



HAL
open science

Etude du développement cardiaque et ses facteurs de transcription par l'utilisation des cellules souches pluripotentes induites

Robin Canac

► **To cite this version:**

Robin Canac. Etude du développement cardiaque et ses facteurs de transcription par l'utilisation des cellules souches pluripotentes induites. Médecine humaine et pathologie. Nantes Université, 2022. Français. NNT : 2022NANU1011 . tel-03872122

HAL Id: tel-03872122

<https://theses.hal.science/tel-03872122>

Submitted on 25 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE DE DOCTORAT DE

NANTES UNIVERSITE

ECOLE DOCTORALE N° 605

Biologie Santé

Spécialité : Biologie Cellulaire, Biologie du Développement

Par

Robin CANAC

Etude du développement cardiaque et ses facteurs de transcription par l'utilisation des cellules souches pluripotentes induites

Thèse présentée et soutenue à Nantes, le 21 Septembre 2022

Unité de recherche : l'institut du thorax (INSERM UMR 1087 / CNRS UMR 6291)

Rapporteurs.res avant soutenance :

Lucile MIQUEROL Directeure de recherche, CNRS, Institut de biologie du développement de Marseille
Stéphane ZAFFRAN Directeur de recherche, INSERM, Marseille Medical Genetics

Composition du Jury :

Président :	Flavien CHARPENTIER	Directeur de recherche, INSERM, l'institut du thorax, Nantes
Examineurs.rices :	Delphine POTIER	Chargée de recherche, CNRS, Centre d'Immunologie de Marseille-Luminy
	Thomas MOORE-MORRIS	Chargé de recherche, INSERM, Institut de Génomique Fonctionnelle, Montpellier
Dir. de thèse :	Patricia LEMARCHAND	Professeure des universités - Praticien hospitalier, Nantes Université, l'institut du thorax
Co-enc. de thèse :	Nathalie GABORIT Jeremie POSCHMANN	Chargée de recherche, CNRS, l'institut du thorax, Nantes Chargé de recherche, INSERM, Center for Research in Transplantation and Translational Immunology, Nantes

Remerciements

Je souhaite tout d'abord remercier Dr. Lucile Miquerol et Dr. Stéphane Zaffran d'avoir accepté le rôle de rapporteur.e et pour le temps passé à relire ce travail de thèse.

Je remercie également Dr. Delphine Potier et Dr. Thomas Moore-Morris d'avoir accepté de faire partie de mon jury de thèse en tant qu'examineur.rice .

J'adresse mes remerciements au Dr. Sigolène Meilhac et au Dr. Matthieu Giraud d'avoir fait partie de mon comité de suivi de thèse. Merci pour vos remarques pertinentes et vos encouragements concernant l'avancée de mon projet de thèse, et aussi de m'avoir fait bénéficier de vos expériences propres lors de nos discussions autour de mon projet professionnel.

Je remercie l'association Genavie et le fonds Marion Elizabeth Brancher pour m'avoir accordé les financements nécessaires à la poursuite de mon travail de thèse au cours d'une quatrième année. Je souhaite remercier plus particulièrement la fondatrice du fonds, Mme Florence Brancher, pour nos échanges intéressants lors de notre rencontre à Nantes.

Je remercie l'ensemble de l'équipe de direction de l'institut du thorax et notamment le directeur Richard Redon. Merci de m'avoir permis d'évoluer pendant plusieurs années au sein du laboratoire, qui offre un contexte de travail favorable unique.

J'adresse également mes sincères remerciements à Flavien Charpentier, chef de l'équipe 2a puis 2, pour avoir accepté de participer à mon jury de thèse et pour m'avoir intégré dans l'équipe depuis mon stage de L3. Merci pour les discussions scientifiques et remarques pertinentes lors de nos réunions d'équipe, mais aussi pour les discussions plus générales (mais tout aussi intéressantes) sur l'état actuel de la recherche et le quotidien de chef d'équipe.

Je remercie chaleureusement ma directrice de thèse, Patricia Lemarchand. Tout d'abord, merci de m'avoir donné l'opportunité d'intégrer ton équipe 7 (il y a maintenant un petit moment) et de m'avoir laissé évoluer jusqu'en thèse sur des sujets de recherche variés. Merci pour tes conseils et remarques scientifiques sur l'organisation et la communication, qui m'ont fait énormément progresser. Je ne manquerai d'ailleurs pas de « rétroplanifier » beaucoup plus à l'avenir ! Merci pour ta disponibilité et d'avoir pris le temps de relire entièrement mon manuscrit de thèse, malgré une dette de gâteaux qui s'accumule, et ton emploi du temps chargé. Merci de m'avoir forcé à prendre du recul dans les moments plus intenses de la thèse, j'ai énormément appris à tes côtés, tant scientifiquement, qu'humainement. Enfin, je souhaite souligner ta bienveillance envers tes étudiants, ton combat pour l'égalité dans le domaine de la recherche, ainsi que la rigueur et le franc-parler qui te caractérisent, car tous ces traits de ta personnalité forcent l'admiration et le respect.

Je souhaite également très sincèrement remercier mon encadrante de thèse, Nathalie Gaborit. Merci de m'avoir accompagné dans toutes les étapes de mon parcours au laboratoire et de m'avoir donné la possibilité d'avoir des projets scientifiquement très riches, me permettant d'aborder un large éventail d'approches et de techniques. Merci pour ton enthousiasme, ton énergie et ton humour au quotidien, qui sont de véritables atouts, et qui m'ont permis de faire avancer les projets dans un cadre de travail très favorable. Merci pour ta grande disponibilité, pour les discussions scientifiques enrichissantes et celles plus personnelles sur mon avenir en recherche, qui étaient tout aussi inspirantes. On aura quadrillé la France en congrès et formations, de Lille à Marseille en passant par Lyon, sans malheureusement pouvoir aller outre-Atlantique en congrès, dans ce pays dont tu m'as tant

fait les louanges ! Je retiendrais surtout tes grandes qualités de scientifique (habileté à diriger des recherches + médaillée du CNRS quand même !) et également tes qualités humaines et d'écoute. Vous êtes avec Patricia des exemples de travail dont je m'inspirerai pour avancer à l'avenir.

J'adresse aussi mes remerciements à Guillaume Lamirault. Merci pour les nombreux échanges scientifiques très enrichissants que l'on a pu avoir. J'ai beaucoup appris de tes réflexions sur les différentes approches que l'on peut donner à une question scientifique ou un projet. Je te remercie également pour ta jovialité et ton humour qui favorisent les relations plus détendues dans l'équipe. Je retiendrai aussi une discussion très inspirante que l'on a eue en congrès à Marseille, sur l'historique du laboratoire et ta vision de l'organisation de la recherche en France et en particulier à Nantes.

Plus globalement, je souhaite vous remercier Patricia, Nathalie et Guillaume en tant que groupe de recherche, pour nous avoir laissé avec Bastien, la possibilité de discuter ouvertement de la conduite des projets, et d'avoir pris en considération nos remarques et nos propositions.

Je remercie sincèrement mon encadrant de thèse, Jeremie Poschmann. Merci de m'avoir laissé la possibilité d'apprendre et de réaliser des techniques nouvelles au sein de ton équipe. Merci d'avoir été très disponibles quand j'en avais le plus besoin lors des optimisations infructueuses et de l'analyse des données. Merci aussi d'avoir été à l'écoute et ouvert aux propositions que j'ai pu faire. Merci également d'avoir trouvé le temps de relire mon manuscrit de thèse, malgré un emploi du temps bien rempli.

Je souhaite remercier Aurore Girardeau pour son soutien technique et culinaire sans faille. Merci de m'avoir initié au « magnifique » monde du western-blot, dont tu es la référence inégalée. Ton apport dans mon travail de thèse est considérable et je t'en remercie très sincèrement. Merci pour ton efficacité, ta rigueur et la qualité de ton travail qui ont permis d'avancer rapidement certaines parties du projet. La qualité de ton travail a également été retrouvée dans les nombreux gâteaux, biscuits, crêpes, ..., que tu as partagés (ou offert quand il y avait du chocolat ;)) avec nous ! Merci !

Je remercie également Virginie Forest pour m'avoir appris la culture des cellules souches et leur différenciation cardiaque, qui nous ont donné du fil à retordre pendant ces années !

Je veux aussi vivement remercier Cynthia Fourgeux pour m'avoir appris le travail concret qui se cache derrière le mot « ChIP-Seq ». Merci pour tous tes conseils et ton soutien depuis mon M2 jusqu'à maintenant. Merci d'avoir toujours été disponible et bienveillante envers moi, particulièrement dans les périodes plus compliquées.

Je souhaite remercier les anciens doctorants de l'équipe avec qui j'ai eu le plaisir de travailler et d'apprendre.

Zeina Al Sayed, merci pour tout ce tu m'as apporté, tant techniquement, qu'humainement. Tu es une des personnes très importantes que j'ai pu rencontrer au cours de ma thèse. Tu m'as passé le flambeau des IRX et j'espère avoir été à la hauteur de tes exigences ;) . Merci pour tous tes conseils et notamment ceux sur les meilleurs kebabs de Nantes !

Stephan de Waard, le seul l'unique et à jamais mon premier « thésard-encadrant » ! Merci pour toutes les expériences partagées depuis mon M1, jusqu'à encore il y a quelques jours. De la salle de culture, à celle de manipe, en passant par le RU, les restos, les bars ou même les stades de foot, on aura passé des moments forts et importants à mes yeux. Malgré une apparence nonchalante et ton second degré aiguisé, j'ai rencontré en toi une personne fiable, honnête et sur qui on peut compter. Merci pour tout ça.

Claire Castro, merci pour ta gentillesse caractéristique et ta bienveillance. J'ai eu l'honneur de partager ton bureau pendant une énoooooorme période de... 2 jours. ~~Merci~~ le COVID ! Tu as été une belle rencontre et un exemple de doctorante modèle, merci.

J'adresse également mes remerciements chaleureux au Dr. Bastien Cimarosti, mon acolyte et frère d'arme. On aura vécu beaucoup de moments variés, heureux ou plus difficiles, au cours du M2 et de la thèse. Tu as toujours été égal à toi-même, honnête, sincère et quelqu'un sur qui on peut compter, merci pour ça. Merci aussi pour ta maladresse, ton tact légendaire, nos « clash » de bureau et tes blagues dont toi seul a le secret. Ces années intenses ont forgé notre amitié et j'ai l'outrecuidance de croire qu'elle durera un bon moment ;) .

Je remercie tous les autres membres de l'équipe 2 avec qui j'ai vraiment apprécié travailler et échanger, Gildas Loussouarn, Isabelle Baró, Céline Marionneau, Mickaël Derangeon, Agnès Tessier Agnès Hivonnait, Béatrice Ollivier, Adrien Lesage, Marine Sarlandie et Floriane Bibault.

Je remercie Alban Gaignard, mon parrain de l'institut du thorax. Même si je ne t'ai pas vraiment sollicité, je savais que je pouvais compter sur toi.

Je souhaite remercier également Amandine Caillaud et Karim Si-Tayeb pour leur apport dans la partie édition génétique du projet. Merci pour votre gentillesse et de m'avoir appris et partagé vos protocoles et expériences de CRISPR-Cas9. Merci pour les nombreux conseils et votre disponibilité.

Je veux aussi remercier le groupe de Julien Barc avec Estelle Baron, Adrien Foucal et Enora Fremy pour leur aide précieuse dans la réalisation de l'étude multi-omique, en particulier pour l'ATAC-Seq. Merci pour tous vos conseils techniques et votre aide en bioinformatique.

Je souhaite remercier les membres de la plateforme génomique (GenoA), Audrey Donnart, Marine Cornec et Stéphanie Bonnaud sans qui toutes les analyses omiques nécessitant du séquençage haut débit n'auraient pas été possibles. Je remercie aussi le côté bioinformatique (BiRD) de la plateforme et plus particulièrement Eric Charpentier. Merci de m'avoir enseigné les bases de la bioinformatique, ainsi que pour ta disponibilité pour débloquer les problèmes sur le cluster.

Je remercie la plateforme de microscopie MicroPICell. Merci à Philippe Hulin de m'avoir formé à la microscopie confocale. Merci à Steven Nedellec pour ta formation et ton aide constante dans nos tentatives de mise en place du micropatterning des cardiomyocytes dérivés d'iPS. Merci à Perrine Paul-Gilloteaux pour ton aide pour l'analyse de vidéos avec Musclemotion, ainsi que pour l'analyse d'image de cardiomyocytes.

Je remercie aussi la plateforme iPS de Nantes et plus particulièrement Laurent David et Anne Gaignerie pour leurs nombreux conseils techniques et scientifiques dans la réalisation de ces projets, et notamment pour la publication des lignées iPS.

Je remercie en parallèle tous les membres du groupe de travail DevStem, sur les cellules souches à Nantes. Merci pour les différentes réunions, présentations et conseils pertinents pour la réalisation des projets associés aux cellules souches.

J'adresse mes remerciements au secrétariat de l'institut du thorax et plus spécifiquement à Corinne Mandin et Aurélie Garnier. Vous m'avez accompagné au laboratoire depuis mon arrivée en stage de L3 et jusqu'à présent. Je vous remercie pour votre efficacité ainsi que pour votre bonne humeur permanente.

Je remercie aussi Martine Le Cunff, Maud Chétiveaux, Nathalie Cressan, Davy Halaré et Guillaume Souchet pour leur travail de l'ombre mais qui nous permet de travailler dans un confort

certain. Merci pour votre sympathie et votre bonne humeur. Merci aussi Guillaume pour tous les bons plans goûter au labo ;)

Je remercie chaleureusement Eléonore Dijoux et Yoann Combot. Merci à tous les deux pour tous les très bons moments passés ensemble de la fac au labo, en passant par les bars, les restos et les boîtes. J'ai été très heureux de voir notre amitié grandir au fur et à mesure de la thèse, et je vous remercie aussi d'avoir été là dans les moments plus difficiles.

Je remercie également Antoine Persello. On aura passé de sacrés moments depuis tout ce temps ! Je nous revois encore dans le bureau avec Stephan quand tu préparais ta présentation pour les 48h avec des animations humoristiques un peu douteuses ;). Merci pour tes jeux de mots, ta grande gentillesse et ton honnêteté. Je garde encore l'espoir qu'on finira millionnaires un de ces jours !

Je remercie très sincèrement Manon Baudic et Michelle Geryk pour leur bonne humeur, leur sympathie et les moments partagés en salle de culture, dans les bars ou à l'étranger ;) Thank you so much for your humor and kindness, and Michelle for discussing naturally with me even with my poor english accent... ^^ I was lucky to meet both of you and thank you for being there when it was more complicated for me. Děkuji !

Je remercie Constance Delwarde, Milène Freineau, Enora Le Questel et Thomas Dupas, la « teem Koh-Lanta », merci de m'y avoir intégré et pour les bonnes soirées passées ensemble, parfois jusqu'au bout de la nuit (à Nantes et/ou Marseille ;). Je remercie également Ludivine Lopez pour sa gentillesse et sa facilité naturelle à organiser des sorties dans les bars. Merci aussi à Pierre Olchesqui, ton M2 pendant le COVID n'a pas été simple mais ton aide dans l'analyse vidéo a été importante.

Je remercie les autres membres de la salle iPS de l'institut, Diane Atticus, Lise Bray et Marine Delamare. On a partagé des bons moments, notamment pendant nos « mercredis gras », mais aussi des moments plus compliqués à lutter entre les pénuries de plastiques, les contaminations et la facilité déconcertante qu'ont ces petites cellules iPS à mettre nos plans à la poubelle.

J'adresse aussi mes Remerciements à ma meilleur(e) collègue de bureau Cyrielle Jajkiewicz. Merci pour ta sympathie et surtout bravo d'avoir réussi à supporter deux thésards de 4^{ème} année un peu déjantés et pas loin du craquage à certains moments. Je crois que tu es le parfait exemple de la phrase « qui aime bien châtie bien ». J'aurais aussi voulu te parler d'un logiciel super intéressant pour toi mais je ne me souviens plus lequel... ;)

Je remercie aussi les chercheurs titulaires « rookies ». Jérôme Montnach pour les discussions scientifiques et celle sur la vie de post-doc. Merci aussi pour l'analyse R des tracés de contraction. Simon Ducheix et Antoine Rimbart pour votre sympathie et votre humour. Merci aussi pour cette soirée/nuit mémorable des 48h 2021, le plongeur n'était pas loin ;)

Je remercie profondément Virginie Vignard pour sa bonne humeur, son humanité et sa bienveillance caractéristique. Merci d'avoir toujours été la même en toutes circonstances, et même quand tu étais ensevelie sous le travail. Je n'oublierai pas les longues discussions qu'on a pu avoir, merci d'avoir été là.

Amandine, je souhaite très sincèrement te remercier de m'avoir accompagné pendant ma thèse. Tu resteras la belle rencontre de ces dernières années. Je pense réellement que je n'en serai jamais arrivé au bout si tu n'avais pas été là et je t'en remercie. Merci pour les très beaux et bons moments, et je n'oublierai jamais la bienveillance et le soutien que tu m'as apportés. Je te souhaite le meilleur pour l'avenir.

Nala - 🐹, je te remercie toi aussi ma boule de poils, pour avoir supporté le rythme effréné de ma dernière année de thèse et pour m'avoir accompagné à toute heure du jour et de la nuit. Ton apport a été au moins équivalent à celui de Josette, la belle machine à café du bureau !

Merci à mes amis d'enfance, les copains sur qui on peut compter, la seconde famille. Merci d'avoir toujours été présent, même quand moi je l'étais le moins. Je n'aurai pas assez de ces quelques lignes pour vous dire à tous combien vous comptez pour moi. Je vous aime !

Jerem, mon frerot, merci pour tout ce que tu as fait pour moi. Pour tous les moments qu'on a passés ensemble et qui m'ont sorti la tête du labo pour me faire relativiser et retourner à l'essentiel. Merci aussi à Mélanie dit Bibiche, ta joie de vivre débordante est « trop chou » à voir. Et non je n'ai pas poussé de fonte à la salle ;) Mewen, merci à toi aussi pour m'avoir aidé à traverser les épreuves avec ton humour (et parfois malgré toi ^^) et ton sens de la fête ! Merci aussi à Negia de canaliser la bête pour que nos retrouvailles soient fameuses à chaque fois ! Merci à vous Pierro et Laura, pour tous les moments partagés, et il y en a eu des mémorables ces dernières années ! Merci à vous aussi Momo et Sarah (et également à Eva) pour tous les bons souvenirs et ceux qui viendront. Merci également à Flora, Alexis, Charline, Toto, Thib dit René, Dodo et Coco. Vous êtes toujours égaux à vous-même, même si on ne se voit pas souvent et c'est d'un soutien unique.

Enfin, je souhaite remercier ma famille sans qui je n'aurais pas pu me dépasser et en arriver là aujourd'hui.

Merci au sud, Tatie, Christophe, Julie et Axel et Gaël, merci d'avoir été présent et votre soutien pendant ma thèse. Les coupures de temps de quelques semaines à vos côtés l'été dans le sud m'ont été essentielles. Bien que la distance nous sépare, sachez que cela ne réduit en rien tout l'amour que je vous porte.

Merci à ma grand-mère. Ma petite Mamie Niquette, comment pourrais-je t'oublier, MERCI ! Merci pour ton soutien inébranlable et toutes tes attentions. Mes « petites bestioles » vont bien, mais je préfère les laisser à quelqu'un d'autre maintenant ;)

Merci à mes frères, Quentin et Gautier (je passerai vos surnoms très flatteurs), je ne vous le dis pas assez souvent mais je vous aime, merci d'être là, merci pour tout. Merci à toi aussi Chloé pour ton soutien et d'être présente pour mon frère.

Merci à mes parents. Merci pour tout votre soutien et amour, bien que ce monde « le labo » vous semble un peu flou, j'ai de la chance de vous avoir. Je vous aime.

A mes grands-pères Marcel et Roger,

Merci de m'avoir toujours encouragé, j'aurais été très heureux de vous montrer le résultat du travail acharné dont je vous parlais tant. Merci de m'avoir beaucoup appris, vous avez été et vous resterez pour moi des modèles de bonté, d'humour et de travail. J'espère vous rendre fiers de là où vous êtes.

Table des matières

Abréviations	10
Liste des figures	14
Liste des tableaux	15
Contexte de recherche	17
I. Introduction	19
I.1. Physiologie cardiaque	20
I.2. Développement cardiaque	23
I.2.A. Développement pré-cardiaque : de la pluripotence au mésoderme cardiogénique	23
I.2.B. Du croissant cardiaque à la formation du tube linéaire cardiaque	27
I.2.C. Formation des chambres cardiaques : focus sur l'hétérogénéité ventriculaire transmurale.....	29
I.2.D. Système de conduction électrique cardiaque et facteurs de transcription.....	32
I.2.E. Impact de la naissance sur les cardiomyocytes	35
I.3. Structuration de la chromatine et régulation transcriptionnelle	36
I.3.A. Organisation de la chromatine.....	36
I.3.B. Résumé des étapes de la transcription des ARNm	43
I.3.C. Lien entre la transcription et la conformation 3D de la chromatine	44
I.3.D. Eléments <i>cis</i> régulateurs	45
I.3.E. Marques épigénétiques.....	49
I.4. Les facteurs de transcription	54
I.4.A. Mécanisme d'action	54
I.4.B. Les facteurs de transcription Iroquois homeobox	63
I.4.C. Stratégies d'étude des facteurs de transcription cardiaques	71
II. Objectifs	83
III. Matériels et Méthodes	87
III.1. Méthodes spécifiques au projet 1	88
III.2. Méthodes spécifiques au projet 2	89
III.2.A. Etude des domaines d'interaction d'IRX5	90
III.2.B. Etude de la localisation d'IRX5 sur le génome	91
III.2.C. Génération du nouveau modèle d'étude iPS IRX5 ^{KO}	94
III.2.D. Approche multi-omique pour l'étude d'IRX5	95
III.2.E. Analyse des mécanismes de fibrose cardiaque.....	98

IV.	Résultats.....	101
IV.1.	Projet 1 : Réseau global des facteurs de transcription du développement cardiaque et famille Iroquois homeobox	103
IV.1.A.	Article 1 Deciphering transcriptional networks during human cardiac development.....	103
IV.2.	Projet 2 : Fonctions cardiaques spécifiques et redondantes d'IRX3 et d'IRX5.....	147
IV.2.A.	Contexte.....	147
IV.2.B.	Article 2 Generation of three human induced pluripotent stem cell lines with IRX5 knockout and knockin genetic editions using CRISPR-Cas9 system.....	149
IV.2.C.	Résultats complémentaires	157
IV.2.D.	Perspectives pour la poursuite du projet 2	171
V.	Discussion générale	175
V.1.A.	Intégration de la famille Iroquois dans le réseau global de facteurs de transcription du développement cardiaque	178
V.1.B.	Impact des facteurs de transcription IRX3 et IRX5 dans le gradient ventriculaire transmurale d'expression des gènes	182
V.1.C.	Implication d'IRX5 dans l'activité électrique cardiaque	185
V.1.D.	Implication des facteurs de transcription Iroquois dans les mécanismes de fibrose cardiaque.....	188
V.1.E.	Importance d'étudier les mécanismes épigénétiques dans le contexte du développement cardiaque	189
V.1.F.	Intérêts et perspectives de l'étude des réseaux de facteurs de transcription dans le développement cardiaque	192
V.1.G.	Conclusion.....	194
VI.	Annexes.....	197
VI.1.A.	Annexe 1	199
VI.1.B.	Annexe 2	217
VI.1.C.	Annexe 3	223
VI.1.D.	Annexe 4	231
VI.1.E.	Annexe 5	237
VII.	Bibliographie	243

Abréviations

2

2D : *Deux dimensions*

3

3C : *Chromosome Conformation Capture*

3D : *Trois dimensions*

4

4C : *Circular Chromosome Conformation Capture*

A

ADN : *Acide désoxyribonucléique*

ADNc : *ADN complémentaire*

Amhc1 : *Atrial Myosin Heavy Chain 1. Chaîne lourde de myosine atriale de poulet*

ANP : *Hormone peptide natriurétique auriculaire. Codée par le gène NPPA*

ARN : *Acide ribonucléique*

ARN pol II : *ARN polymérase II*

ATAC-Seq : *Assay for transposase-accessible chromatin using sequencing*

B

BAF60c : *Sous-unité cardiaque du complexe SWI/SNF. Codée par le gène SMARCD3*

BMP : *Bone Morphogenetic Protein ou protéines morphogénétiques osseuses. Molécule de signalisation*

Bmp10 : *Bone Morphogenetic Protein 10. Molécule de signalisation*

BMP4 : *Bone Morphogenetic Protein 4. Molécule de signalisation*

C

Cer1 : *Cerberus 1. Molécule de signalisation*

ChIP-Seq : *Chromatin Immunoprecipitation Sequencing*

c-Myc : *MYC Proto-Oncogene. Facteur de transcription*

Co-IP : *Co-immunoprécipitation*

CRISPR-Cas9 : *Clustered Regularly Interspaced Short Palindromic Repeats associated with CRISPR-associated protein 9. Technologie d'édition génétique*

CTCF : *CCCTC-Binding Factor. Facteur de transcription et insulateur protéique*

Cx30.2 : *Connexine 30.2 de jonctions gap. Codée par le gène GJC3*

Cx40 : *Connexine 40 de jonctions gap. Codée par le gène GJA5*

Cx43 : *Connexine 43 de jonctions gap. Codée par le gène GJA1*

Cx45 : *Connexine 45 de jonctions gap. Codée par le gène GJC1*

D

DH5 α : *Bactéries E. coli thermocompétentes*

Dkk-1 : *Dickkopf WNT Signaling Pathway Inhibitor 1. Molécule de signalisation*

DMSO : *Diméthylsulfoxyde*

DSIF : *DRB-sensitivity inducing factor. Impliqué dans la pause de l'ARN polymérase II*

E

EED : *Embryonic Ectoderm Development. Sous-unité du complexe répresseur Polycomb 2*

EGF : *Epidermal Growth Factor. Molécule de signalisation*

EOMES : *Eomesdermin. Facteur de transcription*

EZH1 : *Enhancer Of Zeste 1. Sous-unité du complexe répresseur Polycomb 2*

EZH2 : *Enhancer Of Zeste 2. Sous-unité du complexe répresseur Polycomb 2*

F

FGF2 : *Fibroblast Growth Factor 2. Molécule de signalisation*

FRET : *Fluorescence resonance energy transfer. Technique d'imagerie pour l'étude d'interactions protéiques*

G

GATA : Famille des facteurs de transcription GATA
GATA4 : GATA Binding Protein 4. Facteur de transcription
GATA5 : GATA Binding Protein 5. Facteur de transcription
GJA1 : Gap Junction Protein Alpha 1. Gène codant la connexine 43 (Cx43)
GJA5 : Gap Junction Protein Alpha 5. Gène codant la connexine 40 (Cx40)
GJC1 : Gène codant la connexine 45 (Cx45)
GJC3 : Gène codant la connexine 30.2 (Cx30.2)
GRN : Gene Regulatory Network. Réseau de régulation de gènes
GWAS : Genome Wide Association Study

H

H1 : protéine histone H1
H2A : protéine histone H2A
H2B : protéine histone H2B
H3 : protéine histone H3
H3K122Ac : Acétylation de la lysine 122 de l'histone H3. Marque épigénétique
H3K14Ac : Acétylation de la lysine 14 de l'histone H3. Marque épigénétique
H3K18Ac : Acétylation de la lysine 18 de l'histone H3. Marque épigénétique
H3K27Ac : Acétylation de la lysine 27 de l'histone H3. Marque épigénétique
H3K27Me3 : Triméthylation de la lysine 27 de l'histone H3. Marque épigénétique
H3K36Me3 : Triméthylation de la lysine 36 de l'histone H3. Marque épigénétique
H3K4Me1 : Monométhylation de la lysine 4 de l'histone H3. Marque épigénétique
H3K4Me3 : Triméthylation de la lysine 4 de l'histone H3. Marque épigénétique
H3K64Ac : Acétylation de la lysine 64 de l'histone H3. Marque épigénétique
H3K79Me2 : Diméthylation de la lysine 79 de l'histone H3. Marque épigénétique
H3K8Ac : Acétylation de la lysine 8 de l'histone H3. Marque épigénétique
H3K9Ac : Acétylation de la lysine 9 de l'histone H3. Marque épigénétique

H3K9Me2 : Diméthylation de la lysine 9 de l'histone H3. Marque épigénétique
H3K9Me3 : Triméthylation de la lysine 9 de l'histone H3. Marque épigénétique
H4 : protéine histone H4
HAND1 : Heart And Neural Crest Derivatives Expressed 1. Facteur de transcription
HAND2 : Heart And Neural Crest Derivatives Expressed 2. Facteur de transcription
HCN4 : Hyperpolarization Activated Cyclic Nucleotide Gated Potassium Channel 4. Canal potassique permettant l'automatisme dans le système de conduction cardiaque
HDAC : Histone désacétylase
HEK293 : Human embryonic kidney 293 cells. Lignée cellulaire
hESCs : Human Embryonic Stem Cells. Cellules souches embryonnaires humaines
Hi-C : high-throughput chromatin conformation capture. Technique d'analyse de la conformation de la chromatine
HP1 : Heterochromatin Protein 1. Protéine de l'hétérochromatine 1

I

ID2 : Inhibitor Of DNA Binding 2. Inhibiteur transcriptionnel
Inr : Human initiator core promoter element. Séquence ADN du core promoter favorisant l'initiation de la transcription
iPS : Induced pluripotent stem cell. Cellules souches pluripotentes induites
iPS-CFs : Induced pluripotent stem cell derived cardiac fibroblasts. Fibroblastes cardiaques dérivés d'iPS
iPS-CMs : Cardiomyocytes dérivés de cellules souches pluripotentes induites (iPS)
Iro-C : Locus Iroquois chez la Drosophile
IRX : Famille des facteurs de transcription IRX (Iroquois homeobox)
IRX1 : Iroquois Homeobox 1. Facteur de transcription
IRX2 : Iroquois Homeobox 2. Facteur de transcription
IRX3 : Iroquois Homeobox 3. Facteur de transcription

IRX4 : *Iroquois Homeobox 4. Facteur de transcription*
IRX5 : *Iroquois Homeobox 5. Facteur de transcription*
IRX5-HA : *Lignée iPS exprimant IRX5 couplé à un tag HA*
IRX5^{KO} : *Lignée iPS KO pour IRX5*
IRX6 : *Iroquois Homeobox 6. Facteur de transcription*
IrxA : *Cluster A de gènes Iroquois chez les mammifères*
IrxB : *Cluster B de gènes Iroquois chez les mammifères*
ISL1 : *ISL LIM Homeobox 1. Facteur de transcription*

K

KCHIP2 : *Potassium Channel Interacting Protein 2. Sous unité accessoire de canaux potassiques voltage-dépendants*
Kcnd2 : *Potassium Voltage-Gated Channel Subfamily D Member 2. Gène codant le canal Kv4.2*
KI : *Knock-in ou insertion génétique*
KLD : *Traitement Kinase, Ligase et enzyme de restriction DpnI*
Klf4 : *Kruppel Like Factor 4. Facteur de transcription*
KO : *Knock-out ou invalidation de l'expression d'un gène*
Kv4.2 : *Canal potassique voltage-dépendant*
Kv4.3 : *Canal potassique voltage-dépendant*

L

LADs : *Lamina-associated domain. Domaines associés à la lamina nucléaire*
Lefty 1 : *Left-Right Determination Factor 1. Molécule de signalisation*

M

MEF-2 : *Famille des facteurs de transcription Mef2 (Myocyte Enhancer Factor)*
MEIS2 : *Meis Homeobox 2. Facteur de transcription*
MESP1 : *Mesoderm Posterior BHLH Transcription Factor 1. Facteur de transcription*

N

Nav1.5 : *Canal sodique voltage-dépendant. Codé par le gène SCN5A*
NELF : *Negative Elongation Factor. Impliqué dans la pause de l'ARN polymerase II, : Negative Elongation Factor. Impliqué dans la pause de l'ARN polymerase II*
NKX : *Famille des facteurs de transcription NKX (NK2 Homeobox)*
NKX2-5 : *NK2 Homeobox 5. Facteur de transcription*
Notch : *Famille de récepteurs membranaire impliquée de la voie de signalisation Notch*
Notch1 : *Notch Receptor 1. Molécule de signalisation*
NPPA : *Natriuretic Peptide A. Gène codant le peptide natriurétique auriculaire (ANP)*

O

OCT4 : *Octamer-binding transcription factor 4. Facteur de transcription. Codé par le gène POU5F1*
OGA : *O-GlcNAcase. Enzyme*

P

PAM : *Protospacer Adjacent Motif. Motif reconnu par l'enzyme Cas9 du complexe CRISPR-Cas9*
PBS : *Phosphate Buffered Saline. Tampon phosphate salin*
PCA : *Principal Component Analysis. Analyse de réduction de dimensionnalité en composants principales*
PCR : *Polymerase Chain Reaction. Réaction en chaine de la polymérase*
PIC : *Pre-initiation complex. Complexe de pré-initiation de la transcription*
PITX2 : *Paired Like Homeodomain 2. Facteur de transcription*
POU5F1 : *POU Class 5 Homeobox 1. Gène codant le facteur de transcription OCT4*

Q

qPCR : *PCR quantitative*

R

- RBBP4 : *RB Binding Protein 4*. Sous-unité du complexe répresseur Polycomb 2
- RBBP7 : *RB Binding Protein 7*. Sous-unité du complexe répresseur Polycomb 2
- RNA-Seq : *RNA-Sequencing*. Technique d'analyse transcriptomique
- RPMI : *Roswell Park Memorial Institute medium*. Milieu de culture cellulaire
- RYR2 : *Ryanodine Receptor 2*. Gène codant le récepteur à la ryanodine de type 2 (cardiaque)

S

- SCN5A : *Sodium Voltage-Gated Channel Alpha Subunit 5*. Gène codant le canal sodique Nav1.5
- SERCA2 : *Pompe sarco/endoplasmique Ca²⁺ ATPase du réticulum*
- Shox2 : *Short Stature Homeobox 2*. Facteur de transcription
- SNPs : *Single nucleotide polymorphisms*. Polymorphismes d'un nucléotide
- SOX2 : *SRY (sex determining region Y)-box 2*. Facteur de transcription
- SUZ12 : *Suppressor Of Zeste 12*. Sous-unité du complexe répresseur Polycomb 2
- SVF : *Sérum de veau fœtal*
- SWI/SNF : *SWItch/Sucrose Non-Fermentable*. Complexe remodeleur de chromatine

T

- T3 : *Triiodothyronine*. Hormone thyroïdienne
- TADs : *Topologically Associating Domains*
- Tag HA : *Séquence de l'hémagglutinine grippale humaine*
- TALE : *Three Amino-acid Loop Extension*. Classe d'homéodomains protéiques
- TBX : *Famille des facteurs de transcription TBX (T-Box Transcription Factor)*

- TBX1 : *T-Box Transcription Factor 1*. Facteur de transcription
- TBX18 : *T-Box Transcription Factor 18*. Facteur de transcription
- TBX2 : *T-Box Transcription Factor 2*. Facteur de transcription
- TBX20 : *T-Box Transcription Factor 20*. Facteur de transcription
- TBX3 : *T-Box Transcription Factor 3*. Facteur de transcription
- TBX5 : *T-Box Transcription Factor 5*. Facteur de transcription
- TBXT : *T-Box Transcription Factor T*. Facteur de transcription. Alias : Brachyury ou T
- TFIIA : *Transcription factor IIA*. Responsable de l'initiation de la transcription
- TFIIB : *Transcription factor IIB*. Responsable de l'initiation de la transcription
- TFIID : *Transcription factor IID*. Responsable de l'initiation de la transcription
- TFIIE : *Transcription factor IIE*. Responsable de l'initiation de la transcription
- TFIIF : *Transcription factor IIF*. Responsable de l'initiation de la transcription
- TFIIH : *Transcription factor IIH*. Responsable de l'initiation de la transcription
- TGF- β : *Transforming growth factor β* , ou facteur de croissance transformant β
- Tn5 : *Transposase Tn5 ou Tagmentase*
- TNNI1 : *Troponin I1, Slow Skeletal Type*. Gène codant la troponine I1
- TNNI3 : *Troponin I3, Cardiac Type*. Gène codant la troponine I3
- TSS : *Transcription Start Site*. Point de départ de la transcription

V

- Vmhc1 : *Ventricular Myosin Heavy Chain 1*. Chaîne lourde de myosine ventriculaire du poulet

Liste des figures

Figure 1 : Couplage excitation-contraction dans un cardiomyocyte	20
Figure 2 : Potentiels d'action et courants ioniques cardiaques.....	21
Figure 3 : Développement pré-cardiaque humain.....	24
Figure 4 : Gastrulation et formation du mésoderme humain	25
Figure 5 : Développement cardiaque humain du croissant cardiaque au cœur adulte	28
Figure 6 : Modifications du métabolisme cardiomyocytaire associées à la naissance	36
Figure 7 : Structure de la chromatine	37
Figure 8 : Différentes échelles d'organisation de la chromatine	39
Figure 9 : Modèle de séparation de phase pour la formation de l'hétérochromatine	40
Figure 10 : Formation et structure des TADs	42
Figure 11 : Etapes du cycle de transcription des ARN messagers	43
Figure 12 : Modèles de fonctionnement des enhanceurs	47
Figure 13 : Impact de la charge des nucléosomes sur la structuration de la chromatine.....	51
Figure 14 : Différents modèles d'action des facteurs de transcription dans les interactions enhancer/promoteur	56
Figure 15 : Structuration d'un condensat transcriptionnel	57
Figure 16 : Notion de « dose d'activation » de gènes par l'action des facteurs de transcription	58
Figure 17 : Modèles d'accessibilité des facteurs de transcription à la chromatine	61
Figure 18 : Organisation génomique et structure des facteurs de transcription Iroquois homeobox	63
Figure 19 : Expression cardiaque des facteurs de transcription Iroquois.....	64
Figure 20 : Actions activatrices et inhibitrices d'Irx4 dans le cœur	66
Figure 21 : Rôles d'Irx3 dans l'établissement et la maturation du système de conduction ventriculaire	67
Figure 22 : Rôle d'Irx5 dans le maintien de l'hétérogénéité transmurale du potentiel d'action	69
Figure 23 : Patients atteints du syndrome d'Hamamy et présentant des mutations perte-de- fonction d'IRX5.....	70
Figure 24 : Différences entre les iPS-CMs et les cardiomyocytes adultes	75
Figure 25 : Génération de KO et KI par la technologie CRISPR-Cas9	76
Figure 26 : Techniques « omiques » pour l'étude de facteurs de transcription.....	79
Figure 27 : Vue d'ensemble et organisation des méthodes utilisées pour le projet 1.....	88
Figure 28 : Vue d'ensemble et organisation des méthodes utilisées pour le projet 2.....	89
Figure 29 : Récapitulatif des résultats du projet 1.....	177
Figure 30 : Récapitulatif des résultats du projet 2.....	178
Figure 31 : Proposition de modèle d'implication d'IRX5 dans le développement cardiaque	195

Liste des tableaux

Tableau 1 : Caractéristiques des principaux éléments régulateurs du génome	46
Tableau 2 : Descriptif des principales marques épigénétiques de l'histone H3	53
Tableau 3 : Primers de mutagenèse dirigée	90

Contexte de recherche

Cette thèse a été réalisée au sein de l'unité de recherche de l'institut du thorax (INSERM UMR1087 - CNRS UMR6291, directeur : Richard Redon), dont les axes de recherche principaux portent sur l'étude des maladies cardiovasculaires, métaboliques et respiratoires.

