, Meijboom FJ, Pieper PG. 2005. Inheritance of congenital heart disease. *Neth. Heart J.* 13: 88–91.
- Rosenthal A. 1972. Cardiovascular malformations in Klinefelter's syndrome: report of three cases. *J. Pediatr.* 80: 471–473.
- Rudolph AM, Heymann MA, Spitznas U. 1972. Hemodynamic considerations in the development of narrowing of the aorta. *Am. J. Cardiol.* 30: 514–525.
- Samánek M. 1994. Boy:girl ratio in children born with different forms of cardiac malformation: a population-based study. *Pediatr. Cardiol.* 15: 53–57.
- Samánek M, Slavík Z, Zborilová B, Hrobonová V, Vorísková M, Skovránek J. 1989. Prevalence, treatment, and outcome of heart disease in live-born children: a prospective analysis of 91,823 live-born children. *Pediatr. Cardiol.* 10: 205–211.
- Sanders SJ, Ercan-Sencicek AG, Hus V, Luo R, Murtha MT, Moreno-De-Luca D, Chu SH, Moreau MP, Gupta AR, Thomson SA, Mason CE, Bilguvar K, et al. 2011. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* 70: 863–885.
- Scambler PJ, Carey AH, Wyse RK, Roach S, Dumanski JP, Nordenskjöld M, Williamson R. 1991. Microdeletions within 22q11 associated with sporadic and familial DiGeorge syndrome. *Genomics* 10: 201–206.
- Schleich J-M, Abdulla T, Summers R, Houyel L. 2013. An overview of cardiac morphogenesis. *Arch. Cardiovasc. Dis.* 106: 612–623.
- Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, Maron BJ, Seidman CE, Seidman JG. 1998. Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* 281: 108–111.

- Seo S, Kume T. 2006. Forkhead transcription factors, *Foxc1* and *Foxc2*, are required for the morphogenesis of the cardiac outflow tract. *Dev. Biol.* 296: 421–436.
- Sepulveda JL, Belaguli N, Nigam V, Chen CY, Nemer M, Schwartz RJ. 1998. GATA-4 and *Nkx-2.5* coactivate *Nkx-2* DNA binding targets: role for regulating early cardiac gene expression. *Mol. Cell. Biol.* 18: 3405–3415.
- Siegenthaler JA, Choe Y, Patterson KP, Hsieh I, Li D, Jaminet S-C, Daneman R, Kume T, Huang EJ, Pleasure SJ. 2013. *Foxc1* is required by pericytes during fetal brain angiogenesis. *Biol. Open* 2: 647–659.
- Silversides CK, Lionel AC, Costain G, Merico D, Migita O, Liu B, Yuen T, Rickaby J, Thiruvahindrapuram B, Marshall CR, Scherer SW, Bassett AS. 2012. Rare copy number variations in adults with tetralogy of Fallot implicate novel risk gene pathways. *PLoS Genet.* 8: e1002843.
- Soemedi R, Topf A, Wilson IJ, Darlay R, Rahman T, Glen E, Hall D, Huang N, Bentham J, Bhattacharya S, Cosgrove C, Brook JD, et al. 2012a. Phenotype-specific effect of chromosome 1q21.1 rearrangements and *GJA5* duplications in 2436 congenital heart disease patients and 6760 controls. *Hum. Mol. Genet.* 21: 1513–1520.
- Soemedi R, Wilson IJ, Bentham J, Darlay R, Töpf A, Zelenika D, Cosgrove C, Setchfield K, Thornborough C, Granados-Riveron J, Blue GM, Breckpot J, et al. 2012b. Contribution of Global Rare Copy-Number Variants to the Risk of Sporadic Congenital Heart Disease. *Am. J. Hum. Genet.* 91: 489–501.
- Sommer RJ, Hijazi ZM, Rhodes JF. 2008. Pathophysiology of congenital heart disease in the adult: part III: Complex congenital heart disease. *Circulation* 117: 1340–1350.
- Sparkes R, Chernos J, Dicke F. 2005. Duplication of the 22q11.2 region associated with congenital cardiac disease. *Cardiol. Young* 15: 229–231.
- Srivastava D, Cserjesi P, Olson EN. 1995. A subclass of bHLH proteins required for cardiac morphogenesis. *Science* 270: 1995–1999.
- Srivastava D, Olson EN. 2000. A genetic blueprint for cardiac development. *Nature* 407: 221–226.
- Srivastava D, Thomas T, Lin Q, Kirby ML, Brown D, Olson EN. 1997. Regulation of cardiac mesodermal and neural crest development by the bHLH transcription factor, *dHAND*. *Nat. Genet.* 16: 154–160.
- Sucov HM, Dyson E, Gumeringer CL, Price J, Chien KR, Evans RM. 1994. RXR alpha mutant mice establish a genetic basis for vitamin A signaling in heart morphogenesis. *Genes Dev.* 8: 1007–1018.
- Talner NS, Berman MA. 1975. Postnatal development of obstruction in coarctation of the aorta: role of the ductus arteriosus. *Pediatrics* 56: 562–569.
- Temsah R, Nemer M. 2005. GATA factors and transcriptional regulation of cardiac natriuretic peptide genes. *Regul. Pept.* 128: 177–185.