Ce projet a été conduit dans l'équipe II – Canaux ioniques et cardiopathies, dirigée par Flavien Charpentier et Michel De Waard, dans le sous-groupe composé de Patricia Lemarchand, Nathalie Gaborit et Guillaume Lamirault. Les objectifs d'étude de ce sous-groupe sont de mieux comprendre l'implication des facteurs de transcription dans le développement cardiaque et l'établissement de pathologies, plus particulièrement des troubles du rythme cardiaque, par l'utilisation de cellules souches pluripotentes induites (iPS) humaines.

Cette thèse a été dirigée par Patricia Lemarchand et co-encadrée par Nathalie Gaborit et Jeremie Poschmann.

Patricia Lemarchand, Professeure des universités - Praticien hospitalier (Nantes Université), et Nathalie Gaborit, Chargée de recherche (CNRS), ont participé au développement au sein de l'institut du thorax, du modèle iPS et particulièrement de leur différenciation cardiaque et de leur utilisation pour modéliser des pathologies cardiaques. En partant des travaux précédemment réalisés par Nathalie Gaborit sur les facteurs de transcription Iroquois, les projets de l'équipe se sont focalisés sur l'étude de ces facteurs de transcription dans l'établissement de troubles du rythme cardiaque, en utilisant le modèle iPS, qui a fait l'objet de plusieurs publications.

La poursuite de l'étude du rôle de ces facteurs de transcription dans le développement cardiaque pour la mise en place de ce projet de thèse, a ensuite nécessité l'utilisation d'analyses épigénétiques grâce à une collaboration avec Jeremie Poschmann, chargé de recherche (INSERM) au sein du *Center for Research in Transplantation and Translational Immunology* (UMR1064), expert dans cette thématique.



Introduction

I.1. Physiologie cardiaque

La contraction du cœur résulte d'un influx électrique généré dans le nœud sinusal, au niveau de l'oreillette droite, qui va enclencher la contraction des oreillettes lors de sa propagation jusqu'au nœud atrioventriculaire, à la jonction des oreillettes et des ventricules. La vitesse de conduction y est ralentie, avant le passage de l'influx électrique respectivement dans le faisceau de His, les branches droites et gauches du faisceau de His (le tout situé dans le septum interventriculaire) et le réseau ventriculaire de fibres de Purkinje, pour permettre une synchronisation entre la relaxation des oreillettes et la contraction des ventricules (Miquerol, 2018).

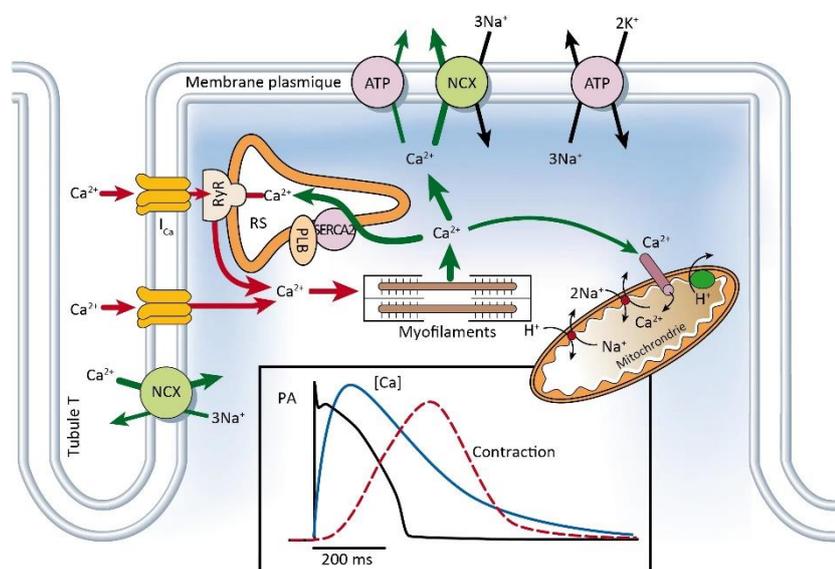


Figure 1 : Couplage excitation-contraction dans un cardiomyocyte

Un influx électrique ou potentiel d'action vient activer les canaux calciques voltage-dépendants dans les tubules T, provoquant une entrée de calcium dans la cellule. L'entrée de calcium extracellulaire va provoquer la libération du calcium du réticulum sarcoplasmique (RS) par l'intermédiaire du récepteur à la ryanodine de type 2 (RyR). Cette augmentation de calcium est nécessaire au coulisement des myofilaments, induisant la contraction. Le calcium est ensuite re-stocké dans le réticulum par la pompe SERCA2, elle-même régulée par le phospholamban (PLB), puis extrudé de la cellule par l'échangeur Na^+/Ca^{2+} (NCX) et d'autres échangeurs ATPasiques (ATP). Adapté de Bers, 2002.

La contraction cardiaque est sous le contrôle d'un couplage excitation-contraction s'établissant à l'échelle cellulaire (Figure 1) (Bers, 2002). Ce couplage débute lors du passage de l'influx électrique, correspondant à un potentiel d'action, généré par une association complexe d'entrée et sortie d'ions à travers la membrane plasmique des cardiomyocytes (Nerbonne and Kass, 2005). Le potentiel d'action est composé de cinq phases : (0) une dépolarisation rapide de la membrane par le courant sodique I_{Na} , suivie (1) d'une phase de repolarisation précoce liée à l'activation du courant potassique I_{to} , puis (2) une phase de

plateau associée principalement au courant calcique I_{CaL} , avant (3) une phase de repolarisation médiée par plusieurs courants potassiques (I_{Ks} , I_{Kr}) et (4) une phase diastolique, associée soit au courant I_f permettant l'automaticité de la génération des potentiels d'action dans les cardiomyocytes nodaux, soit au courant I_{K1} stabilisant le potentiel de repos dans les cardiomyocytes atriaux et ventriculaires. Par ailleurs, des différences de courants ioniques sont observées dans les différents types de cardiomyocytes (nodaux, atriaux et ventriculaires) ou en fonction de la région cardiaque, provoquant des différences de la forme de leur potentiel d'action spécifique (Figure 2) (Harvey and Grant, 2018; Lane and Tinker, 2017). Ainsi, dans l'épaisseur de la paroi des ventricules sont retrouvées des formes différentes de potentiels d'action, associées à la présence ou non du courant I_{to} (Figure 2). Cette hétérogénéité physiologique permet l'établissement du gradient ventriculaire transmural du potentiel d'action et permet une repolarisation appropriée des ventricules (Antzelevitch, 2005).

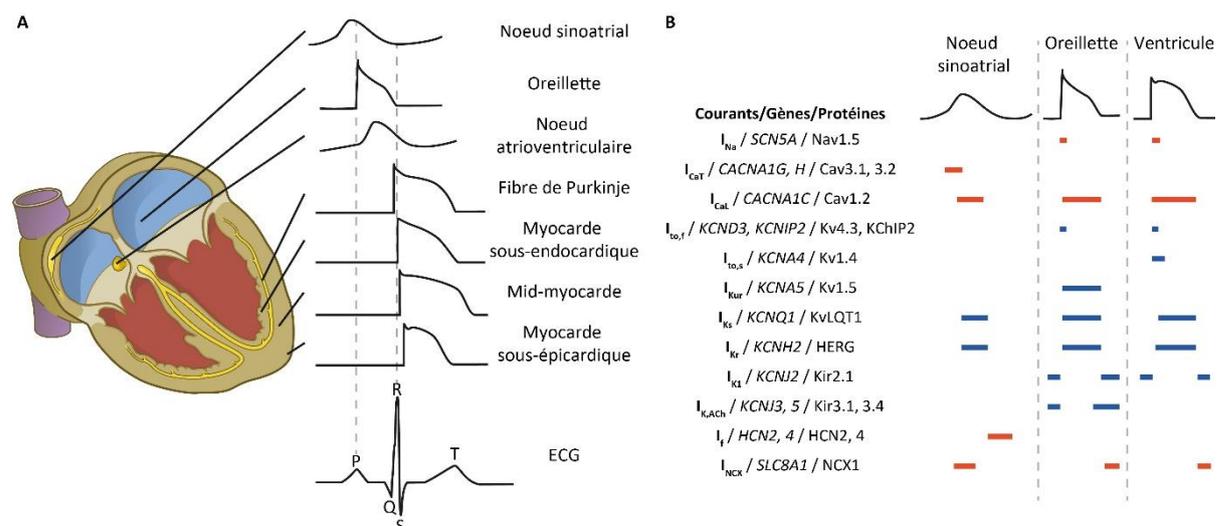


Figure 2 : Potentiels d'action et courants ioniques cardiaques

A. Illustration des différents potentiels d'action générés dans le cœur et leur intégration au sein du tracé typique d'électrocardiogramme (ECG). **B.** Différents courants ioniques sont impliqués dans les différences de potentiels d'action. Adapté de Bers, 2002, Harvey and Grant, 2018, Lane and Tinker, 2017, Nerbonne and Kass, 2005.

L'intégration de l'ensemble des différents potentiels d'action à chaque contraction peut être observée sur l'électrocardiogramme, où l'onde P correspond à la dépolarisation des oreillettes qui reflète leur contraction, où le complexe QRS correspond ensuite à une combinaison entre la repolarisation des oreillettes et la dépolarisation des ventricules et où l'onde T reflète la repolarisation des ventricules (Figure 2). La transmission du potentiel d'action entre les cardiomyocytes s'effectue à l'aide de connexines (Cx) formant des pores entre les cellules, dans des régions spécialisées des cardiomyocytes, les disques intercalaires,

et qui entraînent l'activation des canaux sodiques voltage-dépendants (I_{Na}) (Kleber and Saffitz, 2014). Cette dépolarisation membranaire au niveau des tubules transverses ou tubules T (invaginations de la membrane plasmique) des cardiomyocytes atriaux et ventriculaires induit une augmentation du calcium cytosolique par l'activation de canaux calciques voltage-dépendants de type L (I_{CaL}) et par le relargage de calcium du réticulum sarcoplasmique par l'intermédiaire du récepteur à la ryanodine (Bers, 2002). Le calcium se fixe ensuite à la troponine entraînant la contraction par coulissement des filaments épais de myosine et fins d'actine. Ensuite au cours de la phase de repolarisation membranaire, le calcium cytosolique est re-stocké dans réticulum sarcoplasmique par l'intermédiaire de la pompe SERCA2 (pompe sarco/endoplasmique Ca^{2+} ATPase du réticulum) et est en parallèle éjecté de la cellule principalement par l'échangeur sodium-calcium NCX (Eisner et al., 2017).

En plus de l'hétérogénéité des cardiomyocytes, qui ne correspondent qu'à 35% environ des cellules présentes dans le cœur (Tucker et al., 2020), d'autres types cellulaires sont présents. Cette complexité cellulaire a récemment été mise en évidence lors de la création d'atlas cellulaires du cœur, basés sur des données en cellule unique : parallèlement aux cardiomyocytes, on trouve principalement des fibroblastes, mais aussi des cellules musculaires lisses, des cellules endothéliales, des cellules immunitaires (lymphocytes, monocytes, macrophages, ...) et des neurones (Litviňuková et al., 2020). Les fibroblastes cardiaques (environ 32% des cellules du cœur ; Tucker et al., 2020) sont des cellules importantes dans la communication entre les différents types cellulaires présents dans le cœur et ils ont notamment des interactions proches avec les cardiomyocytes (Nicin et al., 2022). Par leur action régulatrice de la balance production/dégradation de matrice extracellulaire, ils sont des acteurs centraux des mécanismes de fibrose cardiaque (Hall et al., 2021). En effet, lors d'évènements conduisant à la mort de cardiomyocytes (ex : infarctus du myocarde), la capacité de régénération du cœur adulte étant très faible (Bergmann et al., 2009, 2015; Nguyen et al., 2021), des mécanismes de remodelage se mettent en place pour tenter de réparer la région touchée, notamment par la production de matrice extracellulaire non contractile, correspondant à la fibrose, afin de prévenir la rupture du myocarde (Frangogiannis, 2021). Le processus lésionnel provoque une transdifférenciation des fibroblastes quiescents en fibroblastes activés, ou myofibroblastes (α -SMA⁺), qui vont produire de la matrice extracellulaire (principalement du collagène) et bien qu'initialement

bénéfiques pour le maintien de la fonction cardiaque, les effets à long terme d'une fibrose excessive sont néfastes (Hall et al., 2021).

I.2. Développement cardiaque

Le cœur est le premier organe de l'embryon humain à se former pour permettre un apport de nutriments aux cellules et une captation des déchets, et ainsi de poursuivre le développement multicellulaire, puis multi-organe de l'embryon. Dans cette section seront présentées les différentes étapes clés de la formation d'un cœur fonctionnel et plus particulièrement des ventricules. Les données obtenues chez l'humain seront présentées préférentiellement et complétées, en cas d'absence, par des données obtenues dans d'autres organismes. Une attention particulière sera portée aux rôle des facteurs de transcription, qui représentent le cœur du sujet de ce travail.

I.2.A. Développement pré-cardiaque : de la pluripotence au mésoderme cardiogénique

I.2.A.a. Pluripotence

Le développement humain débute par la fécondation et la formation d'un zygote unicellulaire totipotent qui va rapidement se diviser de manière exponentielle (Condic, 2014; Zhai et al., 2022). Ces divisions successives vont conduire, 5 jours après la fécondation, à la formation d'un blastocyste primitif composé : (1) d'une enveloppe cellulaire externe, le trophoctoderme, qui formera le placenta, (2) d'une masse cellulaire interne, constituée de cellules souches pluripotentes, à partir desquelles sera formé le fœtus entier, et (3) d'une cavité remplie de liquide, le blastocœle (Figure 3) (Davidson et al., 2015). Les cellules pluripotentes de la masse cellulaire interne sont actuellement utilisées comme modèle de cellules souches embryonnaires humaines (hESCs ; Vazin and Freed, 2010). La masse cellulaire interne, sous l'influence de la signalisation TGF- β va ensuite, au 7^{ème} jour après la fécondation, se diviser (1) en épiblaste, correspondant aux cellules pluripotentes précurseurs de tous les tissus et (2) en endoderme primitif, qui formera les sacs vitellins (primaire et secondaire) permettant, entre autre, l'apport de nutriments à l'embryon (Ávila-González et al., 2021).

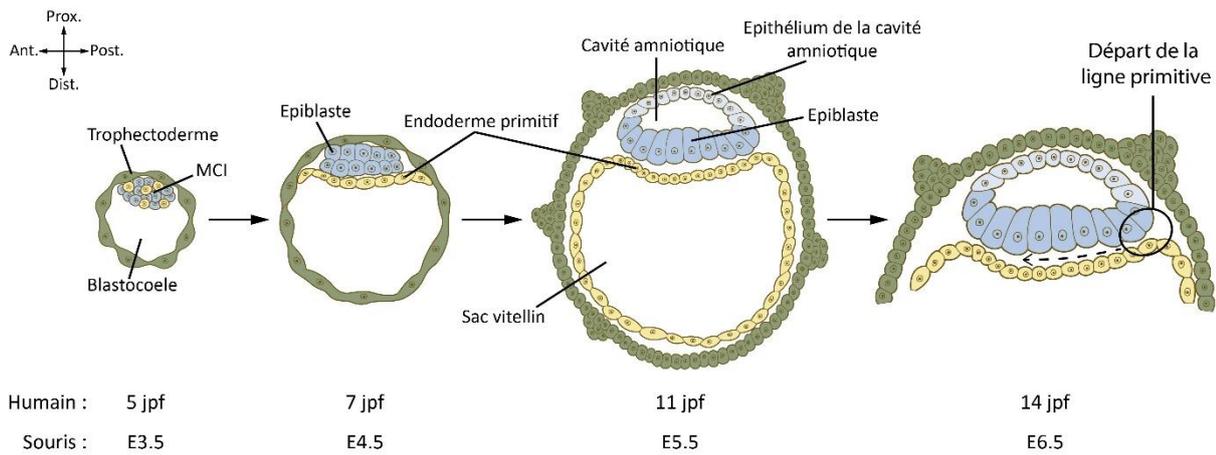


Figure 3 : Développement pré-cardiaque humain

Le nombre de jours post fécondation (jpf) chez l'humain et le stade embryonnaire (E) équivalent chez la souris sont indiqués pour chaque étape. Les axes proximal-distal (Prox./Dist.) et antéro-postérieur (Ant./Post.) sont également indiqués. MCI : Masse cellulaire interne. La ligne pointillée à 14 jpf indique le sens de déplacement de la ligne primitive lors de la gastrulation. Adapté de Chapitre 5 - Carlson, 2018.

Cet état de pluripotence qui est transitoire *in vivo* ou pouvant être stabilisé *in vitro*, est régi par trois facteurs de transcription spécifiques : OCT4¹, SOX2 et NANOG (Li and Belmonte, 2018). Ils fonctionnent ensemble pour réguler positivement leurs propres promoteurs (formant une boucle autorégulatrice interconnectée) ainsi que ceux d'autres gènes permettant le maintien du caractère souche des cellules. Ils inhibent en parallèle l'expression de gènes associés à la différenciation dans d'autres lignages cellulaires (Young, 2011). Ce groupe de facteurs de transcription fonctionne en coopération fonctionnelle avec d'autres facteurs de transcription tel que c-Myc. En effet, il est suggéré que la triade OCT4/SOX2/NANOG cible spécifiquement les gènes de la pluripotence en se fixant à leurs promoteurs pour recruter l'ARN polymérase, alors que c-Myc, sur ces mêmes promoteurs, agit en interrompant la pause de la polymérase pour déclencher la transcription (Yeo and Ng, 2013). La déstabilisation de cette triade par des facteurs extrinsèques va provoquer le passage progressif d'un état de pluripotence à celui de différenciation vers différents types cellulaires.

Entre le 7^{ème} et le 10^{ème} jour après la fécondation, l'embryon va s'implanter dans la muqueuse utérine, les cellules pluripotentes de l'épiblaste vont passer d'un état naïf à « amorcé » et vont différer sur plusieurs aspects comme l'expression spécifique de gènes marqueurs, ou la méthylation globale de l'ADN (Davidson et al., 2015). Les cellules amorcées sont plus enclines à la différenciation contrairement aux cellules naïves qui sont fortement

¹ : Par nomenclature, les symboles des gènes sont en écrits en italique et ceux des protéines en texte droit. Les symboles des gènes et protéines chez l'humain sont écrits en majuscule, et ceux des gènes et protéines chez la souris sont écrits en minuscule.

ancrées dans la pluripotence et, de manière concordante *in vitro* les hESCs amorcées sont capables de bien se différencier en cardiomyocytes battants contrairement aux hESCs naïves (Warrier et al., 2017). La différence entre les états naïf et amorcé est bien décrite chez la souris, mais il manque des modèles adéquats pour décrire dans le détail ces différences chez l'humain (Huang et al., 2014). C'est enfin environ au 11^{ème} jour après la fécondation, que l'épiblaste et l'endoderme primitif vont respectivement former la cavité amniotique et le sac vitellin, par cavitation (Figure 3) (Wamaitha and Niakan, 2018).

1.2.A.b. Gastrulation

Au 14^{ème} jour après la fécondation, l'embryon entre dans l'étape de gastrulation (Figure 4), qui va aboutir à la formation des trois feuilletts embryonnaires : (1) l'endoderme, qui donnera notamment les hépatocytes, les cellules pulmonaires et intestinales, (2) le mésoderme, d'où dérivent notamment les cardiomyocytes, les cellules musculaires lisses et endothéliales et, (3) l'ectoderme, qui est la source des progéniteurs neuronaux et des cellules de l'épiderme par exemple (Williams et al., 2012). La gastrulation débute par la formation de la ligne primitive (ou *primitive streak* en anglais), dans la région postérieure de l'épiblaste (Figure 4A) (Ghimire et al., 2021). La ligne primitive se crée par un détachement de cellules de l'épiblaste qui subissent une transition épithélio-mésenchymateuse (Thiery et al., 2009), et elle se propage linéairement vers la région antérieure de l'embryon. La ligne primitive sert de

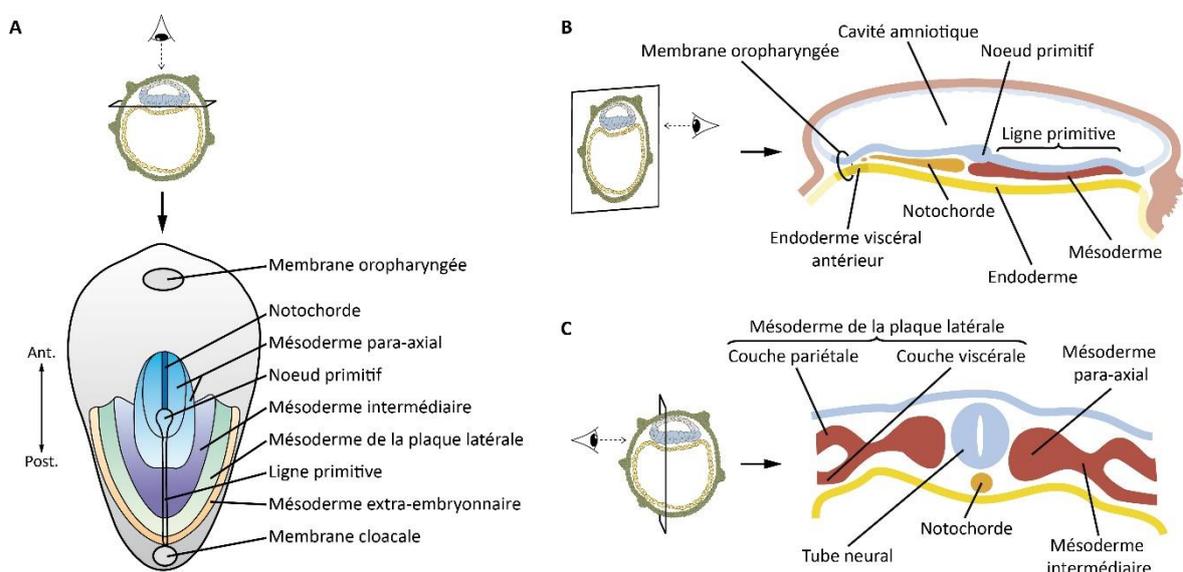


Figure 4 : Gastrulation et formation du mésoderme humain

A. Vue dorsale du disque embryonnaire de l'embryon en gastrulation. L'axe antéro-postérieur est indiqué (Ant./Post.). B. Section sagittale de l'embryon. C. Vue antérieure d'une coupe transversale du disque embryonnaire. Adapté de Chapitre 5 - Carlson, 2018 et Chapitre 5 - Sadler and Langman, 2012.

point d'entrée pour la colonisation, par des cellules de l'épiblaste, de l'invagination qui se crée entre l'épiblaste et l'endoderme primitif (Ghimire et al., 2021).

D'un point de vue moléculaire, la signalisation BMP semble centrale pour la gastrulation et la formation de la ligne primitive chez la souris, où la principale source de BMP est l'ectoderme extraembryonnaire, dérivé du trophoctoderme. Chez l'humain il n'existe pas d'équivalent clair à cette région embryonnaire, cependant chez le macaque l'épithélium de la cavité amniotique est connu pour agir comme un centre de régulation de l'activité BMP. Cette région pourrait donc être la source de la signalisation BMP chez l'humain (Ávila-González et al., 2021; Rossant and Tam, 2022). Des données chez la souris indiquent que cette signalisation est complétée par l'activation des voies de signalisation Wnt (Wnt3) et Nodal (membre de la famille du TGF- β) dans la région postérieure de l'épiblaste et provoque l'apparition et la formation de la ligne primitive. En parallèle, la stabilisation de l'endoderme viscéral antérieur correspondant à l'endoderme primitif situé sous la région antérieure de l'épiblaste (Figure 4B), va engendrer la production d'inhibiteurs de Nodal (Cer1 et Lefty 1) et Wnt (Dkk-1) empêchant la propagation de la ligne primitive dans la région antérieure de l'embryon et provoquant l'apparition du nœud primitif (Chapitre 5 - Carlson, 2018). L'association des signalisations Wnt/Nodal est connue chez la souris pour activer l'expression de facteurs de transcription au sein de la ligne primitive, comme TbxT dont l'activité est nécessaire au déplacement des futures cellules du mésoderme qui vont franchir la ligne primitive (Chapitre 5 - Carlson, 2018; Chapitre 3 - Schoenwolf et al., 2014), ou comme Eomes, marqueur pan-mésodermique (Probst and Arnold, 2017).

1.2.A.c. Mésoderme cardiogénique

Les cellules colonisant la région sous l'épiblaste vont, d'une part, remplacer les cellules de l'endoderme primitif pour former l'endoderme définitif et, d'autre part, se différencier en cellules du mésoderme pour former les mésodermes paraxial, intermédiaire et de la plaque latérale (Chapitre 5 - Sadler and Langman, 2012). Le mésoderme de la plaque latérale va se diviser en deux couches (Figure 4C) : (1) une couche pariétale ou somatique qui donnera les cellules du derme de la peau, des membres et des os et (2) une couche viscérale ou splanchnique permettant la formation des progéniteurs cardiaques notamment (Chapitre 6 - Carlson, 2018; Chapitre 13 - Sadler and Langman, 2012). Au 16^{ème} jour environ après la fécondation, les cellules de la couche viscérale du mésoderme de la plaque latérale exprimant le gène codant le facteur de transcription Mesp1 (Chiapparato et al., 2016), activé par Eomes

(De Bono et al., 2018; Van den Aemele et al., 2012), vont ensuite rapidement migrer vers la région antérieure de l'embryon. Cette migration s'effectue sous l'influence de la balance de signalisation : activation BMP/inhibition Wnt (Chapitre 12 - Schoenwolf et al., 2014) et former le mésoderme cardiogénique (Lescroart and Meilhac, 2012; Meilhac and Buckingham, 2018; Chapitre 12 - Schoenwolf et al., 2014).

I.2.B. Du croissant cardiaque à la formation du tube linéaire cardiaque

I.2.B.a. Champs cardiaques

Entre le 16^{ème} et le 19^{ème} jour après la fécondation, le mésoderme cardiogénique va progressivement, tout au long de la migration des progéniteurs vers la région antérieure de l'embryon, prendre une forme en arc appelée « croissant cardiaque » (Figure 5 – 19 jpf) (Chapitre 6 - Carlson, 2018). Le croissant cardiaque est une structure épithéliale correspondant à la première structure reconnaissable du cœur en développement (Meilhac et al., 2014). Il correspond également au premier champ cardiaque (ou *First Heart Field* en anglais) composé de progéniteurs exprimant des facteurs de transcription centraux dans le développement cardiaque, comme Nkx2-5, les facteurs de transcription de la famille Gata (Gata4/5), Tbx5 ou Hand1. Ces progéniteurs aboutiront quasi-exclusivement à la formation de cardiomyocytes dans le ventricule gauche et une partie des oreillettes (Chapitre 17 - Carlson, 2018; Meilhac and Buckingham, 2018; Schoenwolf et al., 2014).

Il existe également un second champ cardiaque (ou *Second Heart Field* en anglais), situé en position plus médiane et ventrale que le premier, qui est composé de progéniteurs cardiovasculaires multipotents exprimant spécifiquement les facteurs de transcription Isl1 et Tbx1. Ces progéniteurs aboutiront à la production (1) des cardiomyocytes dans le ventricule droit, ainsi que dans les oreillettes et la chambre de chasse (ou *outflow tract* en anglais) du cœur en développement, notamment sous l'impulsion de l'acide rétinoïque et des facteurs de transcription Hox (Bertrand et al., 2011; Stefanovic and Zaffran, 2017). Ils aboutiront également à la production (2) des cellules musculaires lisses et endocardiques (De Bono et al., 2018; Günthel et al., 2018; Kelly et al., 2014). Les cellules du second champ cardiaque vont aussi former le proépicaide (Figure 5 – 21 jpf), correspondant à un pool cellulaire qui engendrera notamment la production (1) de fibroblastes cardiaques, constituant la majorité des cellules interstitielles du cœur et (2) de la couche cellulaire externe du cœur, l'épicaide (Chapitre 17 - Carlson, 2018; Krenning et al., 2010; Moore-Morris et al., 2016). Une étude chez

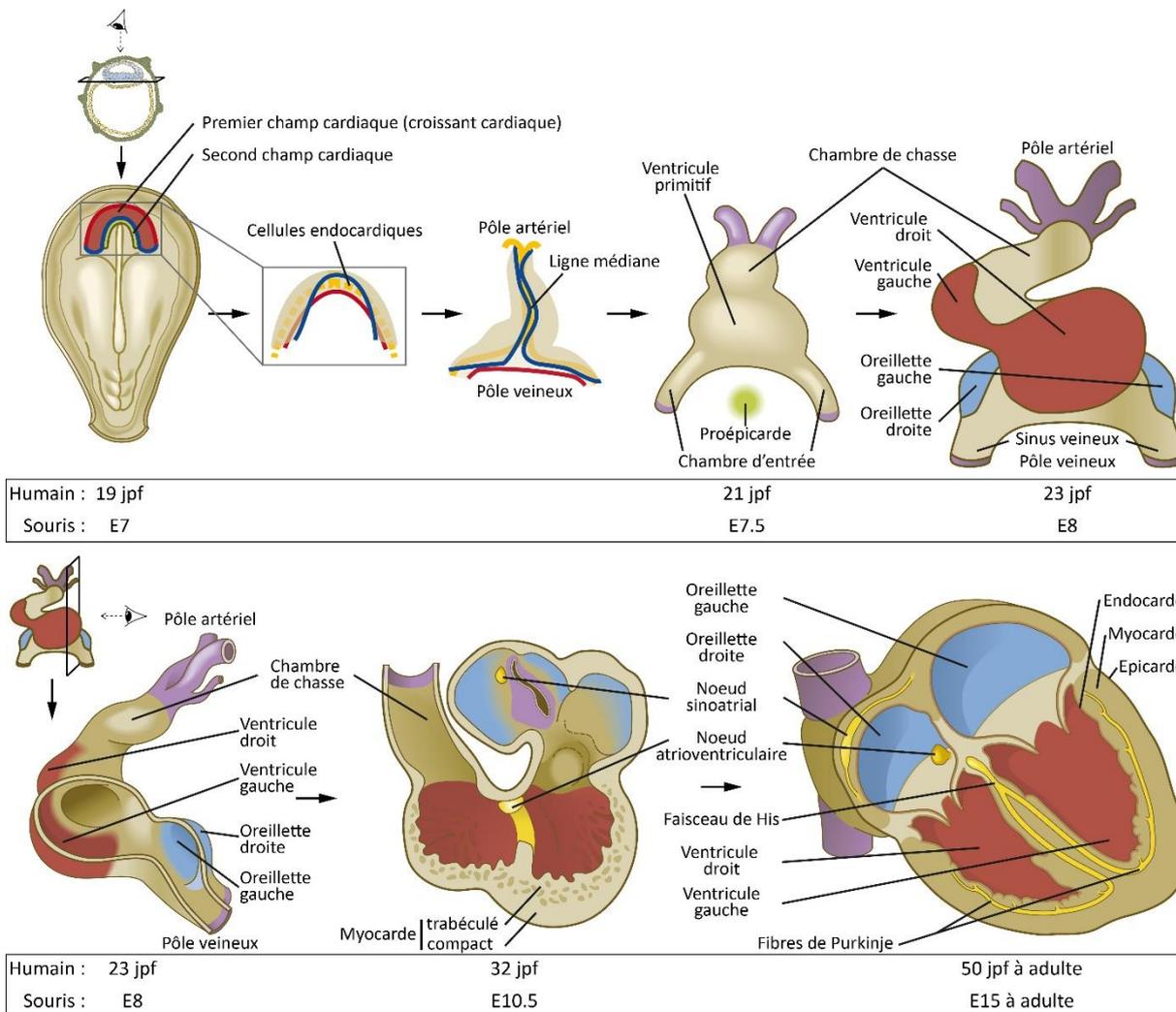


Figure 5 : Développement cardiaque humain du croissant cardiaque au cœur adulte

A partir de 21 jpf les zones beiges correspondent au myocarde primaire et les zones rouges et bleues au myocarde des chambres cardiaques. Le nombre de jours post fécondation (jpf) chez l'humain et le stade embryonnaire (E) équivalent chez la souris sont indiqués. Adapté de Buijendijk et al., 2020, Chapitre 5 et 17 - Carlson, 2018 et Sylva et al., 2014.

la souris indique que dès la gastrulation, différentes sous-populations $Mesp1^+$ temporellement et spatialement distinctes existent et correspondent à des progéniteurs cardiaques engagés dans les différentes régions du cœur (Lescroart et al., 2018). Par exemple, les cellules du premier champ cardiaque dérivent de progéniteurs $Mesp1^+$ « précoces » au début de la ligne primitive, alors les cellules du second champ cardiaque dérivent de progéniteurs $Mesp1^+$ plus tardifs (Lescroart et al., 2014).

1.2.B.b. Tube linéaire cardiaque

Entre le 19^{ème} et le 21^{ème} jour après la fécondation, les cardiomyocytes primitifs en différenciation provenant du premier champ cardiaque vont progressivement fusionner au niveau de la ligne médiane. Ce processus va former le tube linéaire cardiaque renfermant une

couche de cellules endocardiques et une couche de cellules myocardiques séparées par une région matricielle acellulaire, la gelée cardiaque (Günthel et al., 2018). La fusion progressive des cellules du premier champ cardiaque va provoquer la création du péricarde, une cavité remplie de plasma permettant la protection et la lubrification du cœur en développement (Chapitre 17 - Carlson, 2018; Jaworska-Wilczynska et al., 2016). Le tube linéaire cardiaque possède un pôle veineux du côté de la chambre d'entrée du tube (ou *inflow tract* en anglais) et un pôle artériel du côté de la chambre de chasse du tube (Figure 5 – 21 jpf). Le pôle veineux, proche des régions progénitrices des oreillettes, est la source d'un pool de cellules qui aboutira à la formation (1) de la région de confluence des veines au niveau de l'oreillette droite, appelée sinus veineux et (2) d'une partie des cellules composant le futur nœud sinoatrial, responsable de la génération de l'influx électrique cardiaque (van Wijk et al., 2018). La chambre de chasse va également être colonisée par une autre population cellulaire (non présentée ici), les cellules de la crête neurale (Yamagishi, 2021). Une fois le tube complètement formé, les premières contractions cardiaques vont apparaître du côté du pôle veineux (Buijtendijk et al., 2020).

I.2.C. Formation des chambres cardiaques : focus sur l'hétérogénéité ventriculaire transmurale

I.2.C.a. Formation de la boucle ou « looping » cardiaque

Entre le 22^{ème} et le 23^{ème} jour après la fécondation, le tube linéaire cardiaque va commencer à se densifier en cellules, non pas par division des cardiomyocytes primitifs du premier champ cardiaque, mais par l'incorporation de cellules du second champ cardiaque au niveau des deux extrémités du tube (Chapitre 17 - Carlson, 2018; Sylva et al., 2014). En parallèle, le tube cardiaque entre dans une étape de formation d'une boucle (ou *looping* en anglais) et va se déformer, provoquant la formation de la première structure cardiaque asymétrique visible, en passant d'une forme linéaire à une forme en C puis en S (Figure 5 – 23 jpf) (Buijtendijk et al., 2020; Dees and Baldwin, 2018). Le *looping* permet une réorganisation spatiale des différentes régions du tube et va induire le bon positionnement des futurs ventricules et oreillettes. Différents facteurs de transcription ont été décrits comme impliqués dans le *looping* cardiaque, comme *Pitx2*, qui est responsable de l'asymétrie droite-gauche pendant la formation du croissant cardiaque sous l'impulsion de la signalisation Nodal, et qui induit la rotation de la chambre de chasse pendant le *looping* (Campioni et al., 2018).

Cependant, l'implication de Pitx2 ne semble pas indispensable, car sa délétion spécifique dans les cardiomyocytes n'induit pas de défaut de *looping* (Desgrange et al., 2018). Les facteurs de transcription Nkx2-5, Mef-2, Hand1 et Hand2 ont aussi été impliqués dans le *looping*, car le *knock-out* (KO) de ces gènes bloque le développement cardiaque à ce stade (Chapitre 17 - Carlson, 2018). Les valves cardiaques vont commencer à se développer (non présenté ici), juste après le *looping*, entre le 31^{ème} et le 35^{ème} jour après la fécondation (O'Donnell and Yutzey, 2020).

1.2.C.b. Ballooning et apparition des chambres cardiaques

Une étape de gonflement (ou *ballooning* en anglais) va être enclenchée avant la fin du *looping* du tube cardiaque, par une nouvelle vague de prolifération et de différenciation des cardiomyocytes spécifiquement localisés dans le tube pour permettre la création des futures chambres cardiaques (Günthel et al., 2018; Sylva et al., 2014). On distingue ensuite deux types de myocarde : le myocarde primaire et le myocarde des chambres cardiaques. Le myocarde primaire est présent dans le tube cardiaque et son développement est guidé par le facteur de transcription Tbx2. Il est composé de cardiomyocytes dotés d'une activité électrique spontanée, avec de faibles capacités de contraction et de croissance, et d'une vitesse lente de conduction de l'influx électrique. Le myocarde des chambres cardiaques est retrouvé dans les oreillettes et les ventricules et son développement est guidé par le facteur de transcription Tbx5. Il est constitué de cardiomyocytes présentant une organisation sarcomérique bien établie, leur permettant de se contracter fortement, ainsi que de hautes vitesses de conduction de l'influx électrique et de hautes capacités prolifératives, mais sans activité électrique spontanée (Chapitre 17 - Carlson, 2018; Paige et al., 2015). GATA4, TBX5 et NKX2-5 constituent un groupe majeur de facteurs de transcription agissant en coopération dans la régulation de la formation des chambres cardiaques (Moorman and Christoffels, 2003; Paige et al., 2015; Tan and Lewandowski, 2020). Tbx20 est aussi fortement impliqué dans l'établissement des chambres cardiaques, d'une part, en interagissant avec ces facteurs de transcription centraux pour activer l'expression de gènes spécifiques du myocarde des chambres, comme *Nppa*, *Gja1/5* ou *Id2* (Moorman et al., 2010; Sylva et al., 2014). D'autre part, il inhibe l'activité de facteurs de transcription comme Tbx2 et Tbx3 dont la fonction est d'empêcher l'établissement des programmes d'expression génique des chambres cardiaques (Paige et al., 2015). La suppression de l'expression du facteur de transcription Mef2c, majoritairement activé par Isl1 dans le second champ cardiaque, induit des défauts sévères

de développement du ventricule droit et de la chambre de chasse, révélant son rôle fondamental dans la formation de ces structures (Dodou et al., 2004). Les facteurs de transcription Hand1 et Hand2 ont aussi été impliqués dans la mise en place des ventricules (George and Firulli, 2019). Plus spécifiquement, Hand1 est majoritairement exprimé par les cellules du premier champ cardiaque et son activité est principalement concentrée dans le ventricule gauche, alors que Hand2 est majoritairement exprimé dans le second champ cardiaque sous l'impulsion de l'activation par Isl1 et l'invalidation de son expression va induire des défauts majeurs du développement du ventricule droit (Chapitre 17 - Carlson, 2018; Ren et al., 2021).

Le facteur de transcription Iroquois homeobox *Ir4* est également impliqué dans la spécification des chambres et plus particulièrement des ventricules. Par son expression qui y est exclusive, il permet leur spécification en inhibant les gènes impliqués dans la mise en place des oreillettes (Bruneau et al., 2000, 2001).

Entre le 27^{ème} et le 37^{ème} jour après la fécondation, les deux ventricules vont être progressivement de plus en plus délimités par l'apparition du septum interventriculaire, correspondant à une protubérance du myocarde ventriculaire qui commence de l'apex du ventricule primitif et s'étend dans la cavité ventriculaire (Figure 5 – 32 jpf) (Dees and Baldwin, 2018; Chapitre 13 - Sadler and Langman, 2012). En parallèle, la septation interatriale s'effectue pour séparer les deux oreillettes (Chapitre 17 - Carlson, 2018).

1.2.C.c. Développement myocardique des ventricules

La croissance des chambres cardiaques en particulier dans le ventricule, implique l'expansion de sa surface, mais aussi en parallèle, de l'épaississement de la paroi myocardique ou paroi du ventricule. Sur la partie interne de la paroi ventriculaire, appelée myocarde trabéculé, se forment entre le 23^{ème} et le 26^{ème} jour après la fécondation, des trabéculations (petites excroissances) vers la lumière des ventricules en formation. Ces trabéculations deviennent ensuite de plus grandes crêtes reconnaissables (trabécules) entre le 28^{ème} et le 32^{ème} jour après la fécondation (Meilhac, 2018). Les trabécules sont constituées d'endocarde, de gelée cardiaque acellulaire et de myocarde, respectivement de l'intérieur vers l'extérieur de la paroi du ventricule. Il y a ensuite, dans la partie plus externe de la paroi ventriculaire, le myocarde compact recouvert d'une couche de cellules épicrodiques (Figure 5 – 32 jpf).

L'origine cellulaire commune entre le myocarde trabéculé et compact suppose qu'il existe des signaux locaux responsables de leurs spécificités et de leur différenciation

ultérieure (Moorman et al., 2010). L'interaction endocarde/myocarde permet l'apparition des trabécules. La voie de signalisation associée à Notch1, exprimé dans l'endocarde, permet la régulation spécifique du myocarde par effet paracrine en induisant l'activation de gènes comme *Bmp10* par exemple (D'Amato et al., 2018; Günthel et al., 2018). Le myocarde trabéculé engendrera par la suite une partie du système de conduction ventriculaire. Ainsi, des gènes associés à l'activité électrique cardiaque sont aussi exprimés dans les trabécules, comme les gènes codant les connexines 40 et 43, ainsi que le canal sodique Nav1.5, ou encore des facteurs de transcription impliqués dans la régulation de l'activité électrique comme ceux de la famille Iroquois Homeobox, *Irx3* et *Irx5* (Gaborit et al., 2012; Günthel et al., 2018; Moorman and Christoffels, 2003; Zhang et al., 2011).

Dans le cadre de la formation du myocarde compact, c'est l'interaction épicarde/myocarde qui entre en jeu, en induisant la prolifération et la croissance des cardiomyocytes pour permettre la formation de la paroi compacte du ventricule (Moorman et al., 2010). Des cellules multipotentes de l'épicarde viennent coloniser le myocarde et se différencient notamment en fibroblastes cardiaques qui constituent des cellules de soutien du tissu et permettent la sécrétion paracrine de facteurs de croissance induisant un épaissement de la paroi ventriculaire (Günthel et al., 2018; Meilhac, 2018).

Le myocarde va ensuite se compacter de plus en plus, aboutissant à une diminution du nombre de trabécules à partir du 44-48^{ème} jour après la fécondation et va provoquer un épaissement de la paroi ventriculaire (D'Amato et al., 2018). Savoir si cette compaction résulte de la fusion des trabécules ou la croissance du myocarde compact reste encore à être déterminée (Meilhac, 2018). La production de cardiomyocytes est tout de même impliquée dans l'augmentation de l'épaisseur de la paroi, qui devient environ 7,5 fois plus épais pendant la période du 37-42^{ème} jour au 56-60^{ème} jour après la fécondation (Figure 5 – 50 jpf) (Günthel et al., 2018).

1.2.D. Système de conduction électrique cardiaque et facteurs de transcription

Une fois la formation des chambres effectuée, les cardiomyocytes n'ont plus d'activité électrique spontanée, contrairement aux cardiomyocytes du myocarde primaire du tube cardiaque. Le développement de tissus conducteurs spécialisés (faible capacité de contraction et automaticité électrique) est donc nécessaire pour la génération et la transmission de l'influx

électrique et pour permettre la contraction cardiaque du myocarde des chambres (Figure 5 – 50 jpf) (Chapitre 17 - Carlson, 2018). Les différents facteurs de transcription, qui sont essentiels à la formation du tissu conducteur, seront présentés.

1.2.D.a. Nœud sinoatrial

Le nœud sinoatrial, situé sous la jonction de l'oreillette droite et de la veine cave supérieure, est le *pacemaker* cardiaque permettant la genèse de l'influx électrique dans le cœur. Il est dérivé du myocarde primaire et provient d'un pool de cellules progénitrices exprimant le marqueur du second champ cardiaque, *Isl1*, et le marqueur des progéniteurs du sinus veineux, *Tbx18* (Mommersteeg et al., 2010). *Isl1* permet l'activation du facteur de transcription *Tbx3* qui est un répresseur du programme d'expression du myocarde contractile et est impliqué dans le maintien de l'intégrité du système de conduction (Miquerol, 2018). *Tbx18* permet d'exprimer les gènes associés à l'activité de *pacemaker* comme *Hcn4*, responsable du courant dépolarisant permettant l'automatisme, ou des connexines spécifiques, *Cx45* et *Cx30.2*. Les facteurs de transcription majeurs *Tbx5* et *Nkx2-5* ont des rôles opposés dans l'établissement du programme génique spécifique du nœud sinoatrial, *Tbx5* induit l'expression de *Tbx3* en faveur de l'établissement du nœud et *Nkx2-5* active l'expression de gènes spécifiques des chambres comme *Gja1*, *Gja5*, *Scn5a* et *Nppa* (Park and Fishman, 2017; van Weerd and Christoffels, 2016). Le facteur de transcription *Shox2* est central dans le développement du nœud et le maintien de l'activité *pacemaker* par répression de *Nkx2-5* (Ye et al., 2015).

1.2.D.b. Nœud atrioventriculaire

L'influx électrique généré dans le nœud sinoatrial va ensuite traverser les oreillettes par les branches de Bachmann, qui assurent la communication interatriale (Van Campenhout et al., 2013). Ensuite, les oreillettes et les ventricules étant séparés par un tissu fibreux isolant, l'influx est contraint de rejoindre le nœud atrioventriculaire, notamment par l'intermédiaire de voies préférentielles (Christoffels et al., 2010; Yi et al., 2012). Le nœud atrioventriculaire est dérivé de cellules du canal atrioventriculaire embryonnaire et a pour fonction de ralentir l'influx électrique provenant des oreillettes pour permettre de séparer leur contraction de celle des ventricules et assurer ainsi un pompage efficace du sang (Bakker et al., 2010). La balance de facteurs de transcription *Nkx2-5/Tbx2-Tbx3* est nécessaire pour le maintien d'un équilibre entre la production de progéniteurs et la spécification et le maintien de l'intégrité du tissu de conduction. Les mêmes gènes que ceux associés à l'activité de *pacemaker* du nœud

sinoatrial (*Hcn4* et les *Cx45* et *Cx30.2* associées à la conduction lente) y sont également exprimés (Miquerol, 2018). La triade de facteurs de transcription *Gata4*, *Nkx2-5*, *Tbx5* permet de réguler la balance entre génération de tissu contractile ou conducteur en agissant sur des programmes d'expression géniques spécifiques (Bhattacharyya and Munshi, 2020; Park and Fishman, 2017).

1.2.D.c. Système de conduction ventriculaire

L'influx électrique passe ensuite, à vitesse croissante, dans le système de conduction ventriculaire composé respectivement du faisceau de His puis des branches droite et gauche du faisceau de His et enfin, du réseau complexe de fibres de Purkinje. Le système de conduction ventriculaire ne dérive pas d'un champ cardiaque spécifique (Miquerol et al., 2013). Le faisceau de His est situé à la jonction entre les ventricules gauche et droit (respectivement issus du premier et du second champ cardiaque), et dérive de cellules des deux champs cardiaques. Le réseau de Purkinje du ventricule droit provient de progéniteurs du second champ cardiaque, alors que la branche gauche du faisceau et le réseau de Purkinje du ventricule gauche sont issus du premier champ cardiaque (Mohan et al., 2018). La branche droite du faisceau de His est issue des deux types de progéniteurs du premier et deuxième champ cardiaque (Choquet et al., 2021). La balance entre les facteurs de transcription *Tbx5* et *Tbx3* est impliquée dans la mise en place de ce système de conduction ventriculaire. En effet, alors que *Tbx3* inhibe l'expression de gènes spécifiques du myocarde contractile (*Nppa*, *Gja1*, ...), *Tbx5* à l'inverse, induit l'expression des gènes associés à la conduction rapide de l'influx électrique : *Gja5*, *Scn5a*, ... (van Weerd and Christoffels, 2016). Le facteur de transcription Iroquois homeobox *Irx3* est également impliqué dans la régulation de l'expression des connexines spécifiques de la conduction rapide au sein des branches du faisceau de His (voir [partie I.4.B Les facteurs de transcription Iroquois homeobox](#)) (Zhang et al., 2011). Les fibres de Purkinje sont quant à elles issues du ventricule trabéculé et le complexe de facteurs de transcription *Irx3/Nkx2-5/Tbx5* y a été décrit comme impliqué dans leur maturation (Kim et al., 2016; Miquerol, 2018; Park and Fishman, 2017). Au 54-56^{ème} jour après la fécondation, le système de conduction est définitivement dissocié du myocarde contractile.

I.2.E. Impact de la naissance sur les cardiomyocytes

Afin de subvenir aux besoins croissants du fœtus en développement, le cœur doit croître et adapter sa fonction de contraction à l'approche de la naissance. Les modifications cardiomyocytaires présentées ci-après ont lieu en parallèle *in vivo* et sont interdépendantes.

I.2.E.a. Maturation du couplage excitation/contraction cardiomyocytaire

Au cours du développement tardif pré- et post-natal, les cardiomyocytes vont subir des maturations associées à leur activité électrique. D'une part, par l'augmentation de l'expression de gènes codant des canaux ioniques, comme par exemple ceux impliqués dans l'homéostasie calcique (ex : *RYR2*, *CACNA1C*, ...). D'autre part, par des modifications structurelles, comme l'apparition des tubules T, l'augmentation de la taille du réticulum sarcoplasmique ou la formation de disques intercalaires entre les cardiomyocytes (Guo and Pu, 2020). Les cardiomyocytes vont également subir des maturations du système contractile, par accumulation de protéines sarcomériques (Meilhac, 2018), augmentant l'efficacité de la contraction, ainsi que par des changements d'isoformes : par exemple la troponine I1 (*TNNI1*) exprimée dans le cœur fœtal est remplacée par la troponine I3 (*TNNI3*) spécifique des cardiomyocytes du cœur adulte (Sheng and Jin, 2014).

I.2.E.b. Switch métabolique

Après la naissance, la disponibilité en oxygène va provoquer un *switch* métabolique des cardiomyocytes (Figure 6) : la glycolyse anaérobie, avec le glucose ou le lactate comme source d'énergie (nécessaire à la prolifération des cardiomyocytes embryonnaires), est remplacée par la phosphorylation oxydative mitochondriale, avec la β -oxydation des acides gras comme nouvelle source énergétique. L'apparition d'un stress oxydatif suite à l'augmentation de l'activité mitochondriale après la naissance contribue à l'arrêt du cycle cellulaire et à l'hypertrophie cardiomyocytaire (Meilhac, 2018; Piquereau and Ventura-Clapier, 2018).

I.2.E.c. Inversion de la balance prolifération/hypertrophie des cardiomyocytes

Entre la fin du développement prénatal et les premières années de vie après la naissance, la prolifération des cardiomyocytes diminue progressivement et l'hypertrophie physiologique des cardiomyocytes entre en jeu (Figure 6) (Tan and Lewandowski, 2020). Chez la souris, la relocalisation nucléaire du facteur de transcription Meis2 après la naissance va induire une inhibition du cycle cellulaire et donc réprimer la prolifération des cardiomyocytes

(Muralidhar and Sadek, 2016). L'hypertrophie des cardiomyocytes se fait par la suite par division cellulaire mitotique sans séparation cytoplasmique, ce qui entraîne la formation de cellules myocardiques binucléées pour favoriser la croissance du myocarde (Tan and Lewandowski, 2020). Ainsi, au cours des premiers mois de la vie, le volume des cellules myocardiques est augmenté de 30 fois environ chez l'humain.

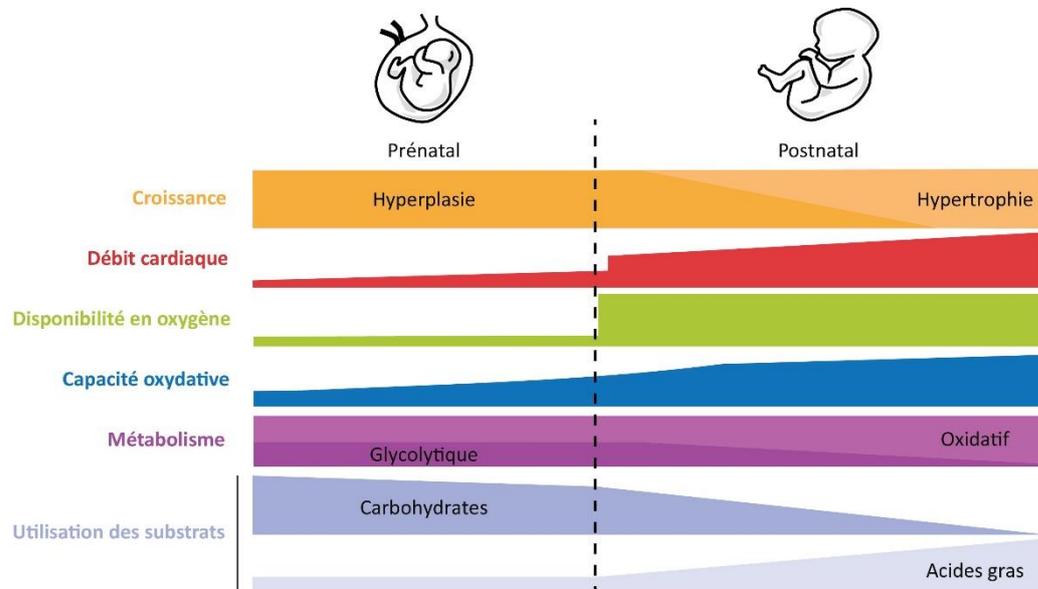


Figure 6 : Modifications du métabolisme cardiomyocytaire associées à la naissance
Adapté de Piquereau and Ventura-Clapier, 2018 et Tan and Lewandowski, 2020.

1.3. Structuration de la chromatine et régulation transcriptionnelle

Il apparaît aujourd'hui clairement que la chromatine n'assure pas uniquement le stockage de l'ADN, mais est aussi très impliquée dans la régulation de l'expression des gènes (Even-Faitelson et al., 2016; Moore-Morris et al., 2018). Dans cette section seront présentées les différentes caractéristiques structurelles et fonctionnelles de la chromatine et leurs implications dans la régulation de la transcription.

1.3.A. Organisation de la chromatine

1.3.A.a. Composition de la chromatine

Si l'on déroulait les trois milliards de paires de bases environ qui composent l'ADN humain, on obtiendrait une molécule de plus de deux mètres, qui est stockée dans un noyau cellulaire d'une taille inférieure à 10µm de diamètre. Ainsi, pour pouvoir entrer dans le volume du noyau, l'ADN est compacté sous forme de chromatine (Doğan and Liu, 2018; McGinty and Tan, 2015). La chromatine est une ribonucléoprotéine correspondant à l'assemblage de

molécules d'ADN enroulées autour de groupes protéiques, les octamères d'histones. Ces octamères composés de paires d'histones H2A, H2B, H3 et H4 respectivement, sont entourés de 145-147pb d'ADN, formant ainsi une structure appelée *nucleosome core particle* (Figure 7) (Flaus, 2011; Luger, 2006). La séquence ADN reliant deux *nucleosome core particles* adjacents est appelée ADN *linker*. Sur cet ADN *linker*, une autre protéine histone est fixée, appelée histone *linker*, correspondant généralement à l'histone H1 (présente dans les régions compactes du génome), elle permet une stabilisation du complexe *nucleosome core particle*/ADN (Felsenfeld and Groudine, 2003). L'association du *nucleosome core particle* et de l'histone linker correspond au chromatosome (Figure 7). Le nucléosome correspond ainsi à l'ensemble *nucleosome core particle* et ADN *linker* (fixé ou non par l'histone *linker*). Cependant, malgré ces définitions, le *nucleosome core particle* est généralement identifié comme étant le nucléosome (McGinty and Tan, 2015). Les histones sont globalement chargées positivement, contrairement à l'ADN qui est chargée négativement. Cependant, il existe une zone en surface du *nucleosome core particle*, où la disposition tridimensionnelle des histones H2A et H2B provoque un regroupement dans l'espace d'acides aminés chargés négativement (acides aspartique et glutamique), formant une poche ou patch acide (*acidic patch* en anglais), qui est impliquée dans les interactions inter-nucléosomes (Paul, 2021).

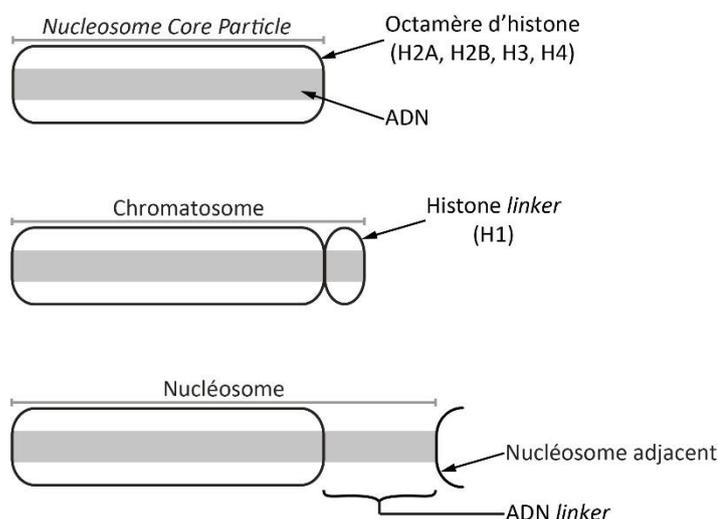


Figure 7 : Structure de la chromatine
Adapté de Flaus, 2011

Le terme de « positionnement des nucléosomes » (ou *nucleosome positioning* en anglais) indique la séquence d'ADN qui les entoure et qui permet d'identifier leur localisation exacte sur le génome. Le positionnement des nucléosomes peut être divisé en deux parties : (1) le positionnement rotationnel, décrivant le côté de l'hélice d'ADN qui fait face aux histones

et (2) le positionnement translationnel, déterminant le point médian du nucléosome par rapport à la séquence d'ADN de 145-147pb qui l'entoure (Flaus, 2011; Liu et al., 2018a). Ce positionnement n'est pas fixe pour tous les nucléosomes, mais peut varier, permettant ainsi la compaction ou l'ouverture de régions spécifiques de l'ADN.

On attribue trois fonctions principales au nucléosome : (1) il permet individuellement le premier niveau de compaction de la chromatine, (2) il sert de base à la fixation d'enzymes et est cible de modifications post-traductionnelles (voir partie I.3.E.b Marques d'histones) participant à la régulation de la chromatine et (3) il permet un niveau de compaction supplémentaire de la chromatine par assemblage avec d'autres nucléosomes (Zhu and Li, 2016).

I.3.A.b. Euchromatine et hétérochromatine

Dans le noyau cellulaire, la chromatine est disposée de manière à ce que chaque chromosome occupe un volume distinct ou territoire chromosomique (Figure 8). Ces territoires sont spécifiques des tissus et des types cellulaires, mais leur fonction reste encore peu claire (Misteli, 2020). Vers le centre du noyau, l'euchromatine aussi appelée compartiment A, correspond à une chromatine relâchée, où les gènes sont actifs. C'est sous cette forme que l'ADN va être transcrit et permettre l'établissement d'un programme d'expression génique spécifique des tissus et des types cellulaires (Morrison and Thakur, 2021). L'euchromatine présente des grandes régions sans nucléosome qui forment ainsi une barrière permettant l'arrêt de l'expansion de l'hétérochromatine et évitant l'inhibition de gènes importants (Allshire and Madhani, 2017). Vers la périphérie du noyau, la chromatine est compactée et les gènes sont inactifs, c'est l'hétérochromatine ou compartiment B. Différentes fonctions de l'hétérochromatine ont été identifiées, comme (1) un rôle dans la spécification des types cellulaires en empêchant l'expression de gènes inadaptés, ou (2) un rôle de protection de l'intégrité de l'ADN en empêchant la survenue de mutations associées à des déplacements de transposons (Allshire and Madhani, 2017).

L'hétérochromatine localisée dans la périphérie du noyau va interagir avec la lamina nucléaire dans ce qu'on appelle des domaines associés à la lamina (Figure 8) (*lamina-associated domain* en anglais) ou LADs, qui ont une taille comprise entre 100 kilobases et 10 mégabases (Guelen et al., 2008). Les LADs sont fortement associés à des régions non transcrites et avec une faible densité en gènes. À l'inverse, entre les LADs, il existe des régions

actives du génome ou inter-LADs. Les LADs d'un même chromosome sont fixés à la lamina nucléaire, alors que les régions inter-LADs d'un même chromosome sont regroupées ensemble vers le centre du noyau (Misteli, 2020). Le rôle du compartiment longeant l'enveloppe nucléaire dans la régulation de la chromatine a été confirmé par différentes expériences complémentaires : (1) des expériences de relocalisation artificielle d'une région interne (active) du noyau vers la périphérie ont montré une répression des gènes de cette région suite à la relocalisation (Reddy et al., 2008) et (2) l'absence de lamina nucléaire provoque une désorganisation des LADs, qui se retrouvent dans une région plus centrale du noyau, avec un impact sur la transcription (Zheng et al., 2018).

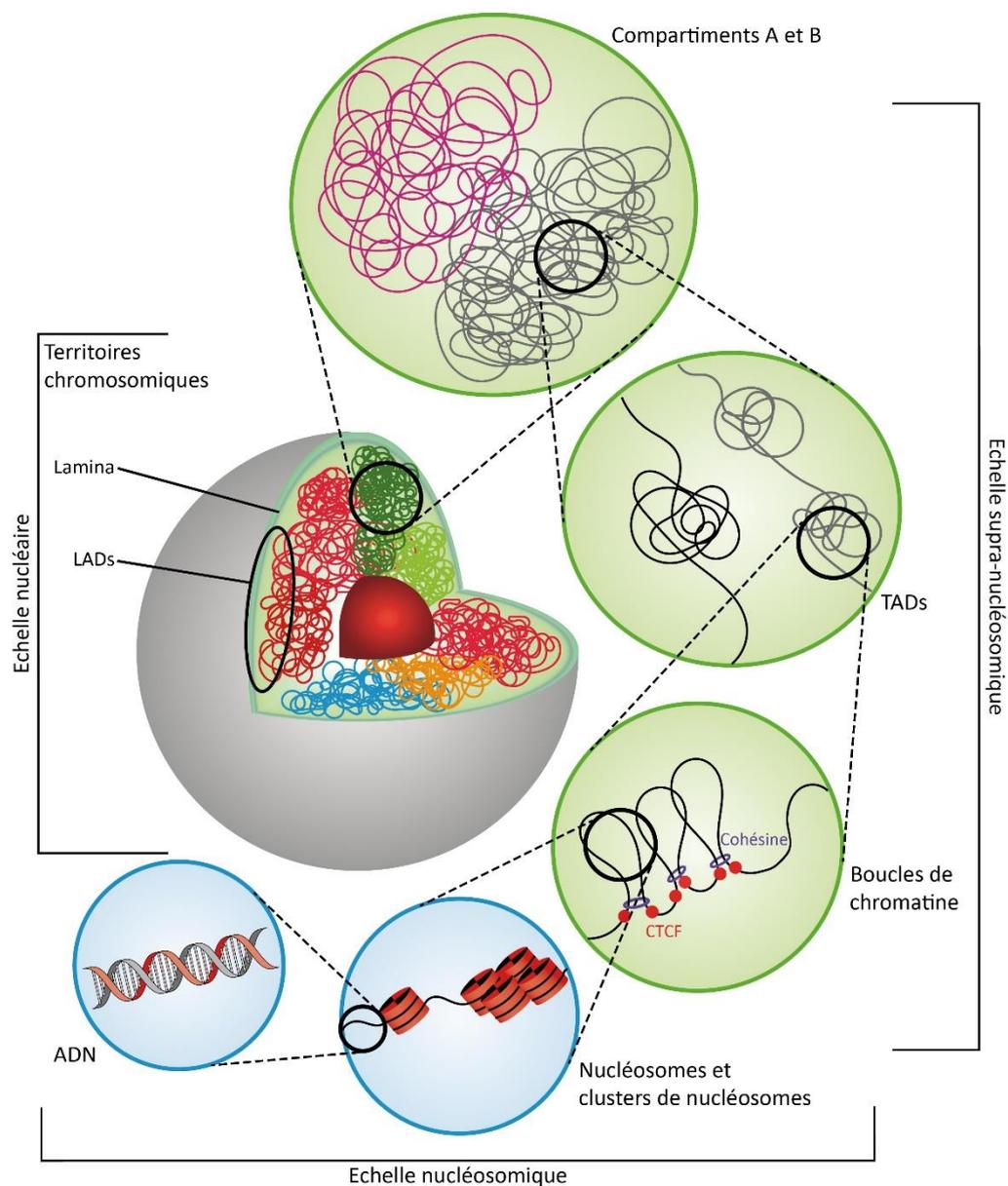


Figure 8 : Différentes échelles d'organisation de la chromatine

LADs : lamina-associated domain, TADs : Topologically Associating Domains. Adapté de Doğan and Liu, 2018.

Pour mieux appréhender le fait que les régions d'ADN ouvertes (euchromatine) ou fermées (hétérochromatine) vont former des structures distinctes (sans membrane), une théorie reposant sur le principe de séparation de phases a été développée (Strom et al., 2017). La base de la séparation de phases est la démixture de populations de protéines distinctes lorsqu'elles sont présentes au-dessus d'un seuil de saturation, provoquant la formation de phases liquide-liquide distinctes et dynamiques (Misteli, 2020; Shin and Brangwynne, 2017). La protéine de l'hétérochromatine, HP1, pourrait être impliquée dans ce phénomène de démixture liquide-liquide (Figure 9). En effet, la mise en place de multiples interactions hydrophobes entre protéines HP1 pourrait permettre la formation de gouttelettes (sans membrane) correspondant aux domaines occupés par l'hétérochromatine, qui ne se mélangeront pas avec l'euchromatine (Sanulli et al., 2019). La phosphorylation de HP1 induit une démixture plus efficace, suggérant un mécanisme potentiel de régulation *in vivo* (Allshire and Madhani, 2017). Il est également supposé que l'accès à ces domaines d'hétérochromatine est régulé par des protéines de transport ou des modifications post-traductionnelles (Strom et al., 2017).

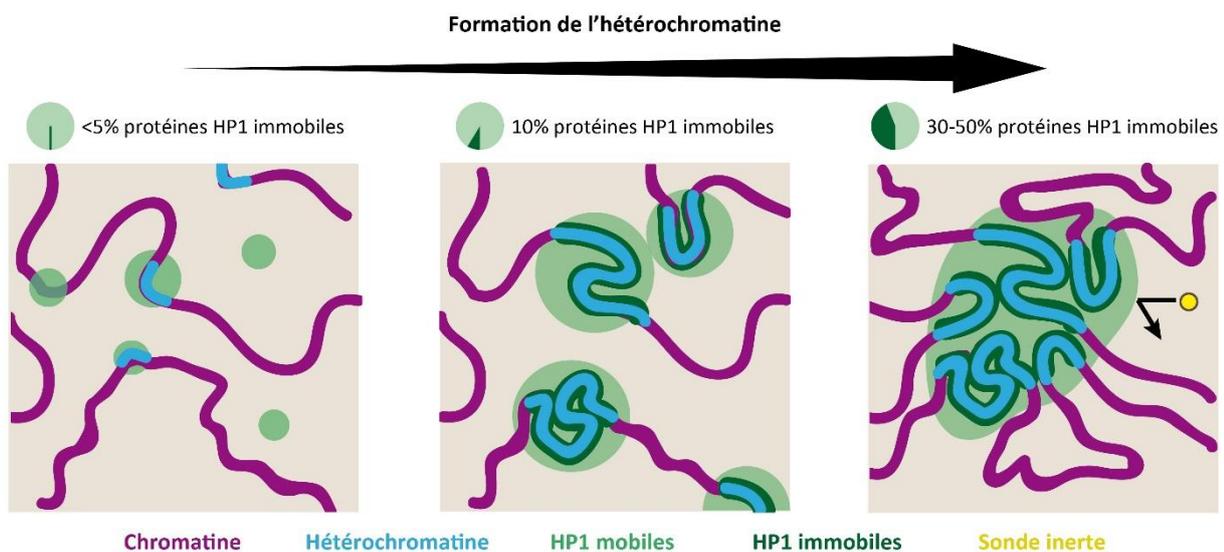


Figure 9 : Modèle de séparation de phase pour la formation de l'hétérochromatine

L'accumulation de la protéine HP1 et la compaction de la chromatine sont favorables à la séparation liquide-liquide induisant la formation de structures d'hétérochromatine sans membranes, inaccessibles à des sondes inertes. Adapté de Strom et al., 2017.

1.3.A.c. Topologically Associating Domains (TADs)

A une échelle plus réduite au sein de l'euchromatine et de l'hétérochromatine se trouvent des domaines structuraux et fonctionnels du génome appelés *Topologically Associating Domains* ou TADs (Figure 8) (Razin et al., 2016). Les TADs correspondent à des

régions de la chromatine, de taille allant de quelques kilobases à plusieurs milliers, qui sont fortement repliées sur elles-mêmes, permettant de nombreuses interactions de chromatine en leur sein (Figure 10A) (Razin and Ulianov, 2017). Le modèle actuel de la formation des TADs est un modèle de boucle-extrusion (Fudenberg et al., 2016) : une boucle d'ADN se forme au travers d'un anneau de cohésine induisant le rapprochement de deux sites de fixation d'un facteur de transcription insulateur (généralement CTCF), qui vont provoquer l'arrêt de la formation de la boucle, une fois arrivés au niveau de l'anneau (Figure 10B). Il existe au sein même des TADs, d'autres niveaux de repliements, appelés sous-TADs (Hsieh et al., 2020; Schmitt et al., 2016), qui constituent des regroupements hétérogènes de boucles transitoires de la chromatine qui interagissent extensivement (Misteli, 2020). Ces différents niveaux de repliements permettent de mettre en contact des éléments de l'ADN dont les séquences sont très éloignées, rendant possible la régulation de l'expression de gènes spécifiques, même à distance (Gonzalez-Sandoval and Gasser, 2016). Les TADs sont globalement bien conservés entre les types cellulaires (Dixon et al., 2012), et en fonction de leur activation ou inhibition, ils permettent l'expression tissu-spécifique de gènes (Ulianov et al., 2016).

Les TADs sont séparés les uns des autres par des régions frontières, qui vont empêcher les interactions de la chromatine entre les TADs adjacents (Kim et al., 2015). Le facteur de transcription CTCF est connu pour jouer le rôle d'insulateur entre les TADs et empêcher les interactions inter-TADs. Certaines régions frontières sont des régions ouvertes (non compactées) et transcriptionnellement actives qui contiennent des gènes de ménage, dont la transcription est forte et n'a pas besoin d'être autant régulée que celle des gènes tissu-spécifiques (Razin and Ulianov, 2017; van Steensel and Furlong, 2019; Ulianov et al., 2016).

Il existe différents exemples de mutations ou variations structurelles de la chromatine conduisant à des défauts de formation de TADs spécifiques (Ibrahim and Mundlos, 2020; Lupiáñez et al., 2015). Ces défauts peuvent être dus notamment à la perte d'une frontière entre les TADs, suite à la mutation d'un site CTCF, provoquant une fusion entre deux TADs qui pourra conduire à des expressions ectopiques de gènes et le développement de pathologies (Figure 10A, C) (Fudenberg et al., 2016).

Bien que la dérégulation de la structure des TADs ait été impliquée dans des pathologies, les effets de variations structurelles sur l'expression de gènes sont ponctuels (Tena and Santos-Pereira, 2021). En effet, différentes études ont montré que la perte de cohésine ou de CTCF induisait une perturbation de la formation des boucles et frontières des

TADs, mais sans modification globale de l'organisation des compartiments A et B, et sans modification profonde de l'expression des gènes (Figure 10C) (Ghavi-Helm et al., 2019; Nora et al., 2017; Schwarzer et al., 2017). Par ailleurs, des exemples de régulations longue-distance entre différents TADs, qui ne tiennent pas compte des frontières, ont été mis en évidence (Galupa et al., 2020; Groff et al., 2018). Ces observations montrent les limites de notre compréhension de la régulation transcriptionnelle par les TADs et suggèrent que d'autres mécanismes de régulation sont impliqués.