- Tevosian SG, Deconinck AE, Tanaka M, Schinke M, Litovsky SH, Izumo S, Fujiwara Y, Orkin SH. 2000. FOG-2, a cofactor for GATA transcription factors, is essential for heart morphogenesis and development of coronary vessels from epicardium. *Cell* 101: 729–739.
- Théveniau-Ruissy M, Dandonneau M, Mesbah K, Ghez O, Mattei M-G, Miquerol L, Kelly RG. 2008. The del22q11.2 candidate gene *Tbx1* controls regional outflow tract identity and coronary artery patterning. *Circ. Res.* 103: 142–148.
- Thom T, Haase N, Rosamond W, Howard VJ, Rumsfeld J, Manolio T, Zheng Z-J, Flegal K, O'Donnell C, Kittner S, Lloyd-Jones D, Goff DC Jr, et al. 2006. Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation* 113: e85–151.
- van BIM, Mooij C, Kapusta L, Heil S, den HM, Blom HJ. 2008. Common 894G>T single nucleotide polymorphism in the gene coding for endothelial nitric oxide synthase (eNOS) and risk of congenital heart defects. *Clin. Chem. Lab. Med.* 46: 1369–1375.
- Veltman JA, Schoenmakers EFPM, Eussen BH, Janssen I, Merkx G, Cleef B van, Ravenswaaij CM van, Brunner HG, Smeets D, Kessel AG van. 2002. High-throughput analysis of subtelomeric chromosome rearrangements by use of array-based comparative genomic hybridization. *Am. J. Hum. Genet.* 70: 1269–1276.
- Verkleij-Hagoort AC, Driel LMJW van, Lindemans J, Isaacs A, Steegers EAP, Helbing WA, Uitterlinden AG, Steegers-Theunissen RPM. 2008. Genetic and lifestyle factors related to the periconception vitamin B12 status and congenital heart defects: A Dutch case-control study. *Mol. Genet. Metab.* 94: 112–119.
- Visootsak J, Aylstock M, Graham JM. 2001. Klinefelter syndrome and its variants: an update and review for the primary pediatrician. *Clin. Pediatr. (Phila.)* 40: 639–651.
- Waldo KL, Hutson MR, Stadt HA, Zdanowicz M, Zdanowicz J, Kirby ML. 2005a. Cardiac neural crest is necessary for normal addition of the myocardium to the arterial pole from the secondary heart field. *Dev. Biol.* 281: 66–77.
- Waldo KL, Hutson MR, Ward CC, Zdanowicz M, Stadt HA, Kumiski D, Abu-Issa R, Kirby ML. 2005b. Secondary heart field contributes myocardium and smooth muscle to the arterial pole of the developing heart. *Dev. Biol.* 281: 78–90.
- Wang D, Chang PS, Wang Z, Sutherland L, Richardson JA, Small E, Krieg PA, Olson EN. 2001. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* 105: 851–862.
- Ward C, Stadt H, Hutson M, Kirby ML. 2005. Ablation of the secondary heart field leads to tetralogy of Fallot and pulmonary atresia. *Dev. Biol.* 284: 72–83.
- Warnes CA, Williams RG, Bashore TM, Child JS, Connolly HM, Dearani JA, Nido P del, Fasules JW, Graham TP Jr, Hijazi ZM, Hunt SA, King ME, et al. 2008. ACC/AHA 2008 guidelines for the management of adults with congenital heart disease: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Develop Guidelines on the Management of Adults With Congenital Heart Disease). Developed in Collaboration With the American Society of

Echocardiography, Heart Rhythm Society, International Society for Adult Congenital Heart Disease, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. *J. Am. Coll. Cardiol.* 52: e143–263.

Weiss MJ, Orkin SH. 1995. GATA transcription factors: key regulators of hematopoiesis. *Exp. Hematol.* 23: 99–107.

Wessels MW, Willems PJ. 2010. Genetic factors in non-syndromic congenital heart malformations. *Clin. Genet.* 78: 103–123.

Wilm B, James RG, Schultheiss TM, Hogan BLM. 2004. The forkhead genes, *Foxc1* and *Foxc2*, regulate paraxial versus intermediate mesoderm cell fate. *Dev. Biol.* 271: 176–189.

Wren C, Birrell G, Hawthorne G. 2003. Cardiovascular malformations in infants of diabetic mothers. *Heart Br. Card. Soc.* 89: 1217–1220.

Yelbuz TM, Waldo KL, Kumiski DH, Stadt HA, Wolfe RR, Leatherbury L, Kirby ML. 2002. Shortened outflow tract leads to altered cardiac looping after neural crest ablation. *Circulation* 106: 504–510.

Yu W, Ballif BC, Kashork CD, Heilstedt HA, Howard LA, Cai W-W, White LD, Liu W, Beaudet AL, Bejjani BA, Shaw CA, Shaffer LG. 2003. Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Hum. Mol. Genet.* 12: 2145–2152.

Yuan S, Zaidi S, Brueckner M. 2013. Congenital heart disease: emerging themes linking genetics and development. *Curr. Opin. Genet. Dev.* 23: 352–359.

Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, Romano-Adesman A, Bjornson RD, Breitbart RE, Brown KK, Carrierio NJ, Cheung YH, et al. 2013. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 498: 220–223.

Zhao Y, Ransom JF, Li A, Vedantham V, Drehe M von, Muth AN, Tsuchihashi T, McManus MT, Schwartz RJ, Srivastava D. 2007. Dysregulation of cardiogenesis, cardiac conduction, and cell cycle in mice lacking miRNA-1-2. *Cell* 129: 303–317.

Zhao Y, Samal E, Srivastava D. 2005. Serum response factor regulates a muscle-specific microRNA that targets *Hand2* during cardiogenesis. *Nature* 436: 214–220.

Zhong TP, Rosenberg M, Mohideen MA, Weinstein B, Fishman MC. 2000. gridlock, an HLH gene required for assembly of the aorta in zebrafish. *Science* 287: 1820–1824.

Thèse de Doctorat

Marta SÁNCHEZ CASTRO

Génétique des malformations cardiaques congénitales : apport de la technique de puce à ADN génomique (Comparative Genomic Hybridisation ; aCGH)

Genetics of congenital heart defects: contribution of array comparative genomic hybridization (aCGH) method

Résumé

Notre projet a consisté à analyser par aCGH 316 patients présentant une malformation cardiaque congénitale (MCC) à type de transposition des gros vaisseaux, tétralogie de Fallot ou coarctation de l'aorte avec pour objectif la détection de microdélétions et microduplications génomiques et ainsi l'identification de nouveaux gènes contribuant à ces MCC. L'identification de nouveaux gènes permet une meilleure compréhension des mécanismes moléculaires et embryologiques à l'origine de ces malformations et un meilleur conseil génétique. Notre étude a conduit à l'identification de 21 microdélétions et 50 microduplications rares de novo ou héritées dont certaines comprennent des gènes candidats aux MCC. L'analyse bioinformatique des données a permis de montrer que de nombreuses microdélétions/microduplications comprennent des gènes possédant des sites de fixation de FOXC1. Ces microdélétions/microduplications pourraient donc altérer l'expression de certains gènes régulés par FOXC1 au cours du développement cardiaque. Par ailleurs, parmi les remaniements identifiés, nous avons détecté une délétion d'environ 1Mb en amont de *SOX9* chez des patients avec malformation cardiaque et syndrome de Pierre Robin. Nous avons montré que cette délétion comprend plusieurs activateurs cardiaques putatifs. Ces résultats suggèrent que la dérégulation de *SOX9* peut être responsable de MCC. Enfin, nous avons identifié une duplication partielle de la région 5' du gène *SEMA3D* donnant lieu à un ARNm tronqué de *SEMA3D*. Ces résultats suggèrent que la portion tronquée de *SEMA3D* pourrait perturber la migration des cellules de la crête neurale au cours du développement et ainsi contribuer aux MCC.

Mots clés

malformations cardiaques congénitales, transposition des gros vaisseaux, coarctation de l'aorte, tétralogie de Fallot, CGH array, CNV, FOXC1, *SOX9*, *SEMA3D*.

Abstract

Our project consists of the analysis by aCGH of a series of 316 patients with Congenital Heart Defects (CHD) such as transposition of the great arteries, tetralogy of Fallot and coarctation of the aorta with the aim to detect genomic microdeletions and microduplications and thus to identify new genes contributing to CHD. The identification of new genes contributes to improve the understanding of the molecular and embryological mechanisms underlying these defects and enable better genetic counseling. Our study led to the identification of rare de novo or inherited 21 microdeletions and 50 microduplications, some of them comprising candidate genes for CHD. Bioinformatic analysis of the data demonstrated that many microdeletions/microduplications include genes with FOXC1 transcription factor binding sites. These microdeletions/microduplications might hamper the expression of genes that are regulated by FOXC1 during heart development. Otherwise, among the rearrangements identified, we detected a ~1 Mb deletion upstream of *SOX9* in patients presenting with heart defects and Pierre Robin syndrome. We have shown that this deletion includes several putative cardiac enhancers. These results suggest that deregulation of *SOX9* might be responsible for CHD. Finally, we identified a duplication of the 5' half of *SEMA3D*, generating a truncated poly-A tailed mRNA of *SEMA3D*. These results suggest that truncated *SEMA3D* may have hampered the migration of cardiac neural crest cells during heart development, and thus contributed to CHD.

Key Words

congenital heart defects, transposition of the great arteries, coarctation of the aorta, tetralogy of Fallot, CGH array, CNV, FOXC1, *SOX9*, *SEMA3D*.