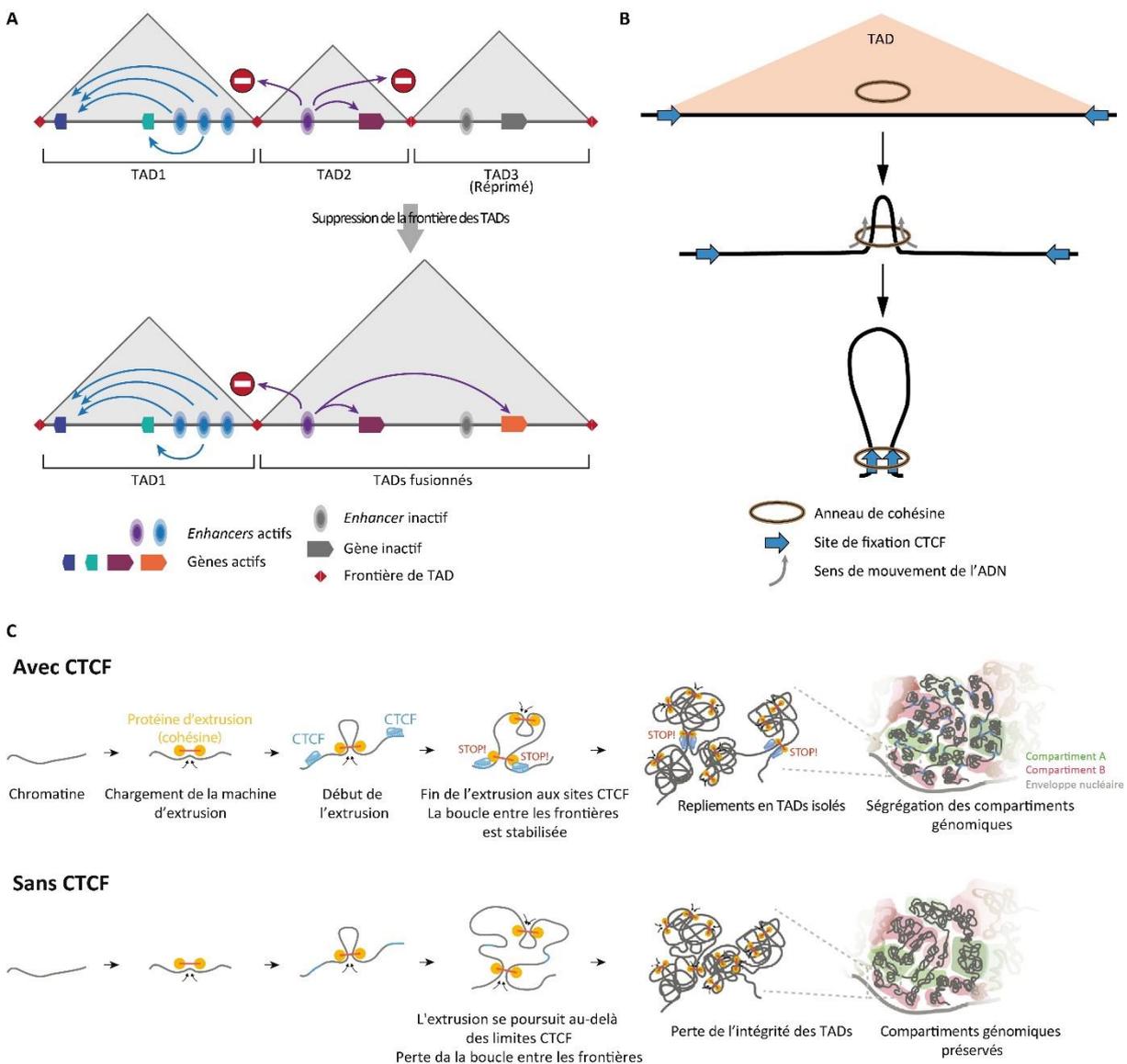


Figure 10 : Formation et structure des TADs

Adapté de Nora et al., 2017, Razin et al., 2016 et Razin and Ulianov, 2017

I.3.B. Résumé des étapes de la transcription des ARNm

Chez les eucaryotes, la transcription des ARNm se fait par l'intermédiaire de l'ARN polymérase II (ARN pol II ; Girbig et al., 2022). L'ARN pol II est une multiprotéine composée de 12 sous-unités qui a besoin du recrutement d'autres partenaires protéiques pour lancer la transcription. L'initiation de la transcription correspond au recrutement, sur le promoteur du gène à transcrire, de l'ARN pol II ainsi que des facteurs d'initiation (aussi connus sous le nom de facteurs de transcription généraux) TFIIA, TFIIB, TFIID, TFIIE, TFIIIF, TFIIH et le complexe médiateur, l'ensemble formant le complexe de pré-initiation de la transcription (Figure 11) (PIC ; Muniz et al., 2021).

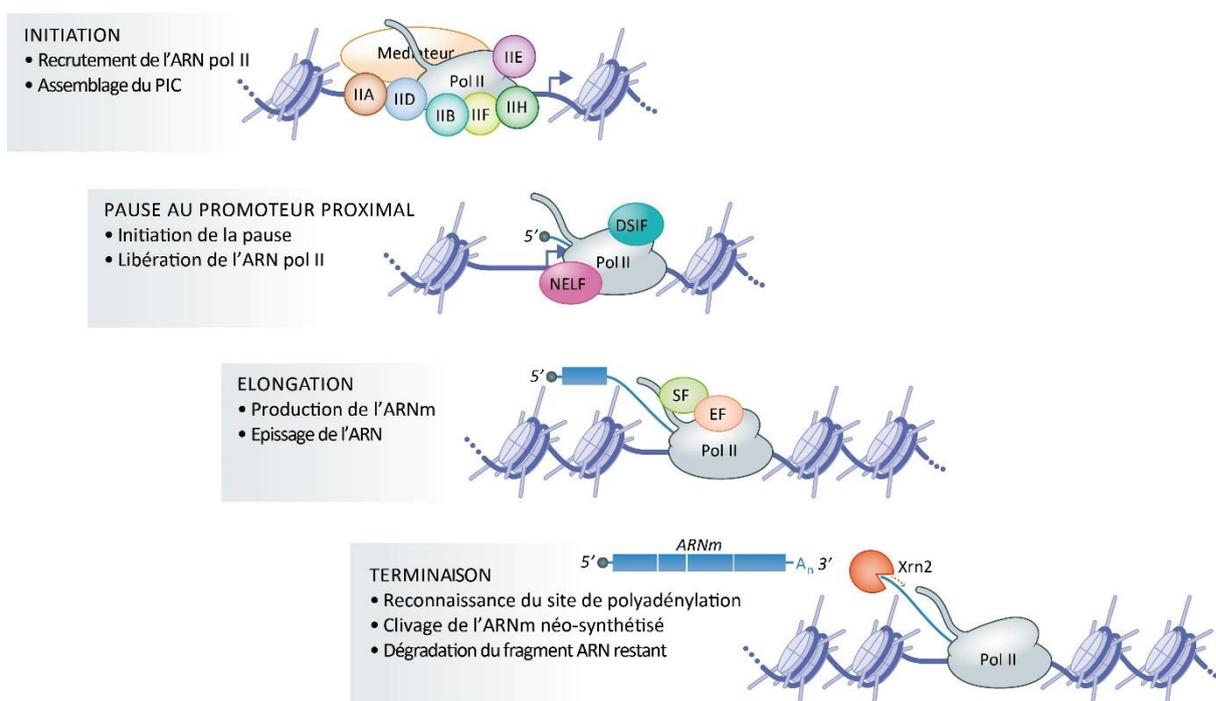


Figure 11 : Etapes du cycle de transcription des ARN messagers

Les facteurs d'initiation de la transcription TFII sont indiqués respectivement IIA, IIB, IID, IIE, IIF et IIH, Pol II : ARN polymérase II, PIC : complexe de pré-initiation de la transcription, NELF et DSIF : facteurs de pause de la polymérase, SF : facteur d'épissage, EF : facteur d'élongation, A_n : queue polyadénylée. Adapté de Muniz et al., 2021.

Le complexe médiateur permet de relier dans un contexte tridimensionnel les éléments régulateurs (ex : *enhancers*) pour favoriser la formation du PIC. Rapidement après le début de la transcription, à environ 20-50pb en aval du point de départ de la transcription (TSS), l'ARN pol II rentre en pause, stabilisée par les facteurs NELF et DSIF. L'apport de facteurs d'élongation, par l'intermédiaire d'autres facteurs de transcription, va entraîner le relargage

de NELF et la conversion de DSIF d'inhibiteur à activateur de l'élongation, entraînant la sortie de la pause de la polymérase puis le début de la phase d'élongation (Figure 11) (Muniz et al., 2021). Lors de cette phase l'ARN pol II doit pouvoir transcrire l'ADN enroulé autour des nucléosomes disposés tout au long du gène. Pour cela, à son arrivée au niveau d'un nucléosome, l'ARN pol II va recruter des facteurs d'élongation et des facteurs remodelers de chromatine pour rendre l'ADN accessible et permettre la poursuite de la transcription (Girbig et al., 2022). La majorité des ARNm possède une queue poly-A, ainsi, la terminaison de la transcription se fait par l'intermédiaire du complexe de clivage et de polyadénylation, qui permet la libération de l'ARNm néo-synthétisé. L'exonucléase Xrn2 dégrade ensuite l'ARN synthétisé après la queue poly-A au sein de l'ARN pol II (Figure 11) (Eaton and West, 2020).

I.3.C. Lien entre la transcription et la conformation 3D de la chromatine

Il existe des interconnexions diverses entre la transcription et la conformation 3D de la chromatine. D'un côté l'activité transcriptionnelle induit des modifications de structuration de la chromatine, par exemple lors de la différenciation des cellules hESCs en cellules spécialisées au cours du développement, l'établissement d'expressions spécifiques de gènes va être nécessaire. Les cellules souches, par leur capacité de pluripotence, expriment un grand nombre de gènes, bien qu'à bas niveau d'expression, ce qui a comme implication sur la structuration de la chromatine la présence d'une grande proportion de régions ouvertes et peu d'hétérochromatine. Lors de la différenciation, ne doivent être exprimés que les gènes spécifiques à la fonction de la cellule en devenir, provoquant ainsi des changements de structuration des TADs et des domaines d'hétérochromatine notamment (Misteli, 2020). Un autre lien entre l'activité transcriptionnelle et la structuration de la chromatine, est la compaction moins importante des TADs présentant une chromatine transcriptionnellement active, comparée aux TADs complètement inactifs, du fait de la présence de la machinerie de transcription (van Steensel and Furlong, 2019; Ulianov et al., 2016).

A l'inverse, bien qu'elle régule la probabilité d'interaction entre des éléments régulateurs et l'accessibilité des promoteurs de gènes pour la fixation de facteurs de transcription, la conformation 3D de la chromatine ne permet pas de déterminer, à elle seule, l'activité transcriptionnelle de tous les gènes de manière binaire (actif/inactif ; Misteli, 2020). Par exemple, la déstabilisation des TADs n'impacte l'expression que d'une petite partie des gènes, mais pas le génome entier, suggérant que la structuration en TADs ne permet pas de

déterminer à elle seule l'activité transcriptionnelle sous-jacente (voir partie I.3.A.c [Topologically Associating Domains \(TADs\)](#)) (Tena and Santos-Pereira, 2021). De plus, Chen *et al.* montrent que la proximité facilitée dans l'espace, entre deux éléments *cis* régulateurs, n'est pas suffisante pour assurer la transcription (Chen *et al.*, 2018).

I.3.D. Eléments *cis* régulateurs

Les éléments *cis* régulateurs correspondent à des séquences généralement non-codantes de l'ADN, qui sont impliquées dans la régulation de la probabilité d'initiation de la transcription de gènes proches ou à distance. Pour cela, les régions régulatrices peuvent être occupées par des facteurs de transcription (Bolt and Duboule, 2020; Wang *et al.*, 2019). Les éléments *cis* régulateurs fonctionnent conjointement à des éléments *trans* régulateurs, qui sont différents de la séquence d'ADN (comme des facteurs de transcription par exemple) et impliqués dans la régulation de la transcription (Wittkopp and Kalay, 2012).

L'ensemble de ces éléments régulateurs partage une fonction commune qui est la régulation de la transcription des gènes (voir partie I.3.B [Résumé des étapes de la transcription des ARNm](#)). Un découpage entre les différentes fonctions que peut avoir une région *cis* régulatrice est réalisé ci-après pour une meilleure compréhension, mais est artificiel car une même région peut présenter en parallèle des caractères de plusieurs catégories ([Tableau 1](#)).

I.3.D.a. Promoteurs

Les promoteurs de gènes sont des régions d'ADN situées en 5' en amont de gènes codants et non-codants pour des protéines. Ils servent de base pour la fixation de l'ARN pol II dans le cas de la transcription d'ARNm (Morrison and Thakur, 2021). Plus précisément, l'ARN pol II va se fixer sur une région particulière du promoteur, le *core promoter*, qui correspond à la séquence minimale permettant la fixation puis l'initiation de la machinerie de transcription (Spitz and Furlong, 2012). Le *core promoter*, d'une taille d'environ 80pb autour du TSS des gènes, n'est pas identique pour tous les gènes, par exemple on estime que 46% des promoteurs possèdent une séquence Inr, et 24% possèdent une séquence TATA-box, qui permettent la fixation de l'ARN pol II (Core and Adelman, 2019; Girbig *et al.*, 2022). Le positionnement des nucléosomes situés dans les promoteurs est bien conservé entre les gènes exprimés. En effet, l'extrémité 5' du premier nucléosome en aval du TSS, appelé « nucléosome +1 », est positionné à environ 40-50pb après le TSS des gènes exprimés (contre

seulement 10pb pour les gènes non exprimés) et est impliqué dans la pause de l'ARN polymérase (Schones et al., 2008; Weber et al., 2014). Le nucléosome -1, est situé à environ 150-180pb en amont des TSS, et son positionnement induit la formation d'une région sans nucléosome (ou *nucleosome-free region* en anglais), qui contient le *core promoter*, et qui est favorable à la fixation de la machinerie de transcription et à la régulation par des facteurs de transcription (Girbig et al., 2022; Hesson et al., 2014; Radman-Livaja and Rando, 2010).

Tableau 1 : Caractéristiques des principaux éléments régulateurs du génome

Caractéristiques	Promoteur	Enhancer	Élément frontière
Accessibilité de la chromatine	+++	+++	+
Activité à distance	+	+++	-
Association avec CTCF	+	+	+++
Association avec la cohésine	+	+	+++
Effet insulateur	+	?	+++
Facteurs de transcription associés			
Ubiquitaires	+++	+	+++
Tissu-spécifiques	+	+++	?
Association avec le PIC	+++	+	-
Transcription			
Unidirectionnelle	+++	-	-
Bidirectionnelle	+	+	-

PIC : Complexe de pré-initiation de la transcription. Adapté de Oudelaar and Higgs, 2021.

1.3.D.b. Enhancers

Les *enhancers* sont de courtes régions d'ADN contenant des motifs de facteurs de transcription qui permettent d'activer la transcription du ou des gènes qu'ils régulent, indépendamment de leur position, de leur distance ou de leur orientation par rapport au promoteur de leur(s) gène(s) cible(s) (Tableau 1) (Oudelaar and Higgs, 2021). En effet, la distance entre un *enhancer* et un gène n'est pas obligatoirement prédictive du lien fonctionnel entre les deux éléments, notamment car l'interaction *enhancer*-promoteur peut s'effectuer par l'intermédiaire de larges boucles de chromatine (voir partie 1.3.A.c *Topologically Associating Domains (TADs)*) (Hamamoto and Fukaya, 2022). De plus, il a été montré que la modification de l'orientation d'un *enhancer* n'a pas d'effet sur sa fonction activatrice (Agrawal et al., 2018). Les *enhancers* sont des régions bien conservées entre les espèces pour un certain nombre d'entre-elles (Kathiriya et al., 2015), cependant elles ne sont pas ubiquitaires au sein d'un individu, mais spécifiques des tissus (Oudelaar and Higgs, 2021).

Différents modèles ont été établis pour décrire les modes d'action supposés des *enhancers* (Figure 12) (Junion et al., 2012; Spitz and Furlong, 2012) :

- Tout d'abord le modèle d'*enhanceosome*, qui proposait que l'*enhancer* soit composé de différents motifs de facteurs de transcription organisés et orientés de manière stricte, permettant la fixation coopérative de différents facteurs de transcription.
- Puis le *billboard model* qui suggère que l'organisation des motifs de facteurs de transcription (espacement et orientation) n'a que peu d'importance, ainsi chaque facteur de transcription est recruté indépendamment via son motif spécifique.
- Enfin le modèle de collectif de facteurs de transcription combine des idées des autres modèles en proposant une coopération de fixation entre les facteurs de transcription sans organisation stricte des leurs motifs. Dans ce modèle l'absence d'un seul facteur de transcription empêche la fixation des autres permettant une régulation fine de l'activité de l'*enhancer*.

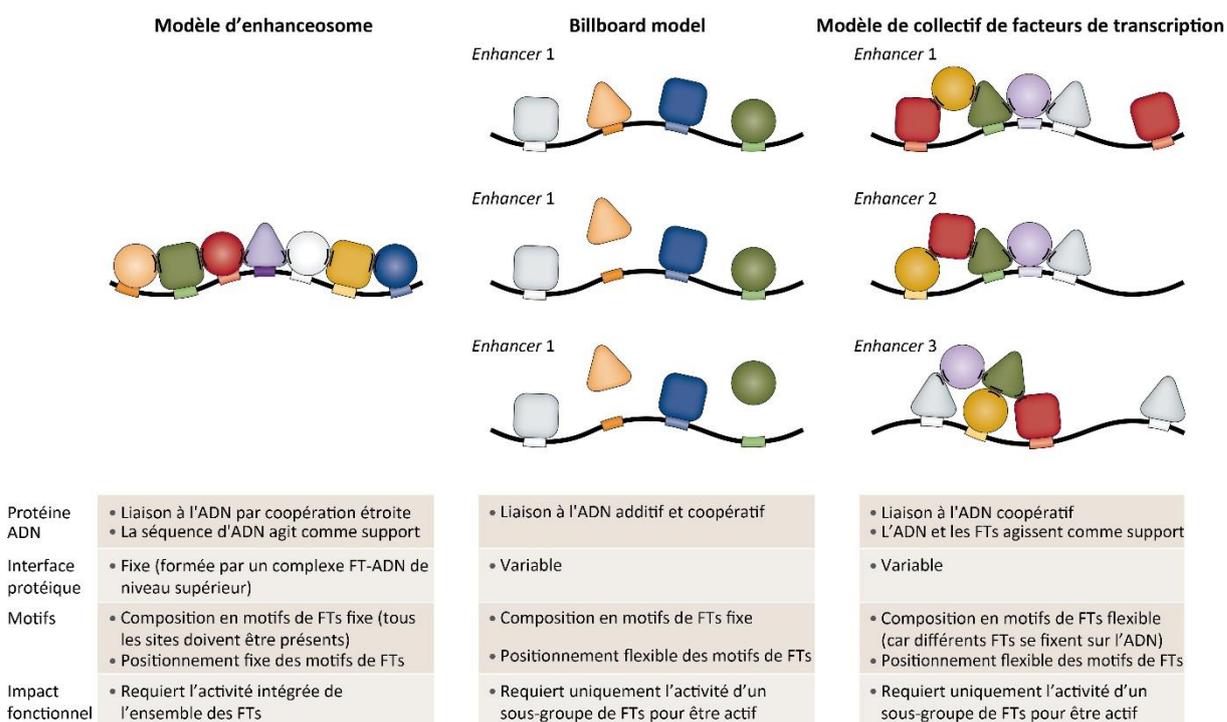


Figure 12 : Modèles de fonctionnement des *enhancers*

Adapté de Spitz and Furlong, 2012

De nombreux promoteurs de gènes ont été montrés comme étant impliqués dans des interactions *enhancer*-promoteur impliquant plusieurs *enhancers* en parallèle (Dunham et al., 2012). Ce type de régulation suggère une redondance d'activation entre différents *enhancers*, rendant ainsi l'activation du gène moins sensible aux mutations perte-de-fonction, comparée à l'activation médiée par un seul *enhancer* (Oudelaar and Higgs, 2021). Il existe aussi des

regroupements physiques de plusieurs *enhancers*, appelés *super-enhancers*, qui vont également réguler différents gènes à distance, indépendamment de leur orientation. Ces *super-enhancers* permettent notamment la régulation de gènes majeurs de l'identité des cellules, comme par exemple le gène *Nanog*, qui est impliqué dans la régulation de la pluripotence dans le développement embryonnaire et dont l'expression est régulée par trois *super-enhancers* différents (Agrawal et al., 2018).

Malgré l'accumulation de résultats décrivant des exemples d'interactions *enhancer-promoter*, les mécanismes associés à la spécificité de ces interactions restent encore peu clairs (Galouzis and Furlong, 2022). En effet, un *enhancer* peut préférentiellement activer des gènes à plusieurs milliers de paires de bases de distance, plutôt que des gènes voisins proches (Chua et al., 2022). Pour tenter d'expliquer ce type d'interaction, Bergman *et al.* proposent un modèle où la majorité des *enhancers* activent de la même manière tous les promoteurs de gènes adjacents et où les activités intrinsèques des *enhancers* et promoteurs (qui sont associées aux motifs de facteurs de transcription présents) se multiplient pour déterminer l'impact sur la transcription (Bergman et al., 2022). Enfin, des exemples de régulation longue-distance, sans interaction physique entre *enhancers* et promoteurs ont été décrits. Ces régulations suggèrent qu'il existe d'autres mécanismes indépendants de la formation de boucles de chromatine et de l'interaction physique d'éléments distants, pour la régulation de l'expression de gènes par les *enhancers* (Alexander et al., 2019; Benabdallah et al., 2019).

La vision globale de la régulation de l'expression génique se complexifie quand on s'intéresse spécifiquement à la dichotomie entre les promoteurs et *enhancers*. En effet, ces deux types d'éléments régulateurs partagent des caractéristiques communes (Tableau 1), comme la présence d'une *nucleosome-free region*, la possibilité de recruter des facteurs de transcription et/ou le complexe de pré-initiation de la transcription (Andersson and Sandelin, 2020). Bien qu'il soit décrit que les *enhancers* vont recruter généralement des facteurs de transcription tissu-spécifiques et agir à distance des gènes, alors que les promoteurs recrutent généralement des facteurs de transcription plus ubiquitaires et sont situés proche des TSS, des promoteurs peuvent servir d'*enhancers* pour d'autres gènes et des *enhancers* peuvent également déclencher la transcription de gènes (Oudelaar and Higgs, 2021). La combinaison des fonctions des promoteurs et *enhancers* permet la création de paysages de régulation

génique (ou *gene regulatory landscapes* en anglais) qui contrôlent les mécanismes complexes et flexibles de l'activité transcriptionnelle (Bolt and Duboule, 2020; Robson et al., 2019).

1.3.D.c. Silencers et insulateurs

Les *silencers* sont des régions du génome aux caractéristiques proches de celles des *enhancers*, à la différence majeure que leur fonction est d'induire une inhibition de l'expression des gènes. L'inhibition est effectuée par l'intermédiaire de facteurs de transcription intrinsèquement inhibiteurs ou s'associant à des cofacteurs inhibiteurs (Maston et al., 2006; Pang and Snyder, 2020).

Les régions insultrices du génome, appelées également régions frontières, sont occupées par des facteurs de transcription insulateurs (comme CTCF) et ont pour fonction d'empêcher les interactions physiques et fonctionnelles entre des régions du génome (Maston et al., 2006). Les insulateurs participent à la régulation transcriptionnelle en induisant le découpage du génome en domaines de régulations (correspondant à des boucles de chromatine ou des TADs à plus grande échelle) et empêchent ainsi les interactions *enhancer-promoteur* inappropriées ([voir partie I.3.A.c Topologically Associating Domains \(TADs\)](#)).

I.3.E. Marques épigénétiques

1.3.E.a. Définition(s) d'épigénétique

Le terme d'épigénétique provient de la combinaison du préfixe « épi » signifiant « sur » ou « au-dessus » et du mot « génétique » faisant référence aux gènes et la séquence d'ADN (Jablonka and Lamb, 2002). Ce terme désigne donc littéralement les événements qui se situent « au-delà » ou « au-dessus » des gènes et de la séquence d'ADN.

La définition d'épigénétique a évolué de nombreuses fois depuis sa première utilisation et ses différentes variantes ont été bien détaillées par Deans and Maggert, 2015. Le terme épigénétique est défini la première fois en 1942 par Conrad Waddington (embryologiste) qui s'intéresse particulièrement aux mécanismes qui contrôlent la différenciation cellulaire (Goldberg et al., 2007). Selon Waddington l'épigénétique correspond à « la branche de la biologie qui étudie les interactions causales entre les gènes et leurs produits qui donnent naissance au phénotype » (Waddington, 2012).

En 1958, David Nanney propose l'hypothèse suivante, les mécanismes épigénétiques seraient des mécanismes auxiliaires des composants génétiques. Selon Nanney, les composants génétiques sont responsables du maintien et de la perpétuation de groupes de gènes

exprimés et non exprimés par le mécanisme de réplication de l'ADN et, les composants épigénétiques permettent le contrôle type-cellulaire spécifique de l'expression des gènes (Deans and Maggert, 2015; Nanney, 1958). Nanney inclut aussi l'idée de la persistance des schémas d'expression type-cellulaire spécifiques au cours des divisions cellulaires, intégrant progressivement la notion d'héritabilité dans la définition d'épigénétique.

En 1994, Robin Holliday va plus loin que Waddington et Nanney dans sa définition de l'épigénétique en y incluant (1) le caractère indispensable de la notion d'héritabilité et (2) la notion que cette héritabilité n'est pas basée sur des modifications de la séquence d'ADN (Holliday, 1994). Cette définition de Holliday va apporter une ambiguïté et des utilisations diverses du terme d'épigénétique dans la communauté scientifique. D'un côté l'utilisation de la définition de Waddington, associée à l'interconnexion entre les éléments du cytoplasme et les éléments génétiques pour produire un phénotype, va être utilisée principalement pour décrire les liens environnement-phénotype. D'un autre côté la définition de Holliday, associée à l'héritabilité et les aspects chromosomiques sera utilisée dans le cas d'études plus moléculaires, comme celles liées à la méthylation de l'ADN ou l'état d'activité de la chromatine (Deans and Maggert, 2015).

En 2007, Adrian Bird prend l'exemple du mécanisme épigénétique de la stabilité des neurones, qui perdure dans la durée sans transmission aux cellules filles (car les neurones ne se divisent pas), pour pondérer la notion d'héritabilité en retirant son aspect obligatoire. Il propose la définition suivante : « l'épigénétique correspond à l'adaptation structurelle des régions chromosomiques de manière à enregistrer, signaler ou perpétuer des états d'activités modifiés » (Allshire and Madhani, 2017; Bird, 2007).

La définition d'épigénétique la plus utilisée actuellement correspond aux mécanismes modifiant l'expression des gènes sans modification de la séquence d'ADN et pouvant également être transmis aux cellules filles lors de divisions cellulaires (Deichmann, 2020; Lind and Spagopoulou, 2018). Cette définition fait principalement référence à la méthylation de l'ADN (qui ne sera pas présentée ici) et aux marques d'histones ([voir partie I.3.E.b Marques d'histones](#)).

1.3.E.b. Marques d'histones

Dans le cadre de la régulation transcriptionnelle, les différents composants de la chromatine vont chacun avoir un rôle important, ainsi, en plus de séquences spécifiques d'ADN ([voir partie I.3.D Eléments cis régulateurs](#)), les protéines histones vont permettre un

autre niveau de contrôle. Chacune des histones possède une queue protéique, composée d'acides aminés chargés positivement (principalement des lysines) (Figure 13), qui vont être la cible de modifications post-traductionnelles (ex : acétylation, méthylation, phosphorylation, ... ; Albin et al., 2019; Strahl and Allis, 2000). Ces modifications, également appelées marques d'histones ou marques épigénétiques, vont être localisées sur des acides aminés spécifiques des queues d'histones et vont renseigner de la conformation tridimensionnelle de la chromatine et de l'activité transcriptionnelle sous-jacente, formant ainsi le code histone (Tableau 2) (Allis and Jenuwein, 2016).

Les marques d'histones ne vont pas simplement servir de marqueur de l'état de la chromatine ou de la transcription, mais elles vont avoir un rôle direct sur la conformation tridimensionnelle de la chromatine et sur le recrutement de protéines spécifiques. Les marques d'histones et les protéines qui leur sont associées vont intervenir à chaque étape de la transcription pour permettre sa régulation (Gates et al., 2017; Morgan and Shilatifard, 2020).

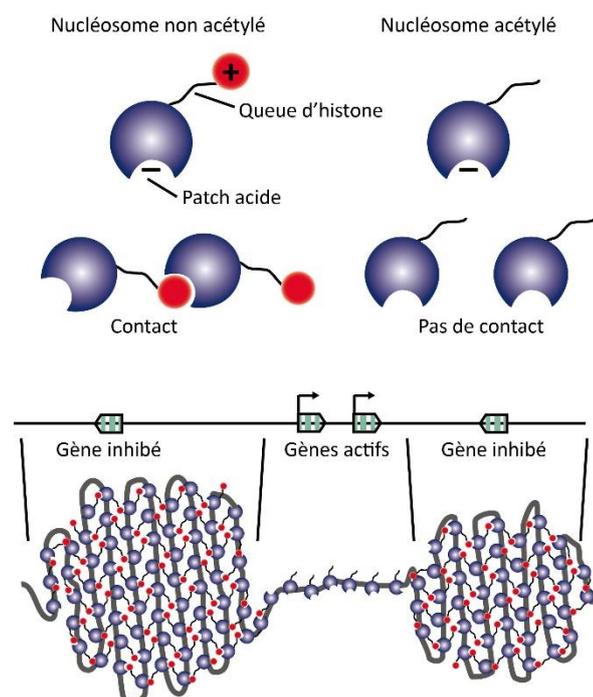


Figure 13 : Impact de la charge des nucléosomes sur la structuration de la chromatine
Adapté de Razin et al., 2016

Par exemple dans le cas de l'acétylation, un groupement acétyle va être positionné sur une lysine d'une queue d'histone, ce qui va provoquer la perte d'une charge positive (Figure 13). La neutralisation de cette charge va ainsi entraîner la déstabilisation du complexe entre la queue d'histone chargée positivement et la poche acide du nucléosome adjacent chargé

négalement, induisant une ouverture de la chromatine favorable à la fixation de la machinerie de transcription (Gibson et al., 2019; Razin et al., 2016). Les marques d'acétylation d'histones sont donc largement considérées comme des marques associées à l'activation de la transcription. A l'inverse, les marques de méthylation d'histones peuvent être associées à la fois à une transcription active ou inactive en fonction du nombre de groupes méthyles et des positions des acides aminés ciblés (Tableau 2). Des expériences de marquage isotopique par impulsion ont permis d'évaluer la durée de demi-vie des différentes marques qui peut varier de quelques minutes à plusieurs jours en fonction des marques (Barth and Imhof, 2010).

Les marques d'histones sont déposées par des protéines appelées *writers* comme les histones méthyltransférases ou acétyltransférases (Hyun et al., 2017). Bien que les marques d'histones soient liées de manière covalente à l'acide aminé cible, elles peuvent être retirées par l'intermédiaire de protéines *erasers*, comme les histones désacétylases (HDAC) par exemple (McKinsey, 2012). Enfin, un troisième type de protéines peut reconnaître et se fixer spécifiquement sur certaines marques d'histones, elles sont appelées *readers* (Atlasi and Stunnenberg, 2017) et présentent un bromodomaine ou un chromodomaine, respectivement pour la fixation spécifique sur des marques d'acétylation ou de méthylation (Barth and Imhof, 2010).

Certaines protéines peuvent présenter simultanément les fonctions de *reader* et *writer*, comme les complexes répresseurs Polycomb 1 et 2 (Golbabapour et al., 2013; Luna-Zurita and Bruneau, 2013). Le complexe multiprotéique Polycomb 2, composé notamment des protéines EZH1/2, SUZ12, EED et RBBP4/7, va agir en tant que *writer*, pour ajouter des marques d'histones associées à la transcription inactive, comme la triméthylation de la lysine 27 de l'histone H3 (H3K27Me3). Le complexe Polycomb 1 va agir en tant que *reader* en se fixant sur ces méthylations pour faciliter la condensation de la chromatine (Goldberg et al., 2007; Schuettengruber et al., 2017). Pour contrebalancer les actions répressives des complexes Polycomb, les complexes Trithorax vont avoir des fonctions inverses en ajoutant des marques associées à la transcription active, comme des marques d'acétylation ou de méthylation spécifiques (Schuettengruber et al., 2011).

Les facteurs de transcription peuvent influencer le paysage épigénétique, par le recrutement de protéines *writer*, *reader* ou *eraser*, et ainsi avoir un impact localisé sur la structuration de la chromatine et la transcription.

Tableau 2 : Descriptif des principales marques épigénétiques de l'histone H3

Marques épigénétiques	Etat transcriptionnel associé	Localisation sur le génome	Références
H3K27Ac	Activation	Promoteurs + <i>Enhancers</i>	Albini et al., 2019; Atlasi and Stunnenberg, 2017; Creighton et al., 2010; Igoikina et al., 2019; Wamstad et al., 2012; Zhang et al., 2022, 2021
H3K9Ac	Activation	Promoteurs + <i>Enhancers</i>	Albini et al., 2019; Igoikina et al., 2019; Karmodiya et al., 2012; Zhang et al., 2022, 2021
H3K14Ac	Activation (en parallèle de marques d'activation) + « Poised » (en parallèle de marques d'activation)	<i>Enhancers</i>	Bannister and Kouzarides, 2011; Karmodiya et al., 2012; Regadas et al., 2021
H3K18Ac	Activation	Promoteurs	Albini et al., 2019; Bannister and Kouzarides, 2011; Luo et al., 2019
H3K8Ac	Activation	Promoteurs	Zhu et al., 2013
H3K64Ac	Activation	Promoteurs (majoritairement) + <i>Enhancers</i>	Atlasi and Stunnenberg, 2017; Di Cerbo et al., 2014; Pradeepa et al., 2016; Zhang et al., 2022
H3K122Ac	Activation	Promoteurs (majoritairement) + <i>Enhancers</i>	Atlasi and Stunnenberg, 2017; Pradeepa et al., 2016; Zhang et al., 2022
H3K4Me1	Activation (en parallèle de marques d'activation) + « Poised » (en parallèle de marques d'activation)	<i>Enhancers</i>	Albini et al., 2019; Atlasi and Stunnenberg, 2017; Bae and Lesch, 2020; Bannister and Kouzarides, 2011; Creighton et al., 2010; Kang et al., 2021; Zhang et al., 2022, 2021
H3K4Me3	Activation	Promoteurs	Albini et al., 2019; Atlasi and Stunnenberg, 2017; Bae and Lesch, 2020; Bannister and Kouzarides, 2011; Igoikina et al., 2019; Karmodiya et al., 2012; Zhang et al., 2021
H3K79Me2	Activation	Sur tout le gène	Albini et al., 2019; Cattaneo et al., 2016; Zhang et al., 2021

H3K36Me3	Controversé (Activation, « Poised », Inactivation)	Sur tout le gène	Albini et al., 2019; Bannister and Kouzarides, 2011; Barral et al., 2022; Chantalat et al., 2011; Sun et al., 2020; Wagner and Carpenter, 2012; Zhang et al., 2021
H3K27Me3	Inactivation	Promoteurs	Albini et al., 2019; Atlasi and Stunnenberg, 2017; Bannister and Kouzarides, 2011; Bogliotti and Ross, 2012; Cai et al., 2021; Igolkina et al., 2019; Kim and Kim, 2012; Zhang et al., 2021
H3K9Me3	Inactivation	Promoteurs	Albini et al., 2019; Atlasi and Stunnenberg, 2017; Bannister and Kouzarides, 2011; Igolkina et al., 2019; Kim and Kim, 2012; Zhang et al., 2021
H3K9Me2	Inactivation	Sur tout le gène + <i>Enhancer</i>	Atlasi and Stunnenberg, 2017; Bannister and Kouzarides, 2011; Zhang et al., 2021

H3K## : correspond à la position des acides aminés lysine sur la queue de l'histone H3 qui vont subir la modification post-traductionnelle indiquée juste après. Ac : acétylation ; Me1,2,3 : mono-, di- et triméthylation respectivement. « Poised » : Etat de la chromatine en « attente », qui pourra être activée plus rapidement que lorsqu'elle est compactée et inactive.

I.4. Les facteurs de transcription

Les facteurs de transcription sont bien connus pour être des protéines se fixant sur des séquences spécifiques de l'ADN et pouvant contrôler positivement ou négativement la transcription de gènes cibles, médiée par l'ARN polymérase (Klemm et al., 2019; Zabet, 2012). Dans cette section seront présentés les différents mécanismes d'action des facteurs de transcription, puis une attention particulière sera portée au rôle des facteurs de transcription Iroquois homeobox dans le cœur et enfin, les principaux modèles et stratégies d'étude des facteurs de transcription cardiaques seront abordés.

I.4.A. Mécanisme d'action

I.4.A.a. Modèle réductionniste

Dans le but de mieux comprendre comment les facteurs de transcription peuvent activer ou inhiber l'expression de gènes spécifiques, des modèles d'action réductionnistes ont d'abord été établis. Historiquement, les facteurs de transcription ont été décrits comme possédant deux caractéristiques principales : (1) la capacité de se lier à l'ADN grâce à des

domaines protéiques spécifiques et (2) la capacité de recruter et stabiliser la machinerie de transcription au promoteur de gènes grâce à un autre domaine, appelé domaine activateur (Latchman, 1997).

Différents domaines protéiques bien conservés au cours de l'évolution permettant aux facteurs de transcription de se fixer spécifiquement à l'ADN ont alors été décrits : comme les homéodomains, les domaines « POU », « ETS » ou en doigt de zinc par exemple (Cheatle Jarvela and Hinman, 2015).

Enfin, lors de l'activation de la transcription par des facteurs de transcription, il avait été proposé que ces derniers se fixent à proximité de gènes cibles (notamment aux promoteurs des gènes) par leur domaine de fixation à l'ADN, pour ensuite permettre la mise en place et la stabilisation du PIC, grâce à leur domaine activateur (Latchman, 1997; Ma, 2011).

1.4.A.b. Modèle intégré

Bien que ces modèles réductionnistes soient valables, ils restent largement incomplets et n'intègrent pas toute la complexité des mécanismes pouvant réguler la transcription comme la structuration tridimensionnelle de la chromatine, les actions combinées de plusieurs facteurs de transcription ou encore les marques épigénétiques.

Les facteurs de transcription ne se fixent pas uniquement aux promoteurs de leurs gènes cibles, mais également au niveau d'autres éléments *cis* régulateurs (voir partie 1.3.D *Éléments cis régulateurs*), pour occuper des centaines voire des milliers de régions génomiques au sein d'un même type-cellulaire. En effet, ce sont les facteurs de transcription qui vont conférer aux éléments *cis* régulateurs une activité régulatrice de la transcription, en permettant la stabilisation ou le blocage de la mise en place du PIC et de la transcription par l'ARN polymérase (Kathiriya et al., 2015). Par exemple, les facteurs de transcription vont se fixer sur des *enhancers* distants de leurs gènes cibles et réguler la mise en place du PIC par la formation de boucles de chromatine (voir partie 1.3.A.c *Topologically Associating Domains (TADs)*). Ils constituent donc le lien central des interactions distantes *enhancer/promoteur* (Panigrahi and O'Malley, 2021). Différents modèles d'interactions *enhancer/promoteur* incluant les facteurs de transcription ont été décrits (Figure 14) (Karr et al., 2022) :

- Le modèle de contact stable, correspondant au modèle classiquement présenté d'interaction entre *enhancers* et promoteurs, qui consiste à la formation d'une boucle

de chromatine assurée par des contacts stables entre des facteurs de transcription et l'ADN.

- Le modèle de « *kiss-and-run* », décrit des situations où le facteur de transcription n'est présent aux abords du promoteur que lors d'un court contact *enhancer*/promoteur.
- Le modèle de communication par diffusion propose que la communication entre *enhancers* et promoteurs s'établisse dans un contexte de proximité, mais sans interaction physique, par diffusion au promoteur d'un facteur de transcription activé par une enzyme recrutée au niveau de l'*enhancer*.

Les deux premiers modèles proposent des mécanismes faisant intervenir des interactions physiques entre *enhancer* et promoteur, à l'inverse, le modèle de communication par diffusion pourrait expliquer comment des régulations « longue distance » sans interaction physique entre *enhancers* et promoteurs, peuvent avoir lieu (Alexander et al., 2019; Benabdallah et al., 2019).

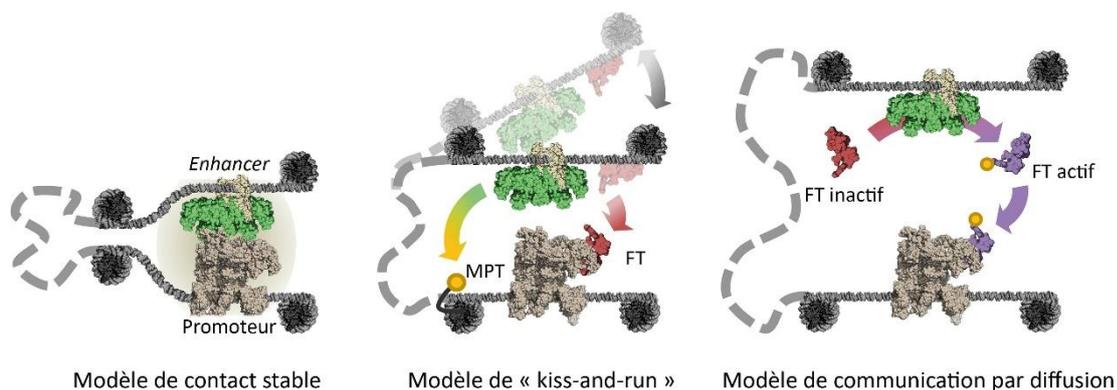


Figure 14 : Différents modèles d'action des facteurs de transcription dans les interactions *enhancer*/promoteur
Le promoteur est occupé par la machinerie de transcription (beige) et l'*enhancer* par un coactivateur (vert), qui peut être un autre facteur de transcription ou une enzyme (modèle par diffusion). MPT : modification post-traductionnelle, FT : facteur de transcription. Adapté de Karr et al., 2022.

Les facteurs de transcription vont se fixer spécifiquement aux *enhancers* et promoteurs par reconnaissance de petits motifs de fixation spécifiques sur l'ADN (*binding sites* ou *binding motifs* en anglais), d'une taille allant de 6 à 12pb environ (Lambert et al., 2018). Il a été montré que la fonctionnalité des motifs est corrélée à leur conservation au cours de l'évolution (les motifs fonctionnels sont plus conservés que ceux dont la fonctionnalité n'a pas été prouvée) et que les motifs associés à l'activation de gènes sont localisés plus proches des TSS que ceux associés à l'inhibition des gènes (Whitfield et al., 2012). Les motifs décrits de facteurs de transcription correspondent à des séquences consensus (séquences surreprésentées de manière générale sur l'ADN), mais peuvent présenter des variations mineures de séquence *in*

vivo, entraînant des changements d'affinité du facteur de transcription pour le motif (Boeva, 2016). D'une part, des variations génétiques rares ou fréquentes d'un nucléotide, appelées SNPs (*Single Nucleotide Polymorphisms* en anglais), peuvent être présents dans ces motifs et altérer la régulation de la transcription, aboutissant à des modifications phénotypiques pouvant être pathologiques (Altshuler et al., 2010; Sachidanandam et al., 2001). Les SNPs associés à des modifications d'expression de gènes correspondent à des régions eQTL (*expression Quantitative Trait Loci* en anglais, Nica and Dermitzakis, 2013). D'autre part, il existe des sites de fixation de facteurs de transcription de faible affinité, aussi appelés motifs fantômes. Ils n'ont pas la capacité de maintenir, à eux seuls, une fixation assez longue des facteurs de transcription pour provoquer l'activation ou l'inhibition de la transcription de gènes (Boeva, 2016). Dans certains contextes tridimensionnels de la chromatine, des regroupements dans l'espace de ces sites fantômes de faible affinité permettent la concentration de nombreux facteurs de transcription dans des régions spécifiques de l'ADN. Cette concentration de facteurs de transcription favorise le phénomène de séparation de phase (voir partie I.3.A.b *Euchromatine et hétérochromatine*) (Boija et al., 2018), créant ainsi des hubs ou condensats de facteurs de transcription qui vont pouvoir réguler l'expression des gènes de manière spécifique (Figure 15) (dépendant du nombre et de la diversité des motifs ; Crocker et al., 2016; Kribelbauer et al., 2019).

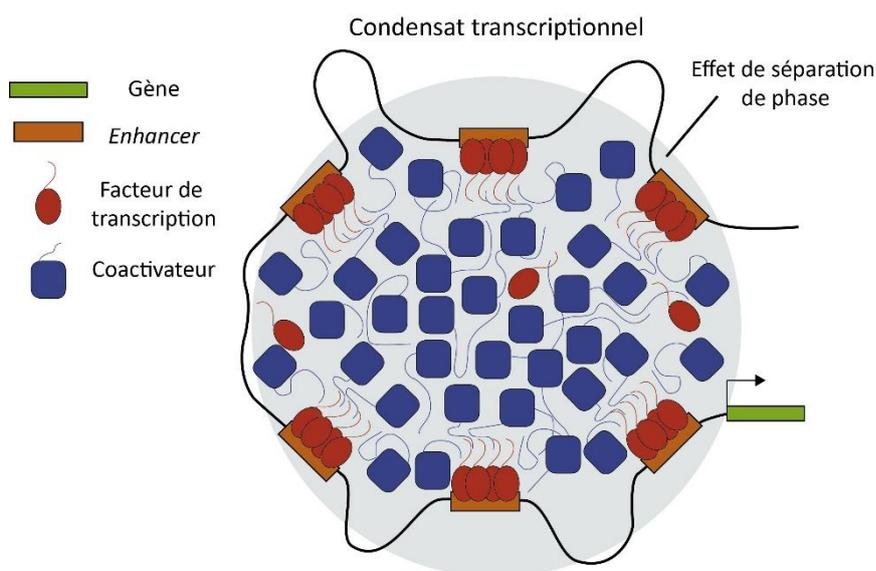


Figure 15 : Structuration d'un condensat transcriptionnel

Le positionnement 3D de différents *enhancers* provoque le regroupement de nombreux facteurs de transcription, favorisant l'effet de séparation de phase qui stabilise ces interactions et permet l'activation spécifique de gènes. Adapté de Boija et al., 2018.

Les éléments *cis* régulateurs, en particulier les enhancers, sont composés de plusieurs motifs de facteurs de transcription différents (voir partie I.3.D.b Enhancers). Cette diversité de motifs confère aux *enhancers* (1) une spécificité d'action dans le cas où tous les facteurs de transcription recrutés sont indispensables (pas de redondance fonctionnelle entre eux), ou (2) une tolérance aux mutations génétiques dans le cas où la fonction des facteurs de transcription recrutés est partiellement redondante (Figure 16) (Spivakov, 2014).

Pour stabiliser les boucles de chromatine entre *enhancers* et promoteurs, et exercer des fonctions spécifiques, les facteurs de transcription ne vont pas agir de manière isolée, mais vont s'associer à différents partenaires protéiques.

Chez les eucaryotes, de nombreux facteurs de transcription ne peuvent pas se fixer seuls à l'ADN et forment donc des homodimères pour pallier ce problème (comme c'est généralement le cas sur des motifs palindromiques, Mitsis et al., 2020). De plus, la formation de complexes multiprotéiques entre facteurs de transcription est nécessaire dans les régions composées de sites de fixation de faible affinité, pour un renforcement mutuel de leur fixation par coopération (Spitz and Furlong, 2012). Différents facteurs de transcription peuvent également être recrutés de manière séquentielle et dynamique sur un site donné (alternant entre les états fixés et non-fixés) et ils semblent donc co-résidents du site, si on tient compte de leur occupation moyennée au sein d'une population cellulaire (Voss and Hager, 2014). Des effets synergiques sur la transcription peuvent être observés au sein des complexes multiprotéiques de facteurs de transcription, du fait d'actions combinées des membres du

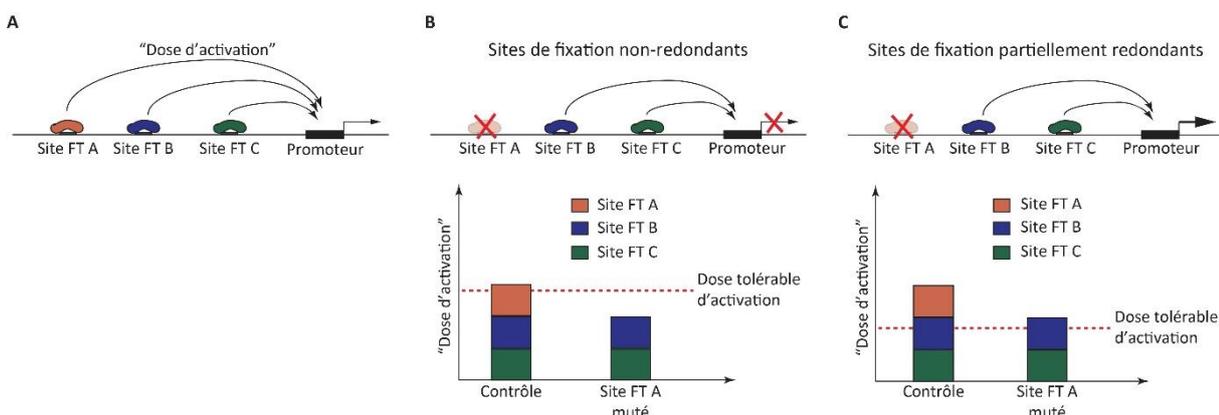


Figure 16 : Notion de « dose d'activation » de gènes par l'action des facteurs de transcription

A. La présence de différents motifs de facteur de transcription (Site FT) dans les *enhancers* suggère une activation cumulative (« dose d'activation ») sur le promoteur de gènes cibles, suite à la fixation des facteurs de transcription. **B-C.** Dans le cas où tous les facteurs de transcription doivent être fixés, l'activation est très spécifique, mais n'est pas tolérante aux mutations génétiques, contrairement aux sites de fixation partiellement redondants. Adapté de Spivakov, 2014.

complexe à différentes étapes de la transcription (voir partie I.3.B Résumé des étapes de la transcription des ARNm) (Ma, 2011). Un exemple bien décrit de complexe de facteurs de transcription, est celui impliquant GATA4, NKX2-5 et TBX5 dans le cœur en développement (Kathiriya et al., 2015; Luna-Zurita et al., 2016; McCulley and Black, 2012).

Les facteurs de transcription peuvent aussi s'associer physiquement à d'autres cofacteurs, comme notamment dans le cadre de la stabilisation du PIC (Figure 15). En effet, bien que certains facteurs de transcription soient capables de fixer directement le facteur TFIID pour favoriser la fixation du PIC au promoteur, la stabilisation des boucles *enhancer/promoteur* nécessite généralement des interactions avec des cofacteurs, comme le complexe médiateur par exemple (Chen and Pugh, 2021). Les facteurs de transcription peuvent agir à la fois comme activateur et comme répresseur en fonction du contexte et des cofacteurs associés (Ma, 2011). Les cofacteurs sont ainsi appelés coactivateurs ou corépresseurs et contiennent généralement des domaines impliqués dans (1) la liaison à la chromatine, (2) le positionnement des nucléosomes et/ou (3) la modification covalente des histones ou d'autres protéines, comme des facteurs de transcription ou l'ARN polymérase (Lambert et al., 2018; Mitsis et al., 2020). Il existe de nombreux cofacteurs de facteurs de transcription impliqués dans la régulation de la transcription, par exemple, plus de 400 protéines impliquées dans les modifications post-traductionnelles des histones ont été décrites (Lambert et al., 2018).

Les facteurs de transcription ne se fixent pas de manière covalente à l'ADN, mais ont des fixations dynamiques et transitoires sur la chromatine. Des expériences de suivi moléculaire unique (ou « *single-molecule tracking* ») ont démontré qu'un facteur de transcription passe généralement plus de 95% de son temps dans un état non fixé à la chromatine ou engagé dans des interactions non spécifiques, avec des temps de fixation de l'ordre de plusieurs centaines de millisecondes. Seule une petite fraction d'un facteur de transcription donné est liée à des sites spécifiques avec des temps de fixation de l'ordre de quelques dizaines de secondes (Azpeitia and Wagner, 2020; Misteli, 2020). Bien que ces événements de fixation soient transitoires, les états de la chromatine qui en résultent sont stables en raison du remplacement continu des protéines qui se dissocient (Misteli, 2020). Cependant, Gurdon *et al.* ont décrit des temps de fixation bien plus importants (de l'ordre de plusieurs heures ou jours) pour certains facteurs de transcription et ils suggèrent qu'il existe

deux modes de fixation des facteurs de transcription : (1) une fixation à court terme dans les cellules en division et en prolifération et (2) une fixation à long terme dans les cellules qui ne se divisent pas (comme les cellules adultes par exemple) pour garantir l'expression génique dans le temps et l'état différencié des cellules (Gurdon et al., 2020).

Le processus de transcription n'est pas continu mais oscille entre des états « on » et « off », générant des salves de transcription (ou « *transcriptional bursting* »), comme les expériences de transcriptomique sur cellules uniques ont également pu mettre en évidence (Misteli, 2020; Voss and Hager, 2014). Un mécanisme important pour contrôler le nombre de transcrits produits, est de réguler le temps de la fenêtre de transcription, qui se traduit dans le cas d'un facteur de transcription activateur, par une augmentation de la probabilité de déclenchement des salves plutôt que par une augmentation de la quantité de transcrits produits par salve (Ma, 2011). Pour que ces salves soient obtenues il faut (1) tout d'abord mettre en place le PIC et initier la transcription, notamment par des liaisons d'activateurs, des boucles *enhancer/promoteur* et/ou le remodelage de la chromatine et des nucléosomes. Il faut ensuite (2) que le promoteur soit dans un état « ouvert » et (3) que cet état permissif soit maintenu dans le temps, car plusieurs ARN polymérases peuvent être chargées pendant une même salve. En effet, des salves recrutant jusqu'à 200 ARN polymérase ont été décrites chez les eucaryotes. Les salves sont de durée limitée indiquant que le promoteur passe d'un état actif à inactif (Rodriguez and Larson, 2020). Les facteurs de transcription sont donc des régulateurs majeurs du déclenchement des salves de transcription.

1.4.A.c. Lien entre le contexte épigénétique et l'activité des facteurs de transcription

Il existe un impact fonctionnel de la structuration de la chromatine associée aux marques épigénétiques sur l'activité des facteurs de transcription, et inversement (Kathiriya et al., 2015).

La position des nucléosomes va avoir un impact sur le positionnement des facteurs de transcription sur la chromatine dû à un encombrement stérique au site de fixation du facteur de transcription (Spitz and Furlong, 2012). Les facteurs de transcription vont donc préférentiellement se fixer dans des régions ouvertes, sans nucléosome : le génome accessible correspond à environ 2-3% de la séquence totale d'ADN, mais capture plus de 90% des régions liées par les facteurs de transcription (Klemm et al., 2019). La présence de marques

épigénétiques sur les nucléosomes va aussi avoir un impact direct sur la structuration de la chromatine et la fixation des facteurs de transcription (voir partie I.3.E.b Marques d'histones). Par exemple, la présence de marques d'acétylation (ex : H3K27Ac) est associée à l'ouverture de la chromatine augmentant la probabilité de fixation des facteurs de transcription, et à l'inverse, certaines marques de méthylation (ex : H3K27Me3) sont associées à une chromatine inactive et fermée, empêchant la fixation des facteurs de transcription. Il est donc possible de prédire l'affinité de fixation des facteurs de transcription en fonction des marques épigénétiques présentes (Liu et al., 2015). Dans le sens opposé, il est aussi possible de prédire les marques épigénétiques par la présence de motifs de facteurs de transcription. En effet Benveniste *et al.*, ont utilisé la base de données ENCODE, référencant les positions génomiques de nombreuses marques épigénétiques et facteurs de transcription, pour construire un modèle prédictif de la présence de certaines marques épigénétiques en fonction du motif de facteur de transcription présent (Benveniste et al., 2014).

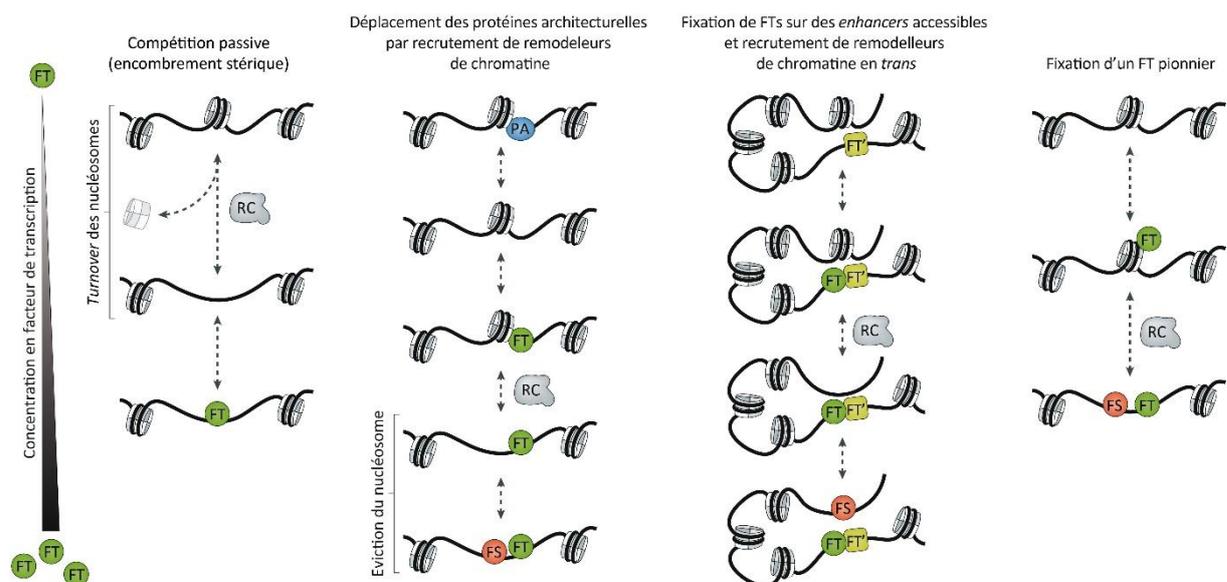


Figure 17 : Modèles d'accessibilité des facteurs de transcription à la chromatine

FT : facteur de transcription, RC : remodelleur de chromatine, PA : protéine architecturale de la chromatine, FS et FT' : facteurs de transcription secondaires. Adapté de Klemm et al., 2019.

Il existe des facteurs de transcription, appelés pionniers (ou « *pioneer transcription factor* »), qui peuvent interagir directement avec la chromatine fermée et créer une ouverture permettant la fixation d'autres facteurs de transcription par la suite (Balsalobre and Drouin, 2022; Voss and Hager, 2014). Le terme « pionnier » fait plus référence à une fonction que peut avoir un facteur de transcription, plutôt qu'à une famille spécifique de facteurs de transcription : par exemple, le facteur de transcription GATA4 a été montré comme étant

pionnier (He et al., 2014; Välimäki and Ruskoaho, 2020). Une fois fixés à la chromatine fermée, les facteurs de transcription pionniers peuvent recruter des partenaires protéiques, comme des cofacteurs modificateurs de marques d'histones (Mitsis et al., 2020), ou des cofacteurs remodeleurs de chromatine, pour induire l'éviction de nucléosomes et rendre la fixation d'autres facteurs de transcription possible (Figure 17) (Bevilacqua et al., 2014; Kathiriya et al., 2015). Ainsi, le complexe multiprotéique SWI/SNF est un complexe remodeleur de chromatine, dont la sous-unité spécifique du cœur, BAF60c, peut interagir avec les facteurs de transcription GATA4, NKX2-5 et TBX5 pour induire l'expression de gènes (Takeuchi and Bruneau, 2009). L'éviction d'un nucléosome réduit ainsi l'encombrement permettant la fixation de facteurs de transcription, qui une fois fixés, empêchent le repositionnement du nucléosome (Figure 17) (Klemm et al., 2019; Spitz and Furlong, 2012). Les facteurs de transcription pionniers peuvent également recruter des complexes répresseurs pour avoir une action opposée et garantir la stabilité de régions réprimées et compactes du génome. La fixation d'un facteur de transcription sur l'ADN va donc avoir des impacts variés sur la structuration de la chromatine environnante, dus notamment au recrutement de cofacteurs spécifiques.

1.4.A.d. Gene regulatory networks ou réseaux de régulation de gènes

Si un facteur de transcription se fixe sur une région régulatrice d'un gène codant un autre facteur de transcription et contrôle son expression, cette relation établit un réseau de régulation (ou *gene regulatory network*, GRN, en anglais) entre les facteurs de transcription (Kathiriya et al., 2015; Thompson et al., 2015). Ces réseaux d'activation et d'inhibition entre les facteurs de transcription sont spécifiques des types cellulaires et sont indispensables pour permettre l'expression coordonnée de gènes, notamment au cours du temps, dans le cas de processus développementaux ou de différenciation cellulaire (Neph et al., 2012; Potier et al., 2014a). L'étude de ces réseaux est donc importante, car des dérégulations vont entraîner le développement de processus pathologiques (Atkins et al., 2016). Dans le développement cardiaque, différents GRN ont été décrits de manière non exhaustive, mais il manque une vision globale reliant fonctionnellement l'ensemble de ces sous-réseaux identifiés (Bruneau, 2013; Kathiriya et al., 2021; Moore-Morris et al., 2018; Olson, 2006; Srivastava, 2006).

I.4.B. Les facteurs de transcription Iroquois homeobox

Les gènes Iroquois homeobox ou *IRX* ont été découverts pour la première fois chez la drosophile dans le cadre d'études des organes sensoriels (Dambly-Chaudière and Leyns, 1992), où trois gènes y ont été identifiés : *araucan*, *caupolican* et *mirror* (Gomez-Skarmeta et al., 1996; McNeill et al., 1997). Des mutations dans le locus iroquois (ou *Iro-C*) provoquent la perte des soies des régions latérales du thorax dorsal des drosophiles, ne laissant apparaître qu'une bande médiane de soies, rappelant les crêtes de cheveux des tribus Iroquoises d'Amérique dont est inspiré le nom du locus (Leyns et al., 1996). Des déficiences de ce locus ont aussi été associées à des défauts du développement cardiaque (Mirzoyan and Pandur, 2013). Des gènes iroquois orthologues ont été retrouvés dans différentes espèces comme le nématode (*Caenorhabditis elegans*), le zebrafish, le xénope ou le poulet (Garriock et al., 2001; Kerner et al., 2009; Ogura et al., 2001; Tan et al., 1999). Chez les mammifères, six gènes *IRX* sont retrouvés disposés dans deux clusters, formés par duplication d'un cluster ancestral de trois gènes : (1) le cluster *IrxA*, composé des gènes *IRX1*, *IRX2* et *IRX4*, situé sur le chromosome 5 chez l'humain et (2) le cluster *IrxB*, composé des gènes *IRX3*, *IRX5* et *IRX6*, situé sur le chromosome 16 chez l'humain (Figure 18) (Kerner et al., 2009; Kim et al., 2012).

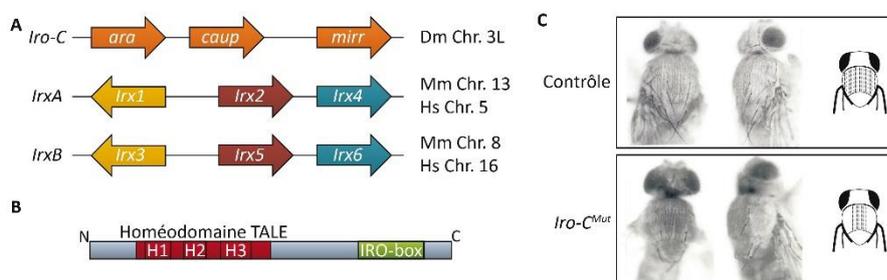


Figure 18 : Organisation génomique et structure des facteurs de transcription Iroquois homeobox

A. Organisation des gènes Iroquois sur l'ADN, en un seul cluster (*Iro-C*) de trois gènes chez la drosophile (Dm) et en deux clusters (*IrxA* et *IrxB*) chez la souris (Mm) et l'humain (Hs). **B.** Structure protéique commune entre les facteurs de transcription Iroquois. TALE : *Three Amino-acid Loop Extension*, H1-H3 : hélices alpha de l'homéodomaine. **C.** Organisation des soies du thorax dorsal chez les drosophiles contrôles et mutées dans le locus *Iro-C* (*Iro-C^{Mut}*). Adapté de Dambly-Chaudière and Leyns, 1992 et Kim et al., 2012.

Tous les membres de la famille des Iroquois homeobox possèdent une région bien décrite pour permettre l'interaction avec l'ADN, l'homéodomaine (Holland et al., 2007; Hu et al., 2018; Kim et al., 2012). L'homéodomaine des protéines *IRX* fait partie de la classe des homéodomains TALE (*Three Amino-acid Loop Extension*). Il est composé de trois hélices alpha et a une taille de 63 acides aminés, contre 60 acides aminés pour les autres classes d'homéodomains, du fait de la présence de trois acides aminés supplémentaires entre les

hélices 1 et 2 (Figure 18) (Bürglin, 1997). Plus précisément, les hélices 2 et 3 ont été caractérisées comme importantes pour la conformation 3D de l'homéodomaine et la fixation à l'ADN respectivement (Bürglin, 2011). Il n'existe pas de motif ADN consensus clair pour la fixation des protéines IRX, mais différentes études proposent des motifs associés aux séquences « ACA » et « TGT » (Anderson et al., 2012; Berger et al., 2008; Biloni et al., 2005; Noyes et al., 2008).

Une autre séquence, de 14 acides aminés, appelée *IRO-box* (Bürglin, 1997), est présente et conservée entre les différents membres de la famille IRX (Figure 18). L'*IRO-box* a été peu étudiée, mais deux études suggèrent qu'elle est impliquée dans les interactions protéine-protéine des IRX (Barrios and Campuzano, 2016; Cavodeassi et al., 2001). L'une suggère une interaction avec la cycline E dans le cycle cellulaire chez la drosophile, et l'autre suggère un rôle dans les interactions protéine-protéine, car sa séquence est très proche de celle de la partie centrale de séquences EGF (*Epidermal Growth Factor*) répétées du récepteur Notch.

Les facteurs de transcription IRX sont impliqués dans des processus développementaux de différents organes comme le système nerveux (Cavodeassi et al., 2001; Lecaudey et al., 2005), les poumons (Doi et al., 2011; Van Tuyl et al., 2006), la rétine (Cheng et al., 2005), les gonades (Jorgensen and Gao, 2005), les membres (Li et al., 2014a) ou le tissu adipeux (Bjune et al., 2020). Du fait de leur implication dans le cycle cellulaire, ils sont aussi associés à des processus cancéreux (Barrios et al., 2015).

Les six gènes Iroquois sont exprimés dans le cœur (Figure 19) et y ont des implications spécifiques et chevauchantes (Kim et al., 2012).

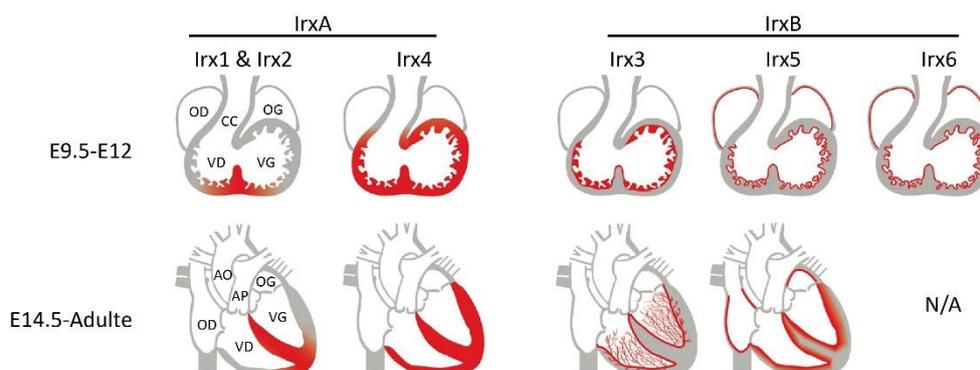


Figure 19 : Expression cardiaque des facteurs de transcription Iroquois

OD : oreillette droite, OG : oreillette gauche, CC : chambre de chasse, VD : ventricule droit, VG : ventricule gauche, AO : artère aorte, AP : artère pulmonaire. Adapté de Kim et al., 2012.

l.4.B.a. Cluster IrxA dans le cœur

IRX1

Le facteur de transcription *Ir1* a une expression cardiaque très similaire à celle d'*Ir2*, localisée à E10.5 chez la souris dans le myocarde trabéculé du septum interventriculaire. Entre E15.5 et E17.5, son expression est principalement retrouvée du côté gauche du septum, dans le faisceau de His et les branches du faisceau (Figure 19). Cette expression suggère qu'il pourrait avoir un rôle dans le développement du septum interventriculaire et du système de conduction ventriculaire (Christoffels et al., 2000). En effet, chez le varan (*Varanidae*), *Ir1* (comme *Ir2*) est retrouvé dans les septa partiels du ventricule, équivalents au septum interventriculaire des mammifères (Hanemaaijer et al., 2019), alors que l'orthologue d'*Ir1* dans le *zebrafish* (*Ir1b*) est exprimé dans le système de conduction en développement, même si le *zebrafish* ne possède pas de septum ventriculaire (Joseph, 2004). Des participations d'*Ir1* dans les mécanismes de remodelage et d'inhibition de la fibrose cardiaque, ont aussi été suggérés dans un modèle de rat insuffisant cardiaque (Zeng et al., 2019). Chez l'humain, bien que trois variants génétiques d'IRX1 aient été retrouvés chez des patients atteints de cardiopathies congénitales structurelles (avec des défauts de septation atriale ou ventriculaire, une transposition des grandes artères ou un retour veineux pulmonaire anormal, Guo et al., 2017), des informations sont encore manquantes sur ses implications au cours du développement cardiaque et dans le cœur adulte humain.

IRX2

L'expression cardiaque d'*Ir2* est concomitante à celle d'*Ir1* dans le septum interventriculaire (Kim et al., 2012). La souris *Ir2*-KO est viable et fertile, sans défaut morphologique du cœur ni de la fonction cardiovasculaire, suggérant l'établissement de mécanismes de compensation d'autres facteurs de transcription Iroquois pour la perte d'*Ir2* (Lebel et al., 2003). A l'inverse, la surexpression d'*Ir2* dans le cœur des souris *Ir4*-KO suggère qu'il exerce un effet partiellement compensateur pour la perte d'*Ir4* (Bruneau et al., 2001). Bien qu'il soit possible qu'un phénotype pathologique apparaisse avec l'âge chez la souris *Ir2*-KO (non investigué), l'absence d'anomalies cardiaques démontre qu'*Ir2* n'est pas indispensable au développement cardiaque.

IRX4

Irx4 est exprimé à partir de E7.5 dans le croissant cardiaque de souris (dans le premier et le second champ cardiaque), puis dans le tube linéaire cardiaque et lors du *looping*, où il est exclusivement exprimé dans les régions progénitrices des ventricules (Figure 19). L'expression d'*Irx4* est maintenue jusqu'au stade adulte et est toujours exclusive aux ventricules (Bruneau et al., 2000; Nelson et al., 2014). Il est d'ailleurs utilisé comme marqueur de sélection des cardiomyocytes ventriculaires dans des protocoles de différenciation cardiaque de cellules souches embryonnaires de souris (Ban et al., 2015; Nelson et al., 2016). La diminution d'expression d'*Irx4* dans les ventricules des souris *Nkx2-5-KO* et *Hand2-KO* suggère que ces deux facteurs de transcription participent à l'activation de son expression (Bruneau et al., 2000). *Irx4* permet la spécification des ventricules en inhibant l'expression de gènes des oreillettes et en activant des gènes des ventricules (Figure 20). Par exemple, *Irx4* active l'expression de la chaîne lourde de myosine ventriculaire (*Vmhc1*) et inhibe l'expression de la chaîne lourde de myosine atriale (*Amhc1*) dans les ventricules, chez le poulet (Bao et al., 1999; Wang et al., 2001).

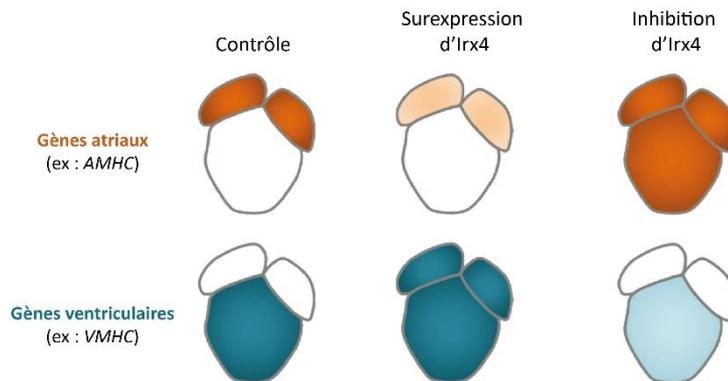


Figure 20 : Actions activatrices et inhibitrices d'*Irx4* dans le cœur

Impact d'une surexpression ou de l'inhibition d'*Irx4* sur l'expression des gènes atriaux (orange) et ventriculaires (bleu). Une couleur foncée indique une expression forte et inversement. Adapté de Kim et al., 2012.

Chez la souris, le KO d'*Irx4* n'induit pas de défauts structuraux du cœur au cours des stades embryonnaires (mécanismes de compensation suggérés notamment par *Irx2*), mais une expression ectopique de gènes atriaux dans les ventricules est observée, comme *Nppa* par exemple (Bruneau et al., 2001). Après 6 mois de vie, la souris *Irx4-KO* développe spontanément une cardiomyopathie avec une diminution de la fraction d'éjection et une hypertrophie cardiaque (Kim et al., 2012). Ces données chez la souris suggèrent qu'*Irx4* n'est pas indispensable au développement ventriculaire, mais y est nécessaire à l'établissement

d'un programme spécifique de gènes (Bruneau et al., 2001). Chez l'humain, deux variants génétiques associés à une perte de fonction d'IRX4 ont été identifiés chez des patients présentant des défauts de septation ventriculaire, confirmant son importance dans le cœur (Cheng et al., 2011).

1.4.B.b. Cluster *IrxB* dans le cœur

IRX3

Irx3 est exprimé à partir de E9.5 chez la souris, dans le myocarde ventriculaire trabéculé, mais pas dans le myocarde compact, ni les cellules endocardiques longeant les trabécules (Figure 19) (Christoffels et al., 2000). *Irx3* est nécessaire au développement du système de conduction ventriculaire, car il y inhibe de manière directe l'expression de la Cx43 (associée à la conduction plus lente de l'influx électrique) en ayant un rôle antagoniste de *Nkx2-5*. En parallèle, il active l'expression de la Cx40 (associée à la conduction rapide), de manière indirecte, probablement en inhibant l'expression d'un répresseur (Figure 21) (Zhang et al., 2011). La souris *Irx3*-KO est viable et sans défaut morphologique du cœur, mais après deux semaines après la naissance : (1) un allongement du QRS est observé chez ces souris, reflétant un ralentissement de la conduction ventriculaire, (2) ainsi qu'un bloc de la branche droite du faisceau de His, qui induit un défaut d'activation du ventricule droit (Zhang et al., 2011). La Cx43 présente une expression ectopique dans les branches du faisceau de His, du fait de l'absence d'*Irx3*, induisant une réduction de la vitesse de conduction par rapport aux

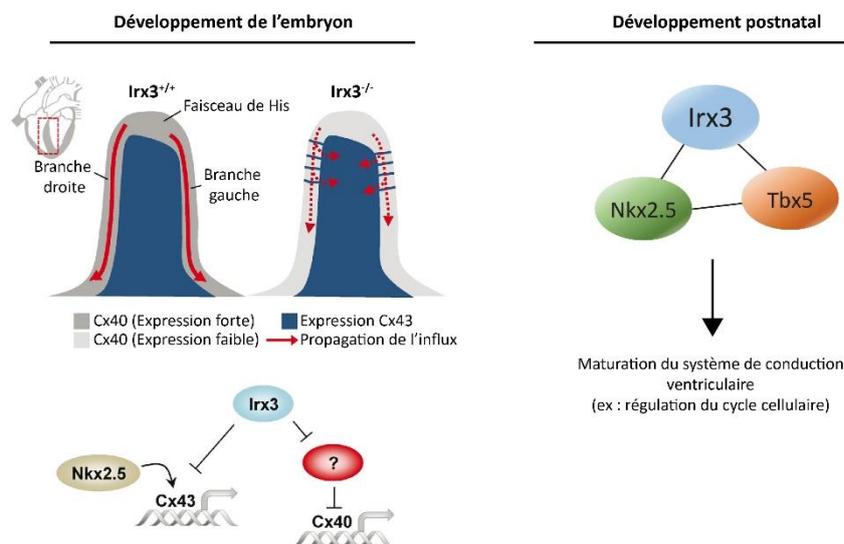


Figure 21 : Rôles d'*Irx3* dans l'établissement et la maturation du système de conduction ventriculaire

Le KO d'*Irx3* (*Irx3*^{-/-}) chez la souris provoque une expression ectopique de Cx43, ralentissant la conduction ventriculaire dans les branches du faisceau de His. Le complexe de facteur de transcription *Irx3*/*Nkx2-5*/*Tbx5* est impliqué dans la maturation postnatale du système de conduction ventriculaire. Adapté de Kim et al., 2012, 2016.

souris sauvages. Ces défauts électriques chez la souris *Irx3*-KO induisent un risque accru de développement d'arythmies (Zhang et al., 2011). La souris *Irx3*-KO présente des niveaux d'expression diminués de *Cx40*, tout comme la souris *Gja5*-KO, et pourtant un phénotype électrique cardiaque moins sévère est retrouvé chez la *Gja5*-KO, suggérant qu'*Irx3* a d'autres effets sur le système de conduction ventriculaire que la régulation des *Cx*. En effet, (1) le KO d'*Irx3* a aussi un impact sur le système His-Purkinje qui se détériore progressivement après la naissance, suggérant un rôle d'*Irx3* dans le maintien de l'intégrité du système de conduction ventriculaire. De plus, (2) *Irx3* coopère avec *Nkx2-5* et *Tbx5* dans la régulation du cycle cellulaire pour permettre une maturation du système de conduction ventriculaire (Figure 21) (Kim et al., 2016). *Irx3* est donc impliqué dans le développement, la maturation et le maintien du système de conduction ventriculaire chez la souris. Chez l'humain, deux variants génétiques d'*IRX3* ont été identifiés chez des patients avec des fibrillations ventriculaires idiopathiques, confirmant le lien observé chez la souris entre des altérations d'*IRX3* et le développement d'arythmies cardiaques (Koizumi et al., 2016). Deux études ont également suggéré une implication d'*Irx3* dans un syndrome cardiaque complexe associé à des risques d'arythmies ventriculaires et de mort subite, le syndrome de Brugada (Barc et al., 2022; Kimura et al., 2016).

IRX5

Chez la souris, *Irx5* a une première expression détectée à E9, dans la partie ventrale du tube cardiaque en *looping*. Puis à E9.5, *Irx5* est exprimé dans l'endocarde longeant les oreillettes et ventricules en formation, mais pas celui du canal atrioventriculaire ou la chambre de chasse (Figure 19) (Christoffels et al., 2000). Dans le cœur adulte, *Irx5* est exprimé au sein du myocarde ventriculaire, dans un gradient décroissant au sein de la paroi ventriculaire : du myocarde sous-endocardique vers le myocarde sous-épicardique, ainsi que dans le faisceau de His et les branches du faisceau (Figure 22A) (Christoffels et al., 2000; Gaborit et al., 2012).

Une cible d'*Irx5* a été identifiée chez la souris, le gène *Kcnd2*, codant la sous-unité α du canal potassique *Kv4.2* responsable du courant I_{to} (rapide) repolarisant. Au sein de la paroi ventriculaire *Kcnd2* a lui aussi une expression en gradient mais qui est inverse par rapport à celle d'*Irx5*. En accord avec ces données, il a été montré qu'*Irx5* inhibe l'expression de *Kcnd2* à l'aide d'un cofacteur, *mBop*, qui recrute à son tour l'enzyme HDAC, pour réprimer épigénétiquement la région du gène *Kcnd2* et empêcher son expression (Costantini et al., 2005). *Irx5* peut agir comme répresseur (avec *mBop*) ou activateur en fonction du contexte,

car dans un modèle de surexpression non cardiomyocytaire, *Irx5* active l'expression de *Kcnd2* (Costantini et al., 2005). Une autre étude suggère que ce mécanisme d'inhibition d'*Irx5* passe par l'interaction avec *Irx4* et non avec *mBop* (Figure 22B) (He et al., 2009). Ces résultats différents indiquent la nécessité d'études supplémentaires pour mieux comprendre ce processus. La souris *Irx5*-KO est viable et sans défaut structurel majeur, mais présente une susceptibilité accrue aux arythmies. Cette susceptibilité a été attribuée à la perte du gradient transmurale du potentiel d'action dans le ventricule qui est associée à la survenue de tachycardies ventriculaires. La perte d'*Irx5* induit une expression ectopique de *Kcnd2* dans le myocarde sous-endocardique induisant un raccourcissement de la durée du potentiel d'action dans cette région, provoquant des défauts de la repolarisation ventriculaire (Figure 22A) (Costantini et al., 2005). En plus de ce défaut électrique, les souris *Irx5*-KO ont une contractilité cardiaque réduite, dont l'origine serait associée à l'activation transcriptionnelle anormale de *Kcnd2*, car le double KO *Irx5*/*Kv4.2* restaure la fonction contractile (Kim et al., 2022). Chez l'humain, l'expression en gradient d'*IRX5* au sein de la paroi ventriculaire est également retrouvée, tout comme le gradient inverse du courant I_{to} . Par contre, ce courant I_{to} est chez l'humain, principalement médié par le canal *Kv4.3* (contrairement à *Kv4.2* chez la souris), en lien avec la sous-unité accessoire *KCHIP2* (Gaborit et al., 2010). Du fait de son expression en gradient (inversé par rapport à *IRX5*) dans de la paroi ventriculaire, *KCHIP2* (codé par *KCNIP2*) est supposé être responsable du gradient transmurale d' I_{to} chez l'humain.

Une précédente étude de l'équipe a par ailleurs permis de montrer un rôle clé pour *IRX5* dans la régulation de la dépolarisation ventriculaire et de la conduction électrique

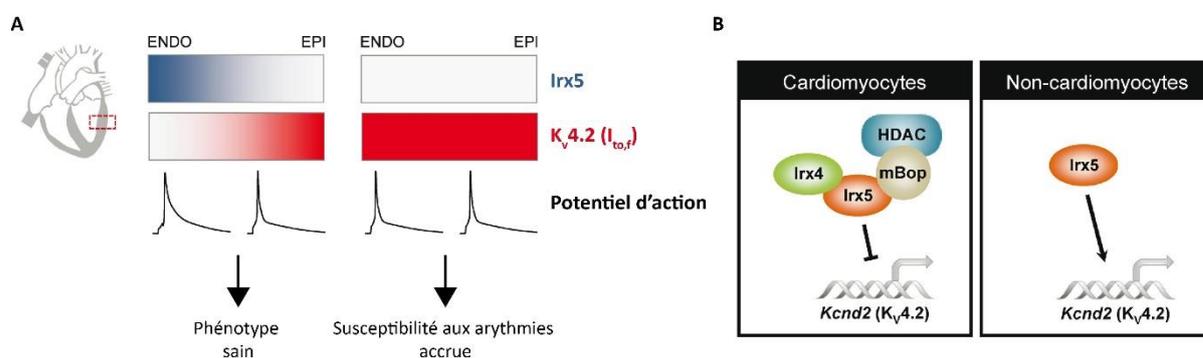


Figure 22 : Rôle d'*Irx5* dans le maintien de l'hétérogénéité transmurale du potentiel d'action

A. Expression du facteur de transcription *Irx5* et du canal potassique *Kv4.2* (responsable du courant $I_{to,f}$) au sein de la paroi ventriculaire, entre les myocards sous-endocardique (ENDO) et sous-épicaire (EPI), ainsi que les potentiels d'action et phénotypes électriques associés chez la souris. **B.** Modèle d'action d'*Irx5* dans les cardiomyocytes avec ses partenaires spécifiques et dans des cellules non-cardiomyocytes, sur l'expression du gène *Kcnd2*. Adapté de Kim et al., 2012.

cardiaque chez l'humain, à travers notamment l'activation de *SCN5A* (codant le canal Nav1.5) par *IRX5*, en cohérence avec le gradient transmural d'expression de *SCN5A* qui est similaire à celui d'*IRX5* (Al Sayed et al., 2021a).

De manière intéressante, *Irx3* et *Irx5* sont exprimés de manière concomitante, potentiellement grâce à des enhanceurs communs proches (Calle-Mustienes et al., 2005; Chua et al., 2022; Tena et al., 2011) et ont des implications en coopération et redondance fonctionnelle dans différents organes (Fu et al., 2018; Li et al., 2014a; Son et al., 2021). La souris *Irx5*-KO étant viable, fertile et ne présentant pas de défauts morphologiques cardiaques majeurs, montre qu'*Irx5* n'est pas indispensable au développement et au maintien de l'intégrité cardiaque, et suggère que son rôle dans ces fonctions est masqué par la redondance fonctionnelle avec d'autres facteurs de transcription Iroquois (Gaborit et al., 2012). De la même manière, la souris *Irx3*-KO est viable, fertile et sans défaut morphologique cardiaque majeur, suggérant également un effet compensateur d'autres facteurs de transcription Iroquois (Zhang et al., 2011). L'étude du double KO pour *IRX3* et *IRX5* chez la souris, létale au stade embryonnaire à E14.5 avec des défauts cardiaques structuraux majeurs, a permis de démontrer qu'effectivement ces deux facteurs de transcription ont des fonctions redondantes et même coopératives (Gaborit et al., 2012).

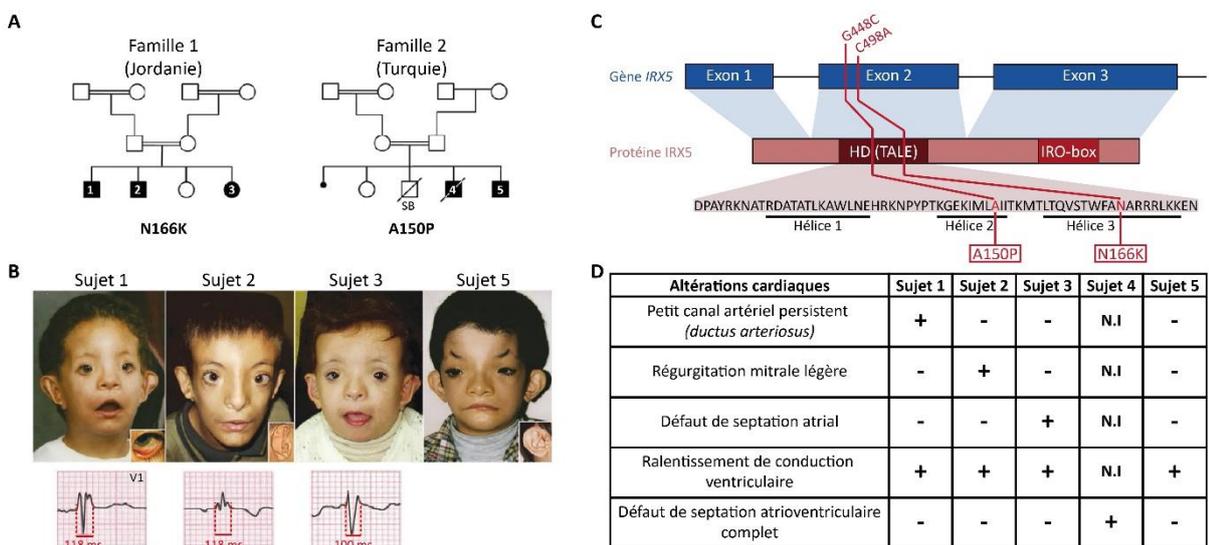


Figure 23 : Patients atteints du syndrome d'Hamamy et présentant des mutations perte-de-fonction d'*IRX5*

A-B. Arbres généalogiques, photos et électrocardiogrammes des patients Hamamy des deux familles décrites dans Bonnard et al., 2012. **C.** Localisation génomique et protéique des mutations Hamamy A150P et N166K. HD : homéodomaine TALE (*Three Amino-acid Loop Extension*). **D.** Altérations cardiaques retrouvées chez les patients Hamamy décrits. N.I : non investigué. Adapté de Bonnard et al., 2012 et Bürglin, 1997.

Chez l'humain, deux mutations localisées dans l'homéodomaine d'IRX5 (A150P et N166K) et provoquant une perte de fonction de la protéine, ont été identifiées pour la première fois en 2012, chez des patients de deux familles différentes (Figure 23A-C) (Bonnard et al., 2012). Les patients de ces familles sont atteints du syndrome d'Hamamy, décrit pour la première fois en 2007 (Hamamy et al., 2007). Ils présentent un phénotype très sévère incluant des altérations multiorganes (défauts craniofaciaux, des os et des membres, ...), dont des défauts cardiaques : principalement un ralentissement de la conduction ventriculaire, commun à tous ces patients (Figure 23D). Le phénotype des patients est plus sévère que le phénotype des souris *Irx5*-KO, suggérant que les mutations Hamamy pourraient avoir des effets dominants négatifs. De plus, le ralentissement de conduction ventriculaire observé chez les patients rappelle le phénotype des souris *Irx3*-KO et *IRX3/IRX5*-KO, suggérant que l'impact des mutations Hamamy pourrait passer par des altérations du couple *IRX5/IRX3*.

IRX6

Irx6 a une expression cardiaque plus faible que celle des autres *Irx* et est localisé dans les mêmes régions qu'*IRX5*, à partir de E10.5 chez la souris (Figure 19) (Mummenhoff et al., 2001). Sa fonction n'a pas encore été étudiée dans le cœur (Hu et al., 2018).

I.4.C. Stratégies d'étude des facteurs de transcription cardiaques

Dans cette partie sont énoncés les principaux modèles utilisés pour l'étude des facteurs de transcription cardiaques, avec une attention particulière pour les modèles basés sur les cellules souches. Les principales techniques de référence et leur intérêt dans l'étude des facteurs de transcription cardiaques sont ensuite abordés.

I.4.C.a. Modèles animaux

Différents modèles animaux sont utilisés pour l'étude des facteurs de transcription cardiaques et prennent généralement place dans le contexte plus large de l'étude du développement cardiaque (Taylor et al., 2021; Zaffran and Frasch, 2002). Ces modèles permettent de pallier le manque d'échantillons du développement cardiaque humain, qui sont rares et difficiles à obtenir.

Les petits animaux ont été et sont toujours utilisés pour l'étude des facteurs de transcription, comme la drosophile (*Drosophila melanogaster*) qui permet de modéliser le développement cardiaque (Piazza and Wessells, 2011). Ce modèle a permis par exemple l'identification de la famille des facteurs de transcription Iroquois (Leyns et al., 1996). Le

poisson-zèbre ou *zebrafish* (*Danio rerio*), ainsi que le xénope (*Xenopus laevis*) ont été notamment utilisés pour l'étude des familles de facteurs transcription GATA, TBX et NKX (Garfinkel and Khokha, 2017; Heicklen-Klein et al., 2005). Le poulet (*Gallus gallus*) a également été montré comme étant un bon modèle d'étude du développement cardiaque (Wittig and Münsterberg, 2020). Enfin, la souris (*Mus musculus*) correspond au principal modèle animal utilisé pour l'étude du développement cardiaque et des facteurs de transcription (Akerberg et al., 2019; Krishnan et al., 2014). Le développement cardiaque et plus particulièrement les réseaux de régulation moléculaires, sont bien conservés entre ces différents modèles et l'humain (Xia et al., 2020). Ils correspondent à des modèles intégrés, où le cœur peut être étudié au sein d'organismes entiers et non de manière isolée et artificielle. Ces modèles sont facilement modifiables génétiquement et permettent également la réalisation de techniques d'analyse globale à haut débit ou « omiques ». Cependant, notamment pour le modèle de souris, il existe des différences dans le développement et la physiologie cardiaque par rapport à l'humain, suggérant que des mécanismes spécifiques des espèces s'établissent, nécessitant ainsi l'utilisation d'autres modèles pour s'approcher du contexte humain (Anzai et al., 2020; Rossant and Tam, 2017).

1.4.C.b. Cardiomyocytes dérivés de cellules souches humaines

Afin d'étudier le développement cardiaque et les facteurs de transcription dans un contexte humain, des modèles de différenciation cardiaque de cellules souches ont été mis au point.

Au cours du développement pré-cardiaque humain, environ 5 jours après la fécondation, des cellules pluripotentes apparaissent au sein de la masse cellulaire interne du blastocyste ([voir partie 1.2.A.a Pluripotence](#)). En 1998 et 2000, deux équipes ont extrait la masse cellulaire interne d'embryons humains après fécondation *in vitro*, pour isoler des lignées de cellules souches embryonnaires pluripotentes, hESCs (Reubinoff et al., 2000; Thomson et al., 1998; Vazin and Freed, 2010). Ces lignées de cellules souches pluripotentes ont des capacités d'autorenouveau et de différenciation dans l'ensemble des types cellulaires dérivés des trois feuillets embryonnaires, et ont été notamment utilisé pour générer des cellules cardiaques (Kehat et al., 2001; Mummery et al., 2003). Ce modèle de différenciation, qui est une source illimitée de cellules cardiaques humaines et notamment de cardiomyocytes (Acimovic et al., 2014), a permis d'ouvrir de nombreuses perspectives d'utilisation dans le contexte de l'étude du développement cardiaque (et du rôle des facteurs

de transcription cardiaques), de la thérapie cellulaire cardiaque ou d'études pharmacologiques (Anderson et al., 2018; Liu et al., 2018b; Williams et al., 2021). Cependant, la rareté et la difficulté d'obtention d'embryons humains, ainsi que les restrictions éthiques associées, sont des limites à l'étude de lignées hESCs issus de patients dans le cadre de stratégies de médecine personnalisée. En effet, bien que des stratégies d'édition génétique soient réalisables sur les hESCs, pour créer des modèles isogéniques (qui ont pour seule différence génétique la modification artificiellement induite), elles ne permettent pas d'étudier les cas où la causalité de la pathologie n'a pas été associée à un variant génétique précis.

En 2006, la première génération de cellules souches pluripotentes induites, ou iPS, a été obtenue par surexpression des facteurs de transcription Oct4, Sox2, Klf4 et c-Myc dans des fibroblastes de souris (Takahashi and Yamanaka, 2006). Un an après, deux équipes différentes ont montré la possibilité de reprogrammer des fibroblastes humains en cellules iPS (Takahashi et al., 2007; Yu et al., 2007). Ces cellules iPS humaines sont très similaires aux hESCs, notamment du point de vue de leur morphologie, de leur capacité de prolifération et de pluripotence ou de leur profil global d'expression de gènes (Choi et al., 2015; Parrotta et al., 2020). Toutefois, différentes études ont montré l'existence de légères différences entre ces deux types de cellules souches notamment en terme de contenu en acide nucléique ou d'expression de sous-groupes de gènes (Chin et al., 2009; Parrotta et al., 2017, 2019). Les cellules iPS présentent un avantage considérable d'accessibilité comparées aux hESCs, car elles peuvent être reprogrammées à partir de cellules provenant d'échantillons biologiques facilement accessibles, notamment des cellules de sang, de peau ou d'urine (Afzal et al., 2017; Agu et al., 2015; Jouni et al., 2015; Li et al., 2016). Par cette possibilité d'obtention facilitée, sans dégradation d'embryons humains, l'utilisation des iPS humaines est ainsi sujette à des contraintes éthiques moins lourdes comparées à celles des hESCs.

Tout comme pour les hESCs, les iPS humaines ont ensuite été utilisées comme source illimitée pour la génération de cellules cardiaques, notamment des cardiomyocytes (iPS-CMs). Il existe différents protocoles, principalement basés sur l'induction de la différenciation des iPS en cellules du mésoderme puis en progéniteurs cardiaques, par l'utilisation de petites molécules (ex : CHIR99021) ou de cytokines (ex : Activine A, BMP4, ...). Ces différents protocoles vont généralement aboutir à la production d'une population mixte de cellules, (1)

à la fois des cardiomyocytes et des cellules non cardiomyocytaires (de type fibroblaste par exemple) et (2) différents sous-types de cardiomyocytes (atriaux, nodaux et ventriculaires, Protze et al., 2019). Des étapes de déplétion en glucose ont été mises au point afin d'enrichir en cardiomyocytes les populations en fin de différenciation. Des protocoles visant à générer plus spécifiquement des cardiomyocytes ventriculaires (généralement le sous-type majoritaire obtenu), nodaux (par régulation de la balance BMP/acide rétinoïque et la surexpression de TBX3) ou atriaux (par l'ajout d'acide rétinoïque) ont également été mis au point (Bekhite and Schulze, 2021; Protze et al., 2019; Swedlund and Lescroart, 2020).

Les modèles dérivés d'iPS permettent l'étude de pathologies et de mutations génétiques cardiaques par reprogrammation de cellules de patients en cellules pluripotentes, qui conservent l'ensemble de leurs spécificités génétiques (Al Sayed et al., 2021b; Yoshida and Yamanaka, 2017). Il est également possible d'éditer génétiquement les cellules iPS, par CRISPR-Cas9 par exemple (Chun et al., 2018). Dans le cadre de l'étude du rôle de facteurs de transcription dans la différenciation cardiaque, la technologie CRISPR-Cas9 a été utilisée pour l'induction de KO (homo ou hétérozygotes) ou pour l'insertion de mutations génétiques de ces gènes observées chez des patients (Ang et al., 2016; Kathiriya et al., 2021). Des études pharmacologiques, ainsi que le développement de nouvelles stratégies de thérapie cellulaire cardiaque sont réalisables avec le modèle de cellules iPS (Lemoine and Christ, 2021; Magdy et al., 2018; Parrotta et al., 2020).

Pour l'étude des facteurs de transcription cardiaques, il est nécessaire de tenir compte des deux principales limites de ces modèles de différenciation cardiaque de cellules iPS : (1) le phénotype immature des cardiomyocytes obtenus en fin de différenciation (Figure 24) et (2) la réalisation de cultures en deux dimensions (2D) qui est éloignée du format 3D de l'organe. Comparés à des cardiomyocytes adultes, les cardiomyocytes dérivés d'iPS n'atteignent pas le même niveau de maturité et semblent être plus similaires à des cardiomyocytes fœtaux (Wu et al., 2021a). En effet, les cardiomyocytes dérivés d'iPS ont une forme arrondie (contrairement aux cardiomyocytes adultes qui ont une forme en bâtonnet), sont généralement mononucléés, n'ont pas de tubules T, ont un profil électrophysiologique fœtal (avec des contractions spontanées), présentent un métabolisme énergétique majoritairement glycolytique (et non associé aux acides gras) et enfin, ont une ultrastructure myofibrillaire désorganisée (Bekhite and Schulze, 2021; Korbassi et al., 2020; Wu et al., 2021a).

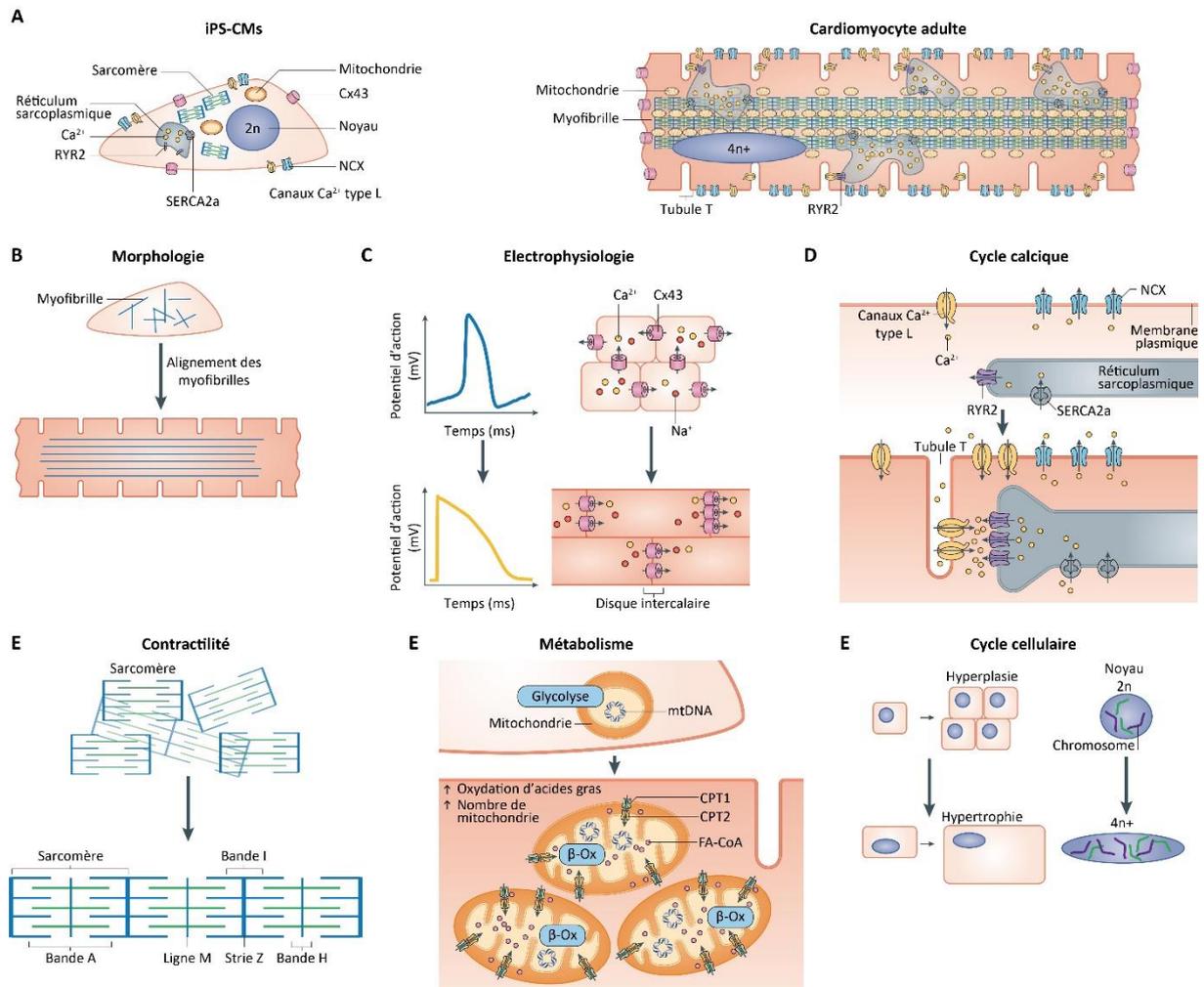


Figure 24 : Différences entre les iPS-CMs et les cardiomyocytes adultes

Différences d'organisation générale (A), morphologiques (B), électrophysiologiques (C), du cycle calcique (D), de la contractilité, métaboliques et du cycle cellulaire, entre les iPS-CM et les cardiomyocytes adultes. RYR2 : récepteur à la ryanodine de type 2, Cx43 : connexine 43, NCX : échangeur Na⁺/Ca²⁺, SERCA2a : pompe calcique du réticulum sarcoplasmique, mtDNA : ADN mitochondrial, β-Ox : bêta-oxydation des acides gras, CPT1-CPT2 : carnitine palmitoyltransferases 1 et 2, FA-CoA : fatty-acyl-CoA esters, n : ploïdie. Adapté de Karbassi et al., 2020.

La différenciation cardiaque des iPS en 2D correspond aussi à une limite du modèle, car cette situation est peu physiologique et ne rend pas compte de l'architecture tissulaire, des communications entre les différents types cellulaires et des contraintes physiques (notamment du flux sanguin) auxquels les cardiomyocytes sont normalement soumis pendant le développement cardiaque (Ahmed et al., 2020; Liu et al., 2022). Ces limites peuvent avoir un impact sur les mécanismes moléculaires observés et sont donc à prendre en compte dans les observations faites au cours et en fin de la différenciation cardiaque des iPS.

Différentes stratégies ont été mises en place pour tenter de faire progresser les cardiomyocytes dérivés d'iPS vers un stade de différenciation plus proche des cardiomyocytes adultes : la culture longue des cardiomyocytes, la stimulation électrique, l'application d'un

stress mécanique ou encore l'ajout d'hormone thyroïdienne (T3) dans le milieu de culture (Ahmed et al., 2020; Wu et al., 2021a) ont été testés. Bien que ces stratégies apportent chacune des améliorations phénotypiques, elles n'aboutissent pas à la production de cardiomyocytes de type adulte. La génération de modèles de différenciation 3D, en plein essor actuellement, permettrait de maturer le phénotype des cardiomyocytes en fin de différenciation et également de s'affranchir des limites liées à la culture 2D (Arslan et al., 2022; Liu et al., 2022). Les différents modèles 3D tels que les *engineered heart tissue* (Ronaldson-Bouchard et al., 2018), les organoïdes cardiaques ou cardioïdes (Cho et al., 2022), ou encore les modèles de « heart-on-chip » (Andrysiak et al., 2021) correspondent à des modèles plus intégrés, constitués de plusieurs types cellulaires et pourraient constituer des outils d'étude pertinents du développement et des facteurs de transcription cardiaques.

1.4.C.c. Techniques d'analyse

Création de modèles par édition génétique

Dans le but d'étudier le rôle de facteurs de transcription cardiaques, notamment au cours du développement, la stratégie classique est la génération de modèles KO (homo ou hétérozygote) pour pouvoir identifier les gènes et voies biologiques impactés par la perte spécifique du facteur de transcription délété (Kathiriya et al., 2015, 2021). Des insertions génétiques ou *knock-in* (KI) sont aussi utilisées, notamment pour la génération de modèles

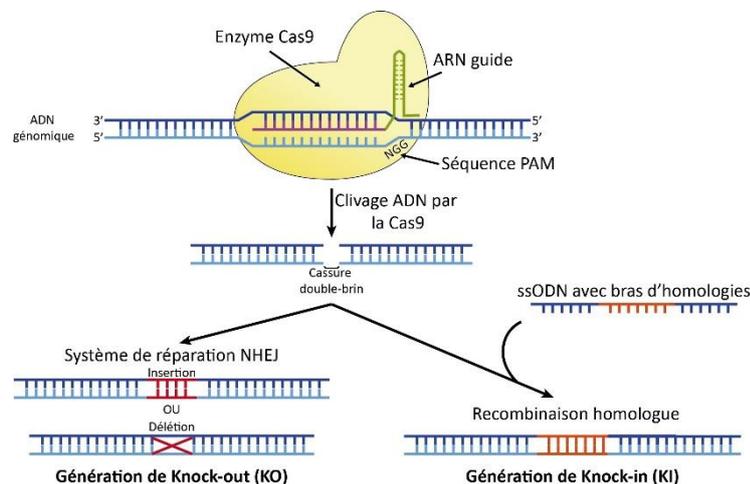


Figure 25 : Génération de KO et KI par la technologie CRISPR-Cas9

La fixation spécifique de l'enzyme Cas9 se fait à l'aide d'un ARN guide, ciblant précisément une zone proche d'une séquence PAM (*Protospacer Adjacent Motif*), généralement de type « NGG ». La PAM va être reconnue par la Cas9 pour induire une coupure double brin. Le système de réparation va tenter de réparer la coupure par NHEJ (*Non-Homologous End-Joining*) mais peut faire des erreurs provoquant par exemple, un changement de cadre de lecture d'un gène dans le cas d'un KO. L'apport d'un fragment d'ADN simple brin (ssODN, *single-stranded oligodeoxynucleotides*) contenant des bras d'homologies entourant une séquence d'intérêt, peut provoquer une recombinaison homologue au site de coupure, induisant l'insertion d'une séquence d'intérêt dans le génome (KI). Adapté de Tian et al., 2017.

isogéniques associés à des mutations génétiques spécifiques de patients ou pour l'ajout d'une séquence tag à un facteur de transcription (Ang et al., 2016). La technique d'édition génétique la plus utilisée actuellement est la technologie CRISPR-Cas9 qui permet le ciblage de régions spécifiques du génome, par l'intermédiaire d'un ARN guide, puis le clivage de l'ADN, grâce à l'enzyme Cas9, qui va reconnaître une séquence spécifique sur l'ADN, appelée séquence PAM (Figure 25) (*Protospacer Adjacent Motif*, De Masi et al., 2020; Tian et al., 2017). La principale limite au CRISPR-Cas9, comme pour les autres techniques d'édition du génome, est liée au risque d'induire des modifications supplémentaires dans des régions non spécifiques (ou *off-target* en anglais ; Fu et al., 2013).

Analyse de l'impact des facteurs de transcription sur le transcriptome

La fonction principale des facteurs de transcription est la régulation de l'expression des gènes. Ainsi, quantifier les niveaux d'expression génique dans des modèles adaptés est nécessaire pour identifier le rôle des facteurs de transcription. L'analyse de l'expression des gènes est principalement conduite par l'étude du transcriptome, qui peut être réalisée dans des modèles KO pour des facteurs de transcription par exemple, afin d'identifier des gènes dont l'expression est impactée par les facteurs de transcription étudiés. Actuellement, il existe différentes techniques d'analyse globale du transcriptome qui sont principalement dérivées du RNA-Seq (Lowe et al., 2017). Le RNA-Seq combine des étapes de fragmentation, rétrotranscription et séquençage d'ARN suivi de l'alignement des fragments séquencés (ou *reads*) sur un génome de référence, pour générer de manière bioinformatique une quantification de chaque ARN de l'échantillon. Il existe (1) des techniques transcriptomiques dites en « *bulk* » qui correspondent à la quantification des ARN issus d'un échantillon composé de plusieurs milliers de cellules (et parfois plusieurs type cellulaires) et (2) des techniques en cellules uniques ou « *single-cell* », qui permettent la quantification des ARN de chaque cellule contenue dans un échantillon de plusieurs centaines ou milliers de cellules (Figure 26A) (Asp et al., 2019; GTEx Consortium, 2013). Ces deux types de techniques ont leurs avantages et inconvénients spécifiques : le *bulk* permet la détection et l'étude de variations d'expression d'un grand nombre de gènes, mais constitue un signal moyenné d'expression entre différents types cellulaires (Altschuler and Wu, 2010; Rassek and Izykowska, 2020) ; le *single-cell* permet l'étude de l'expression de gènes dans différents types cellulaires, mais à cause du phénomène d'expression des gènes en salve au sein d'une cellule (voir partie I.4.A.b *Modèle intégré*), il est généralement retrouvé un grand nombre de niveaux d'expression de gènes à 0, pouvant poser

problème dans les analyses bioinformatiques par la suite (Lähnemann et al., 2020). Ces deux types d'étude transcriptomique peuvent donc être complémentaires pour l'analyse de l'impact des facteurs de transcription cardiaques.

Analyse des sites de fixation des facteurs de transcription sur le génome

Pour avoir un effet sur la transcription de certains gènes, les facteurs de transcription vont se fixer dans des régions spécifiques du génome. L'investigation de ces régions permet d'identifier les gènes qui sont des cibles directes de facteurs de transcription spécifiques et de mieux comprendre leur mode de fonctionnement (Anderson et al., 2018; Ang et al., 2016). La technique de référence pour l'identification des positions génomiques occupées par les facteurs de transcription est le ChIP-Seq (*Chromatin Immunoprecipitation Sequencing*). Le ChIP-Seq consiste : (1) à fragmenter la chromatine, puis (2) immunoprécipiter à l'aide d'anticorps spécifiques d'un facteur de transcription les complexes facteur de transcription-chromatine et (3) séquencer les fragments d'ADN obtenus, avant (4) l'analyse bioinformatique et alignement sur un génome de référence, pour obtenir des pics de *reads* correspondant aux régions ADN fixées par le facteur de transcription étudié (Figure 26B) (Johnson et al., 2007). Les principaux avantages de cette technique sont de pouvoir identifier tous les sites de fixation de facteurs de transcription sur le génome et aussi leurs motifs de fixation, dont l'identification est utile dans l'étude des GRN (Potier et al., 2014b). La qualité des résultats obtenus est sous-tendue par la qualité de l'anticorps utilisé, en effet, la principale limite du ChIP-Seq est l'utilisation d'anticorps qui doivent être très spécifiques des facteurs de transcription d'intérêts pour identifier un maximum de régions génomiques occupées, sans faux-positifs ou faux-négatifs (Park, 2009).

Afin de valider biologiquement la fixation d'un facteur de transcription sur un site précis du génome et d'analyser l'impact biologique de cette fixation, les techniques d'essais de luciférase dans des lignées cellulaires (ex : HEK293) sont couramment utilisées (Ávila-Flores et al., 2019). Le gène codant l'enzyme luciférase, qui va générer un signal de bioluminescence quantifiable, est sous contrôle d'une région d'intérêt de l'ADN. La transfection de lignées cellulaires avec le plasmide luciférase et des plasmides d'expression de facteurs de transcription permet d'identifier l'impact fonctionnel de ces facteurs de transcription sur la région ADN d'intérêt. Cette technique permet une validation biologique, mais offre une vision incomplète du phénomène de régulation, car les facteurs de transcription sont étudiés dans un contexte de surexpression sans la présence de tous leurs cofacteurs, sur une portion d'ADN

réduite, qui s'affranchit de l'impact potentiel d'autres éléments *cis* régulateurs du génome (Hirschenberger et al., 2021).

Analyses complémentaires du contexte épigénétique

Afin de mieux comprendre dans quel contexte les facteurs de transcription se fixent sur l'ADN et l'impact qu'ils peuvent avoir sur le paysage épigénétique, des expériences de CHIP-Seq appliquées aux marques épigénétiques sont possibles (Figure 26B) (Ernst and Kellis, 2017). La technique est similaire à celle appliquée sur les facteurs de transcription, mais cible les régions génomiques où les marques sont déposées, permettant ainsi d'identifier des éléments *cis* régulateurs dans des régions d'intérêts et de caractériser leur état d'ouverture et d'activité (Tableau 2). De plus, pour les marques d'histones il existe des anticorps très spécifiques et généralement validés pour le CHIP-Seq, permettant de pallier la principale limite du CHIP-Seq de facteur de transcription. Cette spécificité d'anticorps rend également possible

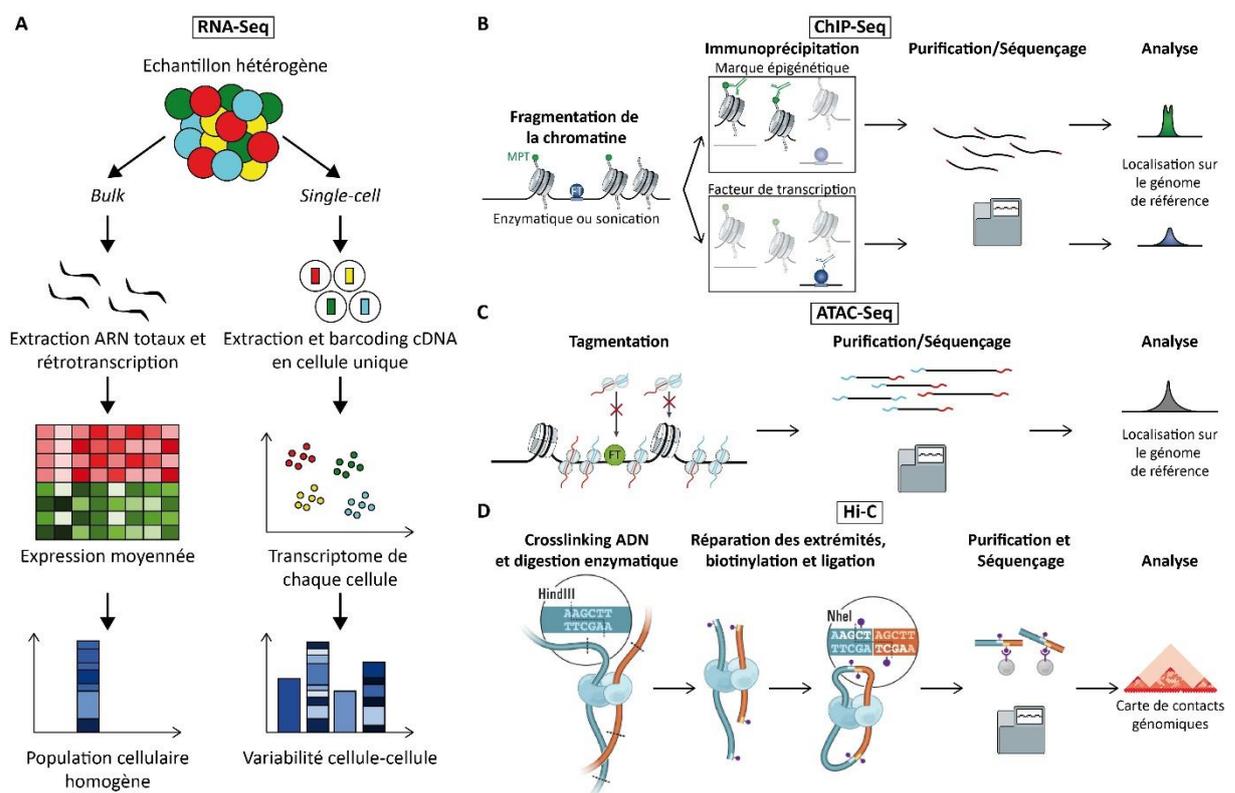


Figure 26 : Techniques « omiques » pour l'étude de facteurs de transcription

A. Comparaison des deux modalités d'analyse transcriptomique par RNA-Seq : en *bulk* ou en *single-cell*. **B.** Principales étapes d'analyse de CHIP-Seq. Par l'utilisation d'anticorps spécifiques, toutes les positions génomiques présentant certaines marques épigénétiques ou étant occupées par un facteur de transcription, peuvent être identifiées. Ces régions sont représentées sous la forme de pics de *reads* (fragments ADN séquencés). **C.** Analyse des régions génomiques ouvertes par ATAC-Seq, identifiées en fin d'analyse sous la forme de pics de *reads* également. **D.** Analyse des contacts physiques de chromatine par Hi-C. Adapté de Bolt and Duboule, 2020, Klemm et al., 2019, Lieberman-Aiden et al., 2009, Rassek and Izykowska, 2020.

la réalisation de CHIP-Seq en *single-cell*, qui permet l'étude du paysage épigénétique spécifiquement au sein de chaque cellule ou type-cellulaire d'un échantillon (Grosselin et al., 2019).

De manière complémentaire au CHIP-Seq de marques épigénétiques, la réalisation d'expériences d'ATAC-Seq (*Assay for transposase-accessible chromatin using sequencing*) permet l'identification de l'ensemble des régions de chromatine ouverte sur le génome, en *bulk* ou en *single-cell* (Galang et al., 2020; Wang et al., 2020). En amont du séquençage et de l'analyse bioinformatique, des transposases hyperactives (Tn5) sont utilisées pour cliver la chromatine et insérer en parallèle des adaptateurs de séquençage Illumina, mais uniquement dans les régions accessibles (non compactées) du génome (Figure 26C) (Klemm et al., 2019). Avec cette technique, des analyses « d'empreintes » de facteurs de transcription (ou *footprinting* en anglais) sont réalisables dans les régions ouvertes de chromatine, où peuvent se fixer les facteurs de transcription (Bentsen et al., 2020). Le séquençage des quelques paires de bases occupées par un facteur de transcription est moins efficace, provoquant ainsi une légère dépression dans le pic d'ATAC-Seq associé à cette région génomique ouverte. En comparant l'intensité des dépressions de ces pics obtenus dans des échantillons contrôles et KO pour un facteur de transcription, il est possible d'identifier des régions plus ou moins occupées par certains facteurs de transcription, du fait du KO du facteur de transcription d'intérêt. Ces données permettent de générer des hypothèses sur de potentiels mécanismes de remodelage, comme par exemple : l'augmentation de la fixation d'un facteur de transcription « B » dans des échantillons KO pour le facteur de transcription « A », peut suggérer un mécanisme de compensation de « B » pour pallier l'absence de « A ». La principale limite de cette étude de *footprinting* est qu'elle est basée sur les motifs consensus connus de facteurs de transcription et ne permet donc pas d'observer les variations de fixation des facteurs de transcription pour lesquels les motifs ne sont pas clairement identifiés.

L'architecture 3D de la chromatine et les contacts physiques entre des régions distantes de l'ADN peuvent être identifiés par la technique de Hi-C qui est un dérivé des techniques de « *chromosome conformation capture* » ou 3C (Oudelaar and Higgs, 2021). L'ADN est d'abord crosslinké, puis digéré avec des enzymes de restriction, avant que les extrémités des fragments d'ADN soient marquées avec de la biotine et liguées ensemble. Les produits de ligation sont ainsi (1) constitués de fragments d'ADN qui étaient à l'origine très proches dans l'espace au sein du noyau et (2) marqués par la biotine. La fragmentation de

l'ADN suivi d'un enrichissement utilisant des billes de streptavidine et enfin d'un séquençage profond, permet la création de cartes de contacts entre les différentes régions de l'ADN (Figure 26D) (Lieberman-Aiden et al., 2009). Contrairement aux autres techniques de « *chromosome conformation capture* » (3C, 4C ou Capture-C), la technique de Hi-C permet d'étudier l'ensemble des contacts de chromatine qui s'établissent.

Analyse des partenaires protéiques des facteurs de transcription

La technique de référence pour l'identification d'interactions physiques entre des facteurs de transcription et des protéines partenaires est la technique de co-immunoprécipitation (Co-IP). Elle consiste à utiliser un anticorps spécifique pour immunoprécipiter un facteur de transcription d'intérêt et d'analyser ensuite les partenaires interagissant physiquement avec lui. L'analyse des cofacteurs peut se faire : (1) soit de manière ciblée, par dénaturation des complexes puis analyse par western-blot, mais nécessite de présupposer des partenaires protéiques potentiels car des anticorps spécifiques de ces partenaires sont nécessaires lors du western-blot ; (2) soit de manière non biaisée, en couplant la Co-IP à des techniques de spectrométrie de masse (Gonzalez-Teran et al., 2022; Kathiriya et al., 2015).

Il existe également des techniques d'imagerie qui ont pour avantage d'étudier les dynamiques des facteurs de transcription dans des cellules vivantes (Liu and Tjian, 2018). Par exemple, la technique de FRET (*fluorescence resonance energy transfer*) permet d'analyser des interactions entre deux protéines d'intérêt de manière continue dans des cellules vivantes et de localiser précisément le lieu de rencontre de ces protéines dans la cellule (Cui et al., 2019). La technique est basée sur un transfert d'énergie d'un fluorophore « donneur », ajouté à l'une des protéines d'intérêt, vers un fluorophore « accepteur », ajouté à l'autre protéine d'intérêt, générant un signal de fluorescence uniquement lorsque les deux protéines interagissent physiquement. Différentes limites du FRET, notamment la faible efficacité de la réaction, ont conduit au développement de nouvelles techniques comme le *Proximity ligation assay* ou PLA (Alam, 2018; Broussard et al., 2013; Cui et al., 2019).

Intégration des données multi-omiques

Les techniques omiques permettent l'étude de différentes modalités (génomique, épigénomique, transcriptomique, ...) sans a priori, mais il est cependant nécessaire de les intégrer ensemble pour obtenir une vision complète des phénomènes étudiés (Leon-Mimila et al., 2019). Lorsqu'elles sont analysées sans a priori et combinées ensuite, ces différentes modalités

permettent des validations croisées de phénomènes biologiques observés (Andersson et al., 2019). Par exemple, l'intégration des données d'épigénétique et/ou de conformation de la chromatine, avec des données de transcriptomique est pertinente. En effet, les observations épigénétiques peuvent renseigner sur les mécanismes de régulation de la transcription mais ne permettent pas une prédiction fiable de l'expression de gènes sous-jacente ([voir partie I.3.C Lien entre la transcription et la conformation 3D de la chromatine](#)), d'où la nécessité d'ajouter les informations de variation d'expression de gènes, fournies par l'analyse transcriptomique (VanOudenhove et al., 2020).

Pour comprendre les mécanismes de régulation de l'expression des gènes, parmi lesquels les facteurs de transcription ont une place prépondérante, il faudrait idéalement intégrer les différentes couches d'informations omiques (génomique, épigénomique, transcriptomique, protéomique, ...) au sein d'un système biologique complexe, car, bien qu'elles jouent des rôles indépendants, elles sont toutes biologiquement liées au sein de la cellule (Jung et al., 2020; Leon-Mimila et al., 2019).



Objectifs

Les facteurs de transcription sont des éléments majeurs du développement cardiaque, puisqu'en interagissant avec l'ADN et en induisant un remodelage de la chromatine, ils gouvernent la mise en place d'un programme d'expression spécifique de gènes. L'enchaînement chronologique des étapes de ce programme requiert l'expression successive de gènes codant des facteurs de transcription, eux-mêmes activés par d'autres facteurs de transcription, formant ainsi un réseau de régulation de facteurs de transcription. Bien que différents sous-réseaux soient connus, il manque une vision globale du réseau de régulation des facteurs de transcription cardiaques. Plus spécifiquement, certains facteurs de transcription, comme ceux de la famille Iroquois qui ont pourtant été montrés comme importants dans le cœur, n'ont pas ou peu été analysés dans des réseaux avec d'autres facteurs de transcription cardiaques.

Le premier projet de cette thèse était donc d'identifier un réseau de régulation global entre les facteurs de transcription cardiaques au cours du développement, afin d'identifier, dans un contexte intégré, les rôles de la famille de facteurs de transcription Iroquois.

Les différents membres de la famille des facteurs de transcription Iroquois ont été montrés comme ayant des fonctions spécifiques mais aussi des fonctions chevauchantes, en fonctionnant en redondance fonctionnelle et en coopération. Plus spécifiquement, du fait de leur expression similaire et des impacts de leur simple ou double KO chez la souris, les facteurs de transcription IRX3 et IRX5 semblent avoir ce mode de fonctionnement redondant et coopératif.

Le second projet de cette thèse était donc d'identifier sans a priori, les fonctions spécifiques et les fonctions redondantes d'IRX3 et d'IRX5 dans un contexte cardiaque humain.



Matériels et Méthodes

III.1. Méthodes spécifiques au projet 1

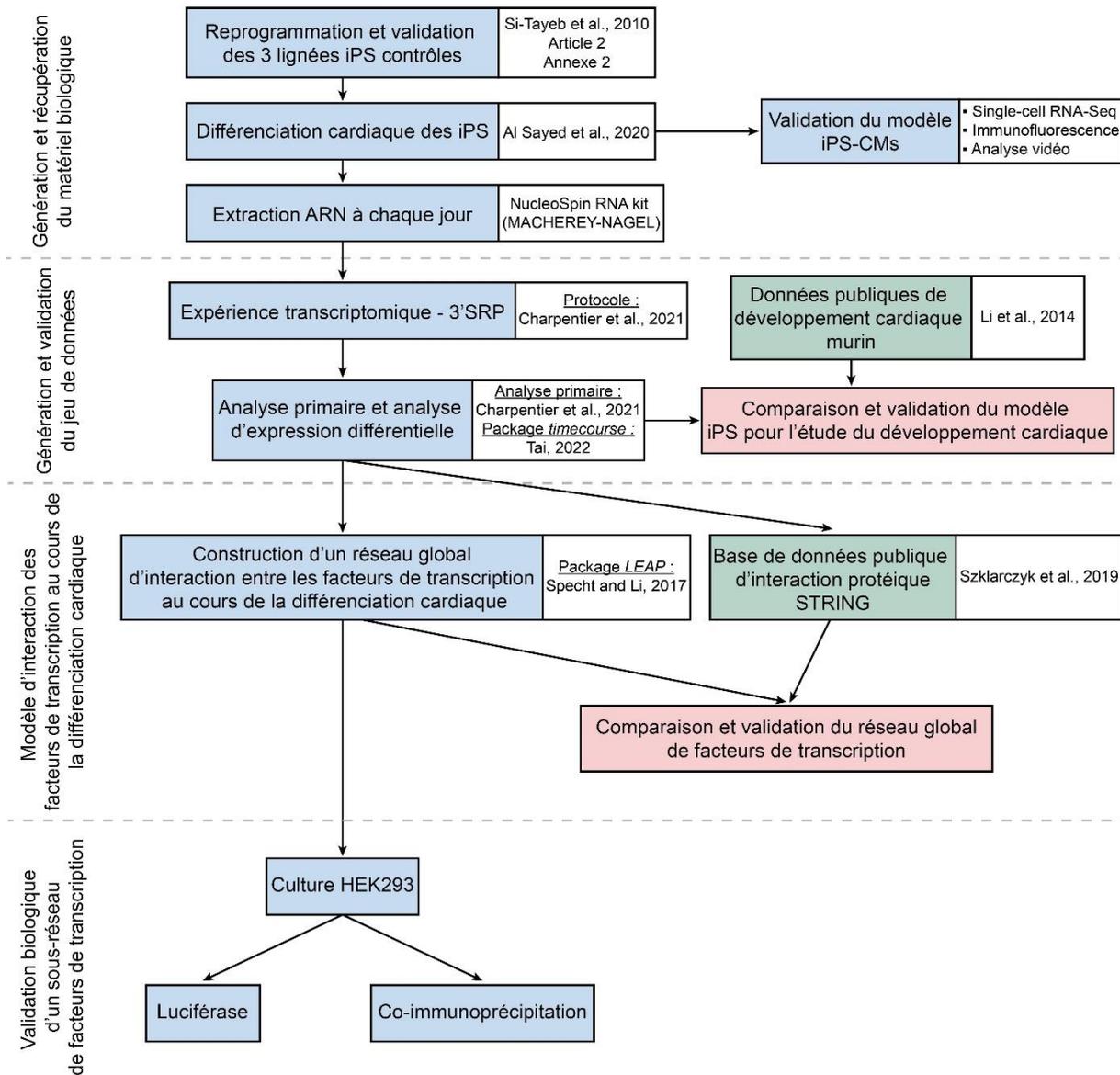


Figure 27 : Vue d'ensemble et organisation des méthodes utilisées pour le projet 1

Les méthodes pour ce projet sont détaillées dans la partie *Materials and Methods* de l'article 1 (IV.1.A - Article 1).

III.2. Méthodes spécifiques au projet 2

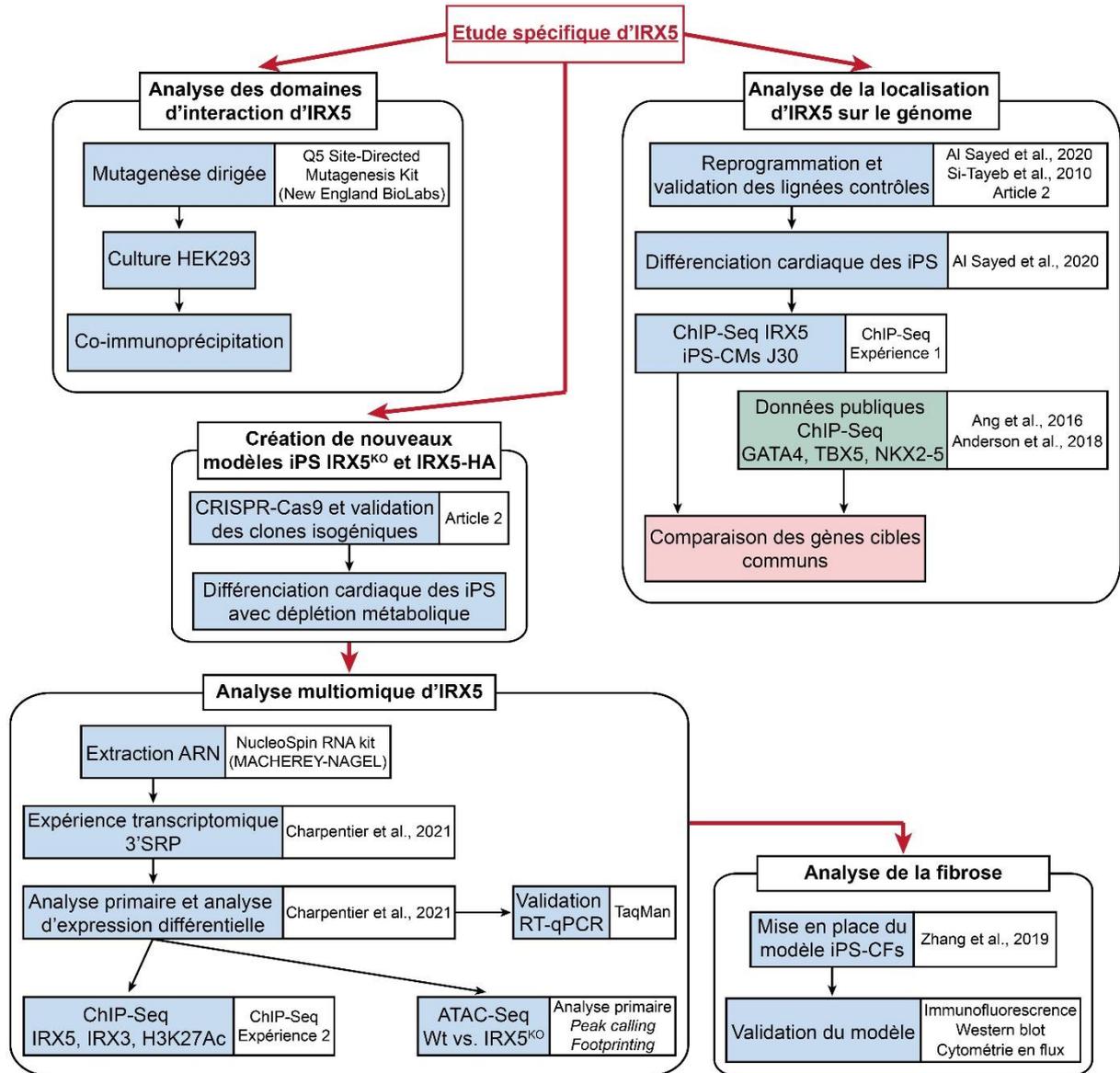


Figure 28 : Vue d'ensemble et organisation des méthodes utilisées pour le projet 2

III.2.A. Etude des domaines d'interaction d'IRX5

III.2.A.a. Mutagenèse dirigée

La réalisation des protéines IRX5 tronquées a été faite en utilisant le kit de mutagenèse dirigée Q5® Site-Directed Mutagenesis Kit (New England BioLabs) et le plasmide IRX5 contrôle (Tableau 3). Brièvement, le plasmide IRX5 contrôle a été amplifié par PCR ([98°C, 30sec, 1 cycle] ; [98°C, 10sec + T°a, 30sec + 72°C, 4min, 25 cycles] ; [72°C, 2min, 1 cycle]) en utilisant des primers spécifiques des régions à retirer et le *Q5 Hot Start High-Fidelity 2X Master Mix*. Les produits de PCR ont ensuite été traités pendant 5min à température ambiante avec le *10X KLD Enzyme Mix* (Kinase, Ligase, DpnI). Des bactéries thermocompétentes DH5 α ont ensuite été transformées avec le produit du traitement KLD, avant amplification. Enfin, les plasmides amplifiés ont ensuite été purifiés avec le kit *QIAGEN Plasmid Plus Midi Kit* (QIAGEN), avant d'être validés par séquençage.

Tableau 3 : Primers de mutagenèse dirigée

Construction plasmidique	Primers de PCR pour mutagenèse	T°a
IRX5 Δ H1	For : 5'-CACCGCAAGAACCCCTAC-3'	67°C
	Rev : 5'-TGTGGCGTTCTTCCGGTA-3'	
IRX5 Δ H2	For : 5'-ACCAAGATGACCCTCACC-3'	64°C
	Rev : 5'-GGTGGGGTAGGGGTTCTT-3'	
IRX5 Δ H3	For : 5'-GAGAATAAAATGACGTGGAC-3'	59°C
	Rev : 5'-GGTCATCTTGGTGATGATG-3'	
IRX5 Mid	For : 5'-GTGCTCGCCAAGCCCAA-3'	66°C
	Rev : 5'-ATTCTCTTTCTTGAGGCGCC-3'	
IRX5 C-ter	For : 5'-ACGCGTACGCGGCCGCTC-3'	72°C
	Rev : 5'-GGCGATCTCTGCCAAAGACCACAGTTTG-3'	

III.2.A.b. Culture et transfection des HEK293

Le protocole de culture et de transfection des HEK293 en vue d'expériences de co-immunoprécipitation, est similaire à celui détaillé dans l'article 1 (IV.1.A - Article 1).

III.2.A.c. Co-immunoprécipitation

Le protocole de co-immunoprécipitation, ainsi que les anticorps utilisés pour les résultats complémentaires sont détaillés dans l'article 1 (IV.1.A - Article 1).

III.2.B. Etude de la localisation d'IRX5 sur le génome

III.2.B.a. Reprogrammation et culture des lignées iPS

Les lignées iPS issues de deux individus contrôles, utilisées pour l'expérience 1 de ChIP-Seq, ont été précédemment reprogrammées (H166Wt dans Al Sayed et al., 2021 (voir Annexes) ; C2a dans Si-Tayeb et al., 2010) par surexpression de facteurs de transcription de pluripotence par virus Sendai ou par lentivirus respectivement. Les lignées hiPSC ont été maintenues à 37°C, 5% CO₂, 21% O₂ dans le milieu *StemMACS™ iPS Brew XF* (Miltenyi Biotec) dans des plaques de culture recouvertes de *Matrigel® hESC-Qualified Matrix* (0,05mg/mL, Corning). A 75% de confluence, les cellules ont été changées de plaque avec du *Gentle Cell Dissociation Reagent* (STEMCELL™ Technologies).

III.2.B.b. Différenciation cardiaque des iPS

Les différenciations cardiaques des iPS ont été réalisées à l'aide de la méthode de « sandwich » de matrigel (Es-Salah-Lamoureux et al., 2016). Brièvement, à 90% de confluence des iPS, une couche de *Growth Factor Reduced Matrigel* (0,033mg/ml, BD Corning) a été ajoutée sur les cellules (Jour -1). Les différenciations ont ensuite été débutées 24h plus tard (Jour 0), en cultivant les cellules avec de l'Activine A (100ng/mL, Miltenyi Biotec) et du FGF2 (10ng/mL, Miltenyi Biotec) pendant 24 heures, dans le milieu RPMI1640 (Life Technologies) supplémenté en B27 sans insuline (Life Technologies), L-glutamine (2mM, Life Technologies), acides aminés non essentiels (1%, Life Technologies) et Pénicilline/Streptomycine (1X, Life Technologies). Le lendemain, le milieu a été changé par du RPMI1640 supplémenté en B27 sans insuline, BMP4 (10ng/mL, Miltenyi Biotec), FGF2 (5ng/mL), L-glutamine (2mM), acides aminés non essentiels (1%) et Pénicilline/Streptomycine (1X), pendant 4 jours. A partir du jour 5, les cellules ont été cultivées dans un milieu RPMI1640 supplémenté en B27 complet (Life Technologies), L-glutamine (2mM), acides aminés non essentiels (1%) et Pénicilline/Streptomycine (1X), qui a été changé tous les deux jours jusqu'au jour 30. Les zones battantes de cardiomyocytes en fin de différenciation ont été dissociées mécaniquement à l'aiguille (23 gauges) et congelées à -80°C dans une solution de PBS (1X) + SVF (20%) + DMSO (10%), avant leur utilisation pour l'expérience 1 de ChIP-Seq.

III.2.B.c. Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 1

Expérimentation

Les échantillons de zones battantes de cardiomyocytes en fin de différenciation préalablement prélevés, ont été lysés par incubation (20min, 4°C) dans une solution de lyse (Tris-HCl 50mM, NaCl 150mM, Triton X-100 1%, Sodium deoxycholate 0.1%, CaCl₂ 5mM, Inhibiteurs de protéases 1X (Sigma-Aldrich)) et l'ADN a été fragmentée par digestion MNase (300U, New England BioLabs) 10min à 37°C. La réaction de lyse a ensuite été stoppée par ajout d'une solution de dilution (Tris-HCl 50mM, NaCl 150mM, Triton X-100 1%, Sodium deoxycholate 0.1%, EGTA 50mM, EDTA 1mM, Inhibiteurs de protéases 1X).

Le contrôle négatif d'immunoprécipitation (échantillon *Input*) a été créé par le mélange de 2% de chaque échantillon. Les échantillons ont ensuite été incubés avec 2µg d'anticorps anti-IRX5 (WH0010265M1, Sigma-Aldrich) sur la nuit à 4°C. Puis ils ont été incubés 4h à 4°C avec des billes de protéine G (*Invitrogen DynabeadsTM Protein G*, Thermo Fisher Scientific), avant 2 lavages avec la solution A (HEPES-KOH (pH 7.5) 50mM, NaCl 150mM, Triton X-100 1%, Sodium deoxycholate 0.1%, EDTA 2mM, SDS 0.1%), 2 lavages avec la solution B (HEPES-KOH (pH 7.5) 50mM, NaCl 300mM, Triton X-100 1%, Sodium deoxycholate 0.1%, EDTA 2mM, SDS 0.1%), 2 lavage avec la solution C (Tris-HCl (pH 8) 10mM, LiCl 0.25M, EDTA 1mM, NP40 0.5%, Sodium deoxycholate 0.5%) et 1 lavage avec la solution D (Tris-HCl (pH 8) 10mM, EDTA 1mM). L'ADN a été élué par incubation dans un tampon d'éluion (Tris-HCl (pH 7.5) 50mM, EDTA 10mM, SDS 1%) pendant 30min à 63°C, suivie d'une incubation d'1h à 63°C en présence de protéinase K (1X, Eurobio). L'ADN a ensuite été extrait au phénol chloroforme avec le kit *MaXtract High Density* (Qiagen), avant purification par précipitation à l'éthanol 100%.

La librairie de séquençage a ensuite été créée en 5 étapes successives, pour tous les échantillons et 200ng d'*Input* :

1. Etape de réparation des fragments (*End repair*) : Incubation de chaque échantillon 20min à 12°C, puis 15min à 37°C dans une solution de *T4 DNA Ligase Buffer* (1X, New England BioLabs), dNTP (0.4mM), *T4 DNA Polymerase* (60U/mL, Enzymatics), *T4 Polynucléotide kinase* (200U/mL, Enzymatics) et *DNA Pol I Klenow fragment* (20U/mL, Enzymatics). L'ADN a ensuite été purifié à l'aide de billes *Ampure SRPI* (BECKMAN COULTER) et d'éthanol 70%, puis élué au Tris-HCl (pH 8) 10mM, avant le passage à l'étape suivante.

2. Etape d'addition de base A : Incubation de chaque échantillon 30min à 37°C dans une solution de *NEBuffer 2* (1X, New England BioLabs), dATP (2mM) et l'enzyme *Klenow Fragment (3'→5' exo-)* (100U/mL, New England BioLabs). L'ADN a ensuite été purifié à l'aide de billes *Ampure SRPI* et d'éthanol 70%, puis élué au Tris-HCl (pH 8) 10mM, avant le passage à l'étape suivante.
3. Etape de ligation des adaptateurs de séquençage : Incubation de chaque échantillon 15min à 25°C dans une solution d'adaptateurs *DNA Adaptors SP* (20nM), *2X Rapid Ligation Buffer* (1X, Enzymatics) et *T4 DNA Ligase* (12 000U/mL, Enzymatics). L'ADN a ensuite été purifié à l'aide de billes *Ampure SRPI* et d'éthanol 70%, puis élué au Tris-HCl (pH 8) 10mM, avant le passage à l'étape suivante.
4. Etape d'enrichissement PCR : Les échantillons ont ensuite été amplifiés et indexés de manière unique (une séquence index unique par échantillon) par PCR ([98°C, 30sec, 1 cycle] ; [98°C, 10sec + 63°C, 30sec + 72°C, 30sec, 15 cycles] ; [72°C, 5min, 1 cycle]), avec le *Phusion High-Fidelity PCR Master Mix with HF Buffer* (Thermo Fisher Scientific).
5. Etape de sélection des fragments d'ADN selon la taille (*Size Selection*) : Des billes *Ampure SRPI* ont été utilisées pour réaliser une sélection négative afin de retirer les fragments d'ADN non digérés, suivie d'une sélection positive des fragments d'intérêt d'environ 300pb.

Les échantillons ont ensuite été regroupés et séquencés en *single-end* par la plateforme GenoBiRD de Nantes, sur HiSeq 2000 (*Rapid Run*).

Analyse

Les échantillons ont été démultiplexés par la plateforme GenoBiRD de Nantes pour obtenir des fichiers *fastq*. Les *reads* ont ensuite été alignés sur le génome de référence humain (hg19) avec l'outil *bowtie2* (Langmead and Salzberg, 2012), puis ceux avec une qualité d'alignement < 30 et les duplicats ont été retirés avant la création d'un fichier BAM par échantillon. La complexité de la librairie a été évaluée selon les normes définies par le consortium ENCODE (score de fraction non redondante, NRF et les scores *PCR bottleneck* 1 et 2, PCB1 et PCB2, Dunham et al., 2012). Le génome de référence a été partitionné en fenêtres de 10pb avec la fonction *bedtools makewindows* (Quinlan and Hall, 2010), afin de pouvoir obtenir le nombre de *reads* dans chaque fenêtre pour chaque échantillon avec la fonction *bedtools coverage*, qui a permis l'obtention d'un fichier BED par échantillon. Les 6 réplicats biologiques ont ensuite été combinés pour améliorer le rapport signal/bruit en utilisant l'outil

NGS-Integrator sur les BED (Wen et al., 2020). Le *peak calling* a été réalisé sur le fichier bedGraph de sortie de *NGS-Integrator*, avec l'outil *MACS2* (fonction *bdgpeakcall*, Zhang et al., 2008). Les fichiers bedGraph ont été convertis en fichiers bigWig avec la fonction *bedGraphToBigWig*, pour visualisation sur le *UCSC Genome Browser*. Les pics IRX5 identifiés ont ensuite été associés aux gènes les plus proches avec le logiciel *GREAT* (McLean et al., 2010), sur le génome hg19 et les paramètres par défauts des règles d'association *Basal plus extension*. Une analyse d'annotation fonctionnelle des gènes associés, a été réalisée par *GREAT* également.

Des données déjà publiées de CHIP-Seq dans des cardiomyocytes dérivés de cellules souches pour GATA4 et TBX5 (Ang et al., 2016), ainsi que pour NKX2-5 (Anderson et al., 2018), ont été utilisées. Les pics obtenus dans ces différentes études ont été associés aux gènes les plus proches également par *GREAT*, avec les mêmes paramètres que précédemment. L'analyse d'annotation fonctionnelle (base de donnée : *Reactome Pathway*) des gènes occupés par IRX5, GATA4, TBX5 et NKX2-5 a été réalisé avec le logiciel *WebGestalt* (Liao et al., 2019), par comparaison au génome entier ou à une liste restreinte de gènes d'intérêt.

III.2.C. Génération du nouveau modèle d'étude iPS IRX5^{KO}

III.2.C.a. Edition génique par CRISPR-Cas9 et validation des clones isogéniques

La stratégie et le protocole d'édition génétique qui ont été utilisés pour générer les lignées iPS IRX5^{KO} et IRX5-HA, ainsi que leur validation, sont détaillés dans l'article 2 ([IV.2.B - Article 2](#)).

III.2.C.b. Différenciation cardiaque des iPS avec déplétion métabolique

Le protocole de différenciation cardiaque avec déplétion métabolique, qui a été utilisé pour l'approche multi-omique du projet 2, était identique à celui sans déplétion détaillé précédemment jusqu'au jour 10 ([III.2.B.b Différenciation cardiaque des iPS](#)). En effet, au jour 10 le milieu a été remplacé par un milieu de déplétion : milieu RPMI1640 sans glucose (Life Technologies) supplémenté en B27 complet et Pénicilline/Streptomycine (1X) pendant 3 jours. Les cellules ont ensuite été dissociées (jour 13) avec une solution de trypsine *TrypLE* (10X, Life Technologies) et passées dans un milieu RPMI1640 supplémenté en B27 complet, L-glutamine (2mM), acides aminés non essentiels (1%), Pénicilline/Streptomycine (1X) et l'inhibiteur de Rho-kinase Y-27632 (STEMCELL™ Technologies). Au jour 14, le milieu a été remplacé par du

milieu de déplétion pendant 3 jours. A partir du jour 17, les cellules ont été maintenues dans le milieu RPMI1640 supplémenté en B27 complet, L-glutamine (2mM), acides aminés non essentiels (1%) et Pénicilline/Streptomycine (1X), et a été changé tous les deux jours jusqu'au jour 30.

Des cardiomyocytes (entre 8 et 25 millions) ont été dissociés au *TrypLE* (10X), puis congelés à -80°C dans une solution de PBS (1X) + SVF (20%) + DMSO (10%), avant leur utilisation pour l'expérience 2 de ChIP-Seq.

III.2.D. Approche multi-omique pour l'étude d'IRX5

III.2.D.a. Extraction d'ARN et RT-qPCR

Les ARN totaux ont été extraits avec le kit *NucleoSpin RNA* (MACHEREYNAGEL) et rétrotranscrits avec le kit *High-Capacity cDNA Reverse Transcription* (Applied Biosystems). Le protocole utilisé pour la validation PCR des nouveaux clones iPS IRX5^{KO} et IRX5-HA est détaillé dans l'article 2 (IV.2.B - Article 2). Dans le cadre des résultats complémentaires, l'amplification en PCR quantitative a été réalisée avec des sondes TaqMan (IRX5 : Hs00373920, KCNIP2 : Hs01552688) marquées au FAM (Applied Biosystems).

III.2.D.b. 3'seq-RNA Profiling (3'SRP)

Expérimentation

Des échantillons d'ARN à J-1 (iPS), J12, J20 et J30 de différenciations cardiaques (avec déplétion métabolique) d'iPS contrôles et IRX5^{KO} ont été prélevés. Puis une librairie de séquençage de 3'SRP a été préparée par la plateforme GenoBiRD de Nantes, selon leur protocole publié (Charpentier et al., 2021).

Analyse

Les échantillons ont été démultiplexés par la plateforme GenoBiRD de Nantes pour obtenir des fichiers *fastq* et l'analyse primaire a été faite avec leur pipeline *Snakemake* (Charpentier et al., 2021). Brièvement, après alignement sur le génome hg38 et le contrôle qualité de séquençage des échantillons, les matrices de comptage normalisées ont été générées par le package R *DESeq2* (Love et al., 2014). L'analyse des gènes différentiels entre les échantillons contrôles et IRX5^{KO} pour chaque jour respectivement, a été réalisée avec *DESeq2* également. Les clustering heatmap et les analyses en composantes principales (PCA) pour chaque jour ont été générés par le pipeline *Snakemake*.

La PCA globale sur tous les échantillons a été réalisée avec le package R *FactoMineR* (Lê et al., 2008), sur la matrice de comptage entière, centrée sur la moyenne et transformée en logarithme.

L'analyse d'annotation fonctionnelle a été réalisée avec le package R *ClusterProfiler* (Wu et al., 2021b), sur la base des termes de processus biologiques (*GO Biological Process*) provenant de la base de données *org.Hs.eg.db* (v3.14.0). Les processus biologiques significativement enrichis (p-value corrigée par Bonferroni < 0.05), par rapport au transcriptome de référence, ont été regroupés par similarité sémantique avec le logiciel REVIGO (Supek et al., 2011).

III.2.D.c. Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 2

Expérimentation

L'expérience 2 de ChIP-Seq a été réalisée en suivant un protocole similaire à l'expérience 1 (III.2.B.c Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 1), mais avec les spécificités décrites ci-après.

Des cardiomyocytes à J12 et à J30, issus de différenciations cardiaques contrôles et IRX5^{KO} réalisées avec déplétion métabolique, ont été utilisés pour le ChIP-Seq d'IRX3, d'IRX5 et de H3K27Ac. La fragmentation de l'ADN a été réalisée avec 400U de MNase pour chaque échantillon J12 et J30, en présence de sodium butyrate (5mM). Un échantillon *Input* composé de 5% de 5 échantillons pris au hasard a été créé par jour étudié respectivement. L'immunoprécipitation a été réalisée avec 10µg d'anticorps anti-IRX5 (5µg WH0010265M1, Sigma-Aldrich + 5µg sc-81102, Santa Cruz) ou 10µg d'anticorps anti-IRX3 (sc-166877, Santa Cruz) ou 2µg d'anticorps anti-H3K27Ac (39133, Active Motif), par échantillon. Les lavages des complexes associés aux billes de protéine G ont été réalisés de la même manière que pour l'expérience 1 dans le cas des échantillons H3K27Ac, ou uniquement avec les solutions A et D pour les échantillons IRX3 et IRX5.

Les échantillons ont ensuite été regroupés et séquencés en *paired-end* sur un NovaSeq 6000.

Analyse

Les échantillons ont été démultiplexés par la plateforme GenoBiRD de Nantes pour obtenir des fichiers *fastq* et l'analyse primaire a été faite avec le pipeline d'analyse *FULL-ChipSeq* (<https://gitlab.univ-nantes.fr/foucal-a/full-chipseq>). Les *reads* ont été alignés sur le génome de référence humain (hg19) avec l'outil *bowtie2*, puis ceux avec une qualité

d'alignement < 30 et les duplicats ont été retirés avant la création d'un fichier BAM par échantillon. La complexité de la librairie a été évaluée selon les normes définies par le consortium ENCODE (score de fraction non redondante, NRF et les scores *PCR bottleneck* 1 et 2, PCB1 et PCB2). Le *peak calling* a été réalisé par comparaison entre les échantillons et l'Input correspondant, avec l'outil *MACS2* (fonction *callpeak*, q-value : 0.05). Les échantillons IRX3 et IRX5 à J12 et à J30 n'ont pas passés les contrôles qualité basés notamment sur le nombre de pics obtenus par *MACS2* et l'enrichissement au TSS des gènes obtenus avec l'outil *deepTools* (Ramírez et al., 2016). Pour visualiser les données sur le *UCSC Genome Browser*, les fichiers d'alignement BAM ont été convertis en fichiers bigWig avec *deepTools* (fonction *bamCoverage*) et les fichiers BED de coordonnées des pics (sortie de *MACS2*) ont été convertis en fichiers bigBed avec la fonction *bedToBigBed*. Dans le cas des échantillons H3K27Ac, les pics identifiés ont été associés aux gènes les plus proches avec le logiciel *GREAT*, pour réaliser des comparaisons avec les gènes exprimés dans les expériences de 3'SRP.

III.2.D.d. Assay for transposase-accessible chromatin using sequencing (ATAC-Seq)

Expérimentation

Des échantillons de 50 000 cardiomyocytes ont été prélevés à J12 et à J30 de différenciations cardiaques (avec déplétion métabolique) d'iPS contrôles et IRX5^{KO}. La perméabilisation des cellules par la digitonine et la tagmentation des régions ouvertes d'ADN ont été réalisées avec le kit *Illumina Tagment DNA TDE1 Enzyme and Buffer* (Illumina) par incubation 30min à 37°C. L'ADN a ensuite été purifié avec le kit *DNA Clean & Concentrator-5* (Zymo Research). La création de la librairie a consisté en (1) l'ajout d'adaptateurs de séquençage pour chaque échantillon avec le kit *Nextera Index Kit* (Illumina), suivi (2) d'une amplification PCR ([72°C, 3min, 1 cycle] ; [98°C, 30sec, 1 cycle] ; [98°C, 10sec + 63°C, 30sec + 72°C, 3min, 11 cycles]) avec le kit *Nextera DNA library Prep Kit* (Illumina), et (3) d'une double sélection des fragments d'ADN selon la taille avec des billes Ampure XP (BECKMAN COULTER). La qualité de la librairie a été évaluée par TapeStation, avant quantification avec le kit *NEBNext Library Quant Kit for Illumina* (New England BioLabs). La librairie a ensuite été séquencée en *paired-end* sur un NovaSeq 6000.

Analyse

Le *trimming* et le démultiplexage des échantillons ont été effectués à l'aide de l'outil *bcl2fastq2* (Illumina). L'alignement des *reads* sur le génome de référence humain (hg19) a été

fait avec l'outil *bowtie2*. Les *reads* correctement alignés et non dupliqués ont ensuite été stockés dans des fichiers BAM, avant la conversion au format BEDPE par *deepTools*. Tous les échantillons ont été combinés pour obtenir un seul échantillon par jour de différenciation cardiaque étudié (J12 et J30) et par condition (contrôle et IRX5^{KO}). La complexité de la librairie a été évaluée selon les normes définies par le consortium ENCODE (scores identiques à ceux utilisés pour l'analyse du CHIP-Seq) et le *peak calling* a été réalisé avec *MACS2*, sur les échantillons regroupés.

Pour la visualisation sur le *UCSC Genome Browser*, les fichiers d'alignement BAM ont été convertis en fichiers bigWig et les fichiers BED de localisation des pics ont été convertis en fichiers bigBed, comme précisé pour l'analyse du CHIP-Seq.

L'analyse de footprinting (identification des cinétiques de fixation des facteurs de transcription sur leur motif entre les conditions contrôle et IRX5^{KO}) a été réalisé avec l'outil *TOBIAS* (Bentsen et al., 2020), en utilisant les seuils suivants : $-\log_{10}(pval) : 125$; Score de fixation : 0.1. Seuls les facteurs de transcription avec une expression détectée en 3'SRP (expression ≥ 1 UPM) ont été conservés dans les résultats.

III.2.E. Analyse des mécanismes de fibrose cardiaque

III.2.E.a. Différenciation des iPS en fibroblastes cardiaques

Les fibroblastes cardiaques dérivés d'iPS (iPS-CFs) ont été générés en suivant le protocole de Zhang *et al.* (Zhang et al., 2019). Brièvement, la différenciation a été lancée (J0) à 90% de confluence des iPS, en cultivant les cellules dans le milieu RPMI1640 supplémenté en B27 sans insuline, CHIR99021 (6 μ M, Miltenyi Biotec), L-glutamine (2mM), acides aminés non essentiels (1%) et Pénicilline/Streptomycine (1X), pendant 2 jours. A J2, le milieu a été remplacé sans le CHIR99021, pendant 24h. Puis à J3, le milieu a été changé par du RPMI1640 supplémenté en B27 sans insuline, IWR-1 (5 μ M, Sigma-Aldrich), L-glutamine (2mM), acides aminés non essentiels (1%) et Pénicilline/Streptomycine (1X), pendant 2 jours. A J5 les cellules ont été dissociées à l'acutase (Sigma-Aldrich) et remise en culture à raison de 20 000 cellules/cm², dans le milieu *Advanced DMEM medium* (Life Technologies), CHIR99021 (5 μ M), acide rétinoïque (2 μ M, Sigma-Aldrich) et l'inhibiteur de Rho-kinase Y-27632, pendant 24h. Le lendemain, le milieu a été remplacé sans inhibiteurs de Rho-kinase, pendant 2 jours. A J8, le milieu a été changé par de l'*Advanced DMEM medium* seul, pendant 4 jours. A J12, les cellules ont été dissociées à l'acutase et remise en culture (dilution au 1/3) dans le milieu *Fibroblast*

growth medium 3 (Promocell), FGF2 (15ng/mL) et SB431542 (10 μ M, Sigma-Aldrich), pendant 6 jours. A J18, le protocole est terminé et les fibroblastes ont été utilisés pour différentes applications soit directement (passage 0), soit ultérieurement jusqu'au passage 4 maximum.

III.2.E.a. Immunofluorescence

Les expériences d'immunofluorescences sur les iPS-CFs en fin de différenciation (J18), ont été réalisées selon le protocole présenté dans l'article 2 ([IV.2.B - Article 2](#)), avec l'anticorps primaire anti-Vimentine (1/250, ref : 5741, Cell Signalling) et l'anticorps secondaire anti-rabbit conjugué Alexa Fluor 488 (1/1000, ref : A-21206, Thermo Fisher Scientific). Les immunomarquages ont ensuite été examinés par microscopie inversée à fluorescence (ZeissTM Axiovert 200 M).

III.2.E.b. Western blot

Les expériences de western blot sur les iPS-CFs en fin de différenciation (J18), ont été réalisées selon le protocole présenté dans l'article 2 ([IV.2.B - Article 2](#)), en utilisant 20 μ g de protéine et les anticorps primaire anti-Vimentine (1/250, ref : 5741, Cell Signalling), anti-TBX5 (1/250, sc-515536, Santa Cruz) et anti-IRX5 (1/250, sc-81102, Santa Cruz), ainsi que les anticorps secondaires anti-rabbit HRP (1/10000, ref : 7074, Cell Signaling) et anti-mouse HRP (1/10000, ref : 7076, Cell Signaling).

III.2.E.c. Cytométrie en flux

Les iPS-CFs en fin de différenciation ont été fixés au paraformaldéhyde 4% pendant 10min à température ambiante, avant perméabilisation avec du PBS 1% Saponine, 10min à température ambiante. Les cellules ont ensuite été incubées avec l'anticorps primaire anti- α -SMA (1 μ g/mL, ref : A5228, Sigma-Aldrich) dilué dans du PBS 0.1% BSA 0.1% Saponine, 30min à température ambiante, puis avec l'anticorps secondaire anti-mouse conjugué Alexa Fluor 647 (1 μ g/mL, ref : A-31571, Thermo Fisher Scientific). L'acquisition et l'analyse des données ont été réalisées avec la logiciel FACSDivaTM et le cytomètre LSR II (Becton Dickinson, BD).

IV

Résultats

IV.1. Projet 1 : Réseau global des facteurs de transcription du développement cardiaque et famille Iroquois homeobox

Article 1

Deciphering transcriptional networks during human cardiac development

Robin Canac[#], Bastien Cimarosti[#], Aurore Girardeau, Virginie Forest, Pierre Olchesqui, Jeremie Poschmann, Richard Redon, Patricia Lemarchand, Nathalie Gaborit^{*}, Guillaume Lamirault^{*}

[#] : Co-premiers auteurs

^{*} : Co-derniers auteurs

DOI : 10.1101/2022.08.11.503560

Article soumis à publication dans le journal *Developmental Cell*

- Analyse transcriptomique en cinétique à chaque jour de la différenciation cardiaque de cellules iPS contrôles
- Construction d'un réseau global d'activation et d'inhibition entre les facteurs de transcription impliqués dans la différenciation cardiaque
- Identification d'un réseau d'activation entre les facteurs de transcription IRX3, IRX5 et les facteurs de transcription cardiaques majeurs GATA4, NKX2-5 et TBX5
- Validation biologique d'un fonctionnement en complexe multiprotéique de ces cinq facteurs de transcription, pour l'activation de l'expression du gène *SCN5A*

Contexte de la construction et la réalisation de cette étude :

La réalisation de ce travail a été faite en collaboration avec Bastien Cimarosti (autre doctorant de l'équipe). Des résultats obtenus au cours de nos deux projets de thèse ont été combinés pour aboutir à ce travail en co-premiers auteurs. Nous avons chacun pris en charge des parties respectives, ainsi que travaillé en collaboration sur d'autres, comme expliqué ci-après :

- La réflexion sur les problématiques scientifiques qui sont développées et sur l'intégration des résultats, ainsi que la construction des figures et l'écriture de l'article ont été réalisées en collaboration.
- Les données de transcriptomique en cinétique (*bulk*), l'analyse transcriptomique pour l'identification des vagues d'expression, et la génération et l'analyse du *single-cell* de validation de la différenciation cardiaque a été prise en charge par Bastien.
- Nous avons effectué ensemble la construction, la validation et l'analyse des réseaux de facteurs de transcription.
- L'ensemble du travail focalisé sur l'étude des facteurs de transcription Iroquois, comprenant les validations luciférase et co-immunoprécipitation correspond à la partie que j'ai prise en charge.

1 Deciphering transcriptional networks during human cardiac 2 development

3 Authors

4 Robin Canac,^{1,3} Bastien Cimarosti,^{1,3} Aurore Girardeau,¹ Virginie Forest,¹ Pierre Olchesqui,¹
5 Jeremie Poschmann,² Richard Redon,¹ Patricia Lemarchand,¹ Nathalie Gaborit,^{1,4,5,*} Guillaume
6 Lamirault,^{1,4,5,6,*}

¹ Nantes Université, CHU Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France

² INSERM, Nantes Université, Center for Research in Transplantation and Translational Immunology, UMR 1064, ITUN, F-44000 Nantes, France

³ These authors contributed equally

⁴ These authors contributed equally

⁵ Senior author

⁶ Lead Contact

* Correspondance: nathalie.gaborit@univ-nantes.fr (N.G.), guillaume.lamirault@univ-nantes.fr (G.L.)

7 **Summary**

8 Human heart development is governed by transcription factor (TF) networks controlling
9 dynamic and temporal gene expression alterations. Therefore, to comprehensively
10 characterize these transcriptional regulations, day-to-day transcriptomic profiles were
11 generated throughout the directed cardiac differentiation, starting from three distinct human
12 induced pluripotent stem cell lines from healthy donors (32 days). We applied an expression-
13 based correlation score to the chronological expression profiles of the TF genes, and clustered
14 them into 12 sequential gene expression waves. We then identified a regulatory network of
15 more than 23 000 activation and inhibition links between 216 TFs. Within this network, we
16 observed previously unknown inferred transcriptional activations linking IRX3 and IRX5 TFs to
17 three master cardiac TFs: GATA4, NKX2-5 and TBX5. Luciferase and co-immunoprecipitation
18 assays demonstrated that these 5 TFs could (1) activate each other's expression, (2) interact
19 physically as multiprotein complexes and (3) together, finely regulate the expression of
20 SCN5A, encoding the major cardiac sodium channel. Altogether, these results unveiled
21 thousands of interactions between TFs, generating multiple robust hypotheses governing
22 human cardiac development.

23

24 **Keywords**

25 Stem cell differentiation, Human induced pluripotent stem cells, Heart development,
26 Transcription factor, Gene regulatory networks, Transcriptomics, Transcription factor
27 complexes, Iroquois transcription factors

28 Introduction

29 Heart formation is a complex process that requires spatio-temporal interplay between
30 distinct and interdependent cell types through specific signaling and transcriptional pathways,
31 leading to their differentiation and specification (Meilhac and Buckingham, 2018; Vincent and
32 Buckingham, 2010). Defects in this developmental process result in congenital heart disease
33 as well as in a number of inherited cardiac disorders in adults (Postma et al., 2011). The specific
34 gene expression program, governing the formation of a functional heart, needs precise
35 regulation, in a time-, cell- and space-dependent manner (Asp et al., 2019). This program is
36 mediated by transcription factors (TFs) regulating the expression of other TF-encoding genes
37 and establish specific TF networks, such as between GATA4, NKX2-5 and TBX5 (Bruneau, 2013;
38 Luna-Zurita et al., 2016). These networks control and permanently remodel over time the
39 transcriptional expression program that govern heart development.

40 A thorough understanding of these networks is crucial to gain knowledge on the
41 transcriptional regulations and dysregulations that govern normal and pathological cardiac
42 development, respectively. However, full knowledge of the global TF regulatory network of
43 cardiac development is still missing. For instance, while several studies on Iroquois homeobox
44 TF family (IRX) have shown their key roles on the regulation of adult cardiac electrical
45 conduction (Al Sayed et al., 2021; Bruneau et al., 2001; Costantini et al., 2005; Gaborit et al.,
46 2012; Zhang et al., 2011), their function during human cardiac development has not been
47 investigated yet. Cellular models derived from human induced Pluripotent Stem Cells (hiPSCs)
48 offer a unique opportunity to address these challenges, as they reproduce the cellular
49 differentiation processes which lead stem cells to acquire a cardiac cell phenotype, carrying
50 the genome of either healthy subjects or patients with inherited cardiac diseases.

51 In the present study, we first validated hiPSC cardiac differentiation model as a
52 relevant tool to unravel the global TF regulatory network governing human cardiac
53 development, identifying a network of 216 TFs with time-dependent activations and
54 inactivations. Among these, we identified and biologically validated an undescribed TF
55 regulatory network involving IRX3, IRX5 and three main cardiac TFs, GATA4, NKX2-5 and TBX5.
56 Furthermore, we generated new hypotheses on the potential mechanisms leading to the
57 cooperative effect of these TFs that could form a functional multiprotein complex activating
58 the promoter of SCN5A, encoding the main cardiac sodium channel.

59 **Results**

60 ***Directed cardiac differentiation robustly generates functional cardiac cells***

61 Cardiac differentiation of three hiPSC lines reprogrammed from three healthy donors
62 was used as a cellular model of cardiac development (Figure 1A). After directed cardiac
63 differentiation, all three hiPSC lines expressed cardiac-specific troponin I (Figure 1B), and
64 displayed spontaneous contractions (Figure 1C), demonstrating their capability to form
65 functional cardiomyocytes.

66 Based on single-cell transcriptomic data from 14 520 cells obtained at the end of
67 directed cardiac differentiation (Figure 1D and Supplementary Figure S1), about 95% of the
68 cells could be successfully annotated to one of the 15 cell types described in the developing
69 human fetal heart (Asp et al., 2019), including 34% cardiomyocytes, 21% epicardial cells, and
70 15% fibroblast-like cells. This distribution was similar to previous findings in the adult human
71 heart (Litviňuková et al., 2020). These data indicate that directed cardiac hiPSC differentiation
72 generated the cellular diversity observed in human fetal heart, known to be not only necessary
73 for cardiac function, but also required for the establishment of cardiomyocytes (Meilhac,
74 2018).

75 To investigate how gene expression variations are orchestrated throughout cardiac
76 differentiation, we then generated daily transcriptomic data, from hiPSC stage (D-1) to day 30
77 (D30), for three independent cardiac differentiations of each of the three hiPSC lines (Figure
78 2A). Directed cardiac differentiation was associated to gradual temporal transcriptomic
79 changes, represented on the first principal component (PC1) of the principal component
80 analysis (Figure 2B). PC1 was significantly correlated with time from the onset of cardiac
81 differentiation (spearman correlation coefficient $\rho=0.87$, $p\text{-value} < 2.2 \cdot 10^{-16}$). Cardiac
82 differentiation evolution, represented by PC1 (Figure 2C), showed that 85% of transcriptomic
83 variations were achieved by D14 (Figure 2D). Altogether, these data demonstrate that the first
84 14 days of hiPSC cardiac differentiation represent the ideal time window to investigate the
85 molecular processes that lead to functional cardiac cells.

86 ***Transcriptomic kinetics of hiPSC cardiac differentiation unveiled biological processes***
87 ***involved during cardiac development***

88 Focusing on gene expression changes related to hiPSC cardiac differentiation, the 3
89 000 genes with the most significant expression variation during directed cardiac
90 differentiation (differentially expressed genes; DEG) were identified and grouped into 12
91 clusters, chronologically ordered based on the time point when their expression level changes
92 the most, showing distinct temporal gene expression profiles (Figure 3A). The average
93 temporal expression pattern of each cluster was then compared to transcriptomic data
94 obtained for the same genes from an in vivo reference model of murine cardiac development
95 (Li et al., 2014; Figure 3B). As cardiac cells derived from hiPSCs are usually described as
96 reaching an equivalent of, at the most, E18.5 stage in murine embryonic development
97 (DeLaughter et al., 2016), we restricted the comparison of the hiPSC dataset to murine
98 developmental transcriptomic data obtained between murine embryonic stem cells and E18.5
99 stage. Apart from cluster D, all clusters displayed strikingly similar expression patterns
100 between hiPSC cardiac differentiation and murine cardiac development. Nevertheless, genes
101 of cluster D were associated with gastrulation biological processes (Figure 3B – Cluster D
102 middle panel) which is completed before E7.5. As no data was available between the mESC
103 and E7.5 stages in the murine experiments, relevant gene expression changes associated to
104 this process were likely to be absent in the murine transcriptomic dataset but remained
105 detectable in the daily hiPSCs cardiac differentiation dataset. For all other 11 clusters, hiPSC
106 cardiac differentiation could be confidently matched to sequential gene expression waves that
107 occur during murine cardiac development. Altogether, these clusters recapitulate key steps of
108 cardiac development, including (1) expression decrease of genes related to pluripotency and
109 stemness maintenance (Figure 3B – Cluster A to C), followed by the transient expression of
110 genes related (2) to gastrulation and mesoderm formation (Figure 3B – Cluster D) and (3) to
111 early cardiac development (Figure 3B – Cluster E). These specific patterns were then followed
112 by the successive implementation and persistence over time of gene expression waves that
113 set up the sequential establishment of the functional cardiac phenotype (Figure 3B – Cluster
114 F to L). To confirm these results, similar analyses were conducted on the top 3 000 DEG during
115 murine cardiac development from mouse embryonic stem cells (mESCs) to E18.5
116 (Supplementary Figure S2). This again revealed consistency of gene expression changes during
117 hiPSC cardiac differentiation and during murine cardiac development. Collectively, these

118 analyses demonstrate that hiPSC cardiac differentiation precisely recapitulates transcriptomic
119 processes related to human and mouse cardiac development.

120 ***Prediction of gene regulatory networks governing hiPSC cardiac differentiation***

121 TFs are known to be key players of developmental processes (Luna-Zurita et al., 2016;
122 Ng et al., 2021). Therefore, to elucidate gene regulatory networks that underlie human cardiac
123 development, gene expression analysis was then focused on all 216 TFs that were found to be
124 differentially expressed during the time-course of cardiac differentiation (Figure 4A;
125 Supplementary Table 1). Overall, 69% of these TFs have been already linked to cardiac
126 (patho)physiological phenotypes (Supplementary Table 1). We chose to adapt an expression
127 correlation score involving time delay (LEAP method, see Methods) to capture gene
128 associations that are hidden by time lags (i.e. time delay between the mRNA expression of the
129 source gene and the mRNA expression of its target gene). Using this method on the 216 TFs,
130 we predicted interactions that activated or inhibited the expression of target TFs by source
131 TFs, building a regulatory network. This gene expression-based network included 11 467
132 activating interactions and 11 539 inhibitory interactions (Figure 4B left panel; Supplementary
133 Table 2). We then evaluated the biological relevance of these TF interactions, using the STRING
134 protein-protein interaction (PPI) database to generate an undirected PPI-based network
135 restricted to the 216 TFs (Figure 4B right panel). Interestingly, 182 TFs (84%) were found to
136 share at least one known PPI interaction. This included interactions between TFs belonging to
137 the same gene cluster but also, interactions between TFs from different gene clusters,
138 suggesting coordination between TFs to regulate the successive gene expression waves.

139 Comparing both networks, the gene expression-based network (LEAP-based)
140 contained a greater amount of information than the PPI-based network (STRING-based)
141 (Figure 4B). Indeed, although both networks were generated using the same TF query list, the
142 density (i.e. normalized averaged number of neighbors) of the gene expression-based network
143 was 5.5 fold higher, as compared to the PPI-based network. Deeper analysis showed that
144 about 100% of the nodes and 80% of the links found in the PPI-based network were also found
145 in the gene expression-based network (Figure 4C-D). Moreover, focusing on links between
146 successive expression clusters, more than 76% of those found in PPI-based network were also
147 found in the gene expression-based network (Figure 4E). Further confirming the accuracy of

148 gene expression-based strategy, sub-networks that have been well-described in the literature
149 were also present in both networks: (1) the network composed of the main actors of
150 pluripotency (e.g POU5F1) and early phases of cardiac development (e.g EOMES, MESP1;
151 Figure 4F), and (2) the TF network implicated in cardiogenesis (e.g ISL1, MEF2C; Figure 4G).
152 This validated the relevance of such expression correlation score approach taking into account
153 time delay to comprehensively analyze TFs and their interactions throughout cardiac
154 differentiation. Altogether, while the gene expression-based network confirmed already
155 known and validated interactions, it also inferred 21 530 new interactions unveiling numerous
156 new hypotheses on TF networks potentially critical for cardiac development.

157 ***IRX3 and IRX5 are involved in triggering expression of GATA4, NKX2-5, TBX5 cardiac*** 158 ***transcription factor network***

159 Leveraging this new gene expression-based network to uncover new regulation
160 mechanisms, and based on our previous focus of interest (Al Sayed et al., 2021; Gaborit et al.,
161 2012) we evaluated IRX TF family involvement in the establishment of cardiac developmental
162 processes. Expression levels of the 6 different *IRX* TF genes was analyzed during cardiac
163 differentiation in the three hiPSC lines (Figure 5A). Expression of *IRX6* was undetectable and
164 expression of *IRX1* and *IRX2* did not vary over time. Only *IRX3*, *IRX4* and *IRX5* expression
165 increased significantly between D-1 and D30 of cardiac differentiation. Interestingly, based on
166 their expression profiles, *IRX3* and *IRX5* ranged from the earliest cardiac-specific gene cluster
167 with an expression level that was maintained until the end of cardiac hiPSC differentiation
168 (cluster F). This suggested a potential role for *IRX3* and *IRX5* in the early establishment of gene
169 regulatory networks essential for cardiac fate, and beyond. In contrast, *IRX4* expression was
170 detected in one of the latest clusters (cluster K). Therefore, we then focused on both *IRX3* and
171 *IRX5* TFs.

172 In order to investigate the role of *IRX3* and *IRX5* in cardiac differentiation progression,
173 all their potential TF targets in the subsequent G to L clusters were extracted from the gene
174 expression-based network (Figure 5B). Interestingly, the master cardiac TF genes *GATA4*
175 (cluster G), *NKX2-5* (cluster I) and *TBX5* (cluster L) were individually found to be potential
176 targets of both *IRX3* and *IRX5*. It is well established that *GATA4* acts in a multiprotein complex
177 with *NKX2-5* (cluster I) and *TBX5* (cluster L) cardiac TFs (Ang et al., 2016; Durocher et al., 1997;

178 Luna-Zurita et al., 2016). To further explore potential new interactions, we then focused on
179 the gene expression-based sub-network involving *IRX3*, *IRX5*, *GATA4*, *NKX2-5* and *TBX5*,
180 referred later as IGNIte sub-network (Figure 5C). In the IGNIte sub-network, *IRX3* and *IRX5*
181 were inferred as activators of *GATA4*, *NKX2-5* and *TBX5*, and confirming the literature
182 (Bruneau, 2013; McCulley and Black, 2012; Nemer and Nemer, 2010; Waardenberg et al.,
183 2014), *GATA4* was inferred as activator of *NKX2-5* and both *GATA4* and *NKX2-5* were activators
184 of *TBX5* expression.

185 In order to investigate the biological relevance of these inferred interactions, luciferase
186 assays were conducted on *GATA4*, *NKX2-5* and *TBX5* core promoters (Figure 5D). *IRX3* and
187 *IRX5* proteins were, separately (fold changes 4.2 and 1.5 respectively) or together (fold change
188 4.5), able to bind the promoter of *GATA4* and to activate luciferase expression. A slight
189 tendency towards potentiation of both activating effects is observable when *IRX3* and *IRX5*
190 were present but was not statistically significant. On the *NKX2-5* promoter, *IRX5* alone was
191 able to activate luciferase expression (1.3-fold change), but not *IRX3*, suggesting that the
192 inferred activation of *NKX2-5* by *IRX3* found in the IGNIte sub-network was due to *IRX5*, and
193 that the high similarity between *IRX3* and *IRX5* expression profiles caused the false-positive
194 link to appear. Together, *IRX3* and *IRX5* were able to activate *NKX2-5* promoter, with a
195 tendency towards potentiation too (fold change 1.2 between *IRX5* alone and *IRX3/IRX5*
196 conditions; $p > 0.05$). According to the order of appearance of TFs in IGNIte sub-network, *NKX2-*
197 *5* promoter activation was assessed in the combined presence of *IRX3*, *IRX5* and *GATA4*, which
198 showed an activator effect (1.8-fold change). Although a potentiation tendency was observed
199 when *GATA4* was present in addition with *IRX3* and *IRX5*, this effect was not statistically
200 significant. On *TBX5* promoter, *IRX3* and *IRX5* were able to bind and activate gene expression
201 either individually (2.7- and 1.2-fold change, respectively) or together (3.6-fold change).
202 Potentiation of both activator effects was clearly observable and statistically significant when
203 *IRX3* and *IRX5* were together on the *TBX5* promoter. Finally, considering the joint expression
204 of *IRX3*, *IRX5*, *GATA4* and *NKX2-5* from D10, we proved the activator effect of these TFs on
205 *TBX5* promoter (6.6-fold change), which is statistically increase from the *IRX3/IRX5* condition
206 (1.8-fold change). Collectively, these results biologically validated the new interactions
207 inferred with the gene expression-based network and illustrated the progressive temporal

208 activation of the major TFs GATA4, NKX2-5 and TBX5, by IRX3 and IRX5 during cardiac cell
209 lineage establishment.

210 ***IRX3 and IRX5 physically interact with GATA4, NKX2-5 and TBX5 to control SCN5A expression***

211 As the expression of the IGNiTe sub-network members was maintained until D30 of
212 hiPSC cardiac differentiation (Figure 6A), the functional role of IRX3, IRX5, GATA4, NKX2-5, and
213 TBX5 as a multiprotein complex was evaluated using co-immunoprecipitation and luciferase
214 assays in heterologous expression systems. Luciferase assays were conducted on the
215 promoter of *SCN5A*, a known target of these TFs (Al Sayed et al., 2021; Arnolds et al., 2012;
216 Briggs et al., 2008; Koizumi et al., 2016; Tarradas et al., 2017). According to the chronological
217 order of expression of these five TFs along cardiac differentiation of hiPSCs (Figure 6A), we
218 first investigated the role of IRX3 and IRX5. As previously described (He et al., 2009), IRX3 and
219 IRX5 physically interacted (Figure 6B top panel and Supplementary Figure S3) and could
220 cooperatively activate the *SCN5A* promoter (Figure 6B bottom panel). While IRX3 alone
221 activated the *SCN5A* promoter (2.3-fold change), IRX5 potentiated its effect with a 1.5-fold
222 change. GATA4 was able to physically interact with IRX5 but not with IRX3 (Figure 6C top
223 panels) and when the three TFs were co-transfected, only GATA4 and IRX5 interacted,
224 suggesting a competitive effect between IRX3 and GATA4 to bind IRX5 (Figure 6C bottom left
225 panel). Also, the addition of GATA4 potentiated (1.5-fold change) the activity of the IRX3/IRX5
226 couple on *SCN5A* promoter (Figure 6C bottom right panel). NKX2-5 interacted with both IRX3
227 and IRX5 individually (Figure 6D left panels), but again, when the four TFs were co-transfected
228 we only observed an interaction between IRX5, GATA4, and NKX2-5, suggesting again a
229 competition between IRX3 and IRX5, in favor of IRX5, in these interactions (Figure 6D central
230 panel). NKX2-5 amplified (8.0-fold change) the effect of the IRX3/IRX5/GATA4 trio on the
231 *SCN5A* promoter (Figure 6D right panels). Finally, when IRX3, IRX5, GATA4, NKX2-5 and TBX5
232 were co-transfected a global protein complex could be formed between IRX5, GATA4, NKX2-
233 5 and TBX5, but not with IRX3, even if IRX3 alone was able to interact with TBX5 (Figure 6E left
234 and central panels). However here, TBX5 slightly reduced (-1.6-fold change) the effect of the
235 IRX3/IRX5/GATA4/NKX2-5 quartet on *SCN5A* promoter (Figure 6E right panel) suggesting a
236 down-regulating role of TBX5 in this global complex. Collectively, we showed that following
237 *IRX3, IRX5, GATA4, NKX2-5* and *TBX5* gene expression increase during cardiac differentiation,

238 the direct activation of *SCN5A* expression is under the control of a time-changing multi-TFs
239 complex that controls the temporal expression profile of *SCN5A*.

240 **Discussion**

241 In this study, based on a transcriptomic kinetics study on cardiac differentiation of
242 hiPSCs, we identified the global TF regulatory network that is required for heart development.
243 We notably identified novel time-dependent TF-gene regulations that connect *IRX3* and *IRX5*
244 to the core cardiac *GATA4*, *NKX2-5* and *TBX5* TFs. We also found that these five TFs form
245 protein complexes to regulate target gene expression, such as *SCN5A*. Altogether, this time-
246 course bulk transcriptomic data provided a dynamic model relevant to identify new roles for
247 TFs in developmental processes.

248 ***In vitro modeling of time in cardiac development***

249 This study demonstrates that hiPSC cardiac differentiation is a relevant model to study
250 the successive steps leading to the establishment of the gene expression program during
251 human cardiac development. To date, most studies contributing to the knowledge on heart
252 development and TF regulation have been conducted in animal models, mainly in mice
253 (Krishnan et al., 2014), as access to human embryonic cardiac tissue, is indeed very limited. If
254 regulatory mechanisms of development are overall highly evolutionary conserved, some are
255 human-specific (Anzai et al., 2020; Olson and Srivastava, 1996). Therefore, investigation of
256 human cardiac development also requires suitable human models. hiPSC cardiac
257 differentiation models have proved to generate functional cardiac cells and suggested that
258 punctual time points during this differentiation might reflect some key developmental stages
259 (Al Sayed et al., 2021; Kathiriya et al., 2021; Wamstad et al., 2012). However, fully assessing
260 the relevance of hiPSC cardiac differentiation model to study human cardiac development,
261 requires demonstrating that it thoroughly and accurately reproduces human cardiac
262 development in a temporally coordinated fashion. All phenotypic changes that occur during
263 cardiogenesis are known to be embodied by dynamic alterations in cellular transcriptome. Yet,
264 although the ideal situation would be to compare transcriptomic changes along hiPSC
265 differentiation to the ones occurring during human cardiac development, no public human
266 transcriptomic dataset studying well-distributed stages across the entire cardiac development

267 is available. In the present study, we therefore used murine cardiac transcriptomic data
268 generated from specific stages that appropriately rang from conception to birth (Li et al.,
269 2014), to compare with hiPSC cardiac differentiation data. Their high level of consistency
270 confirmed that our hiPSC cardiac differentiation model accurately reproduces cardiogenesis.
271 An important added value of the present data is that it filled a gap of knowledge on the global
272 gene expression changes that occur daily, between these developmental stages in human
273 cells.

274 A major limitation of hiPSC-derived models is immaturity: cardiac cells produced by
275 current hiPSC differentiation protocols have a fetal-like phenotype far from adult cells
276 (Yoshida and Yamanaka, 2017). Although this limitation does not affect the study of prenatal
277 stages of cardiac development, obtaining mature cardiac cells would broaden the scope of
278 these models to study later stages of development as well as aging processes.

279 ***In vitro modeling of cardiac development-associated cellular diversity***

280 Cardiomyocytes require substantial cell diversity to support both the proper execution
281 of their biological functions and their differentiation, since many signaling pathways regulating
282 their formation are sourced from other cell types (Litviňuková et al., 2020; Meilhac, 2018). In
283 this study, we confirmed that hiPSC cardiac differentiation generates the cellular diversity
284 typically reported in the human fetal heart and thus provides the opportunity to investigate
285 regulatory mechanisms occurring between these different cardiac cell types. However, hiPSC
286 cardiac differentiation in two dimensions does not reproduce the spatial organization of the
287 cell types as observed in the context of a heart. The emergence of more integrated hiPSC-
288 derived models such as cardioids (Hofbauer et al., 2021), will therefore undoubtedly enhance
289 our insights into transcriptional regulation between cardiac cell types.

290 ***Uncovering new regulatory networks using a gene expression kinetics-based strategy***

291 An original aspect of this study was the identification of expression regulations
292 occurring between TFs in a temporal manner. For that we chose to adapt the LEAP
293 bioinformatic tool designed for single-cell data to kinetic transcriptomic bulk data (Specht and
294 Li, 2017). Importantly, with this tool, these gene regulations are oriented, indicating not only
295 the interaction but also which partner is expected to be the target/source. This higher level of

296 information is important to design more efficiently confirmation experiment, and cannot be
297 obtain in traditionally-used protein-protein interaction databases, such as STRING (Szkarczyk
298 et al., 2019). Moreover, our strategy allowed to biologically link genes in a time-dependent
299 manner during cardiac differentiation, and thus provided important new insights on cardiac
300 gene regulatory networks (Wang et al., 2020). Of note, one cannot exclude that some of the
301 inferred links may not reflect biological interactions (e.g. TF does not directly bind to an
302 inferred target gene). Other studies embarked in different strategies to study cardiac gene
303 regulation. For instance, Gonzalez-Teran et al. combined, PPI data associated with GATA4 and
304 TBX5 TFs, and genetic data generated on patients presenting congenital heart diseases (CHD)
305 to identify CHD candidate genes (Gonzalez-Teran et al., 2022). This integrated strategy of PPI
306 data and CHD-associated genetic data could be a relevant complementary approach of our
307 chronological gene expression-based strategy in order to identify new CHD-associated TF
308 regulatory networks and to offer a better understanding of cardiac disease underlying
309 mechanisms.

310 ***Activation cascade of GATA4, NKX2-5, TBX5 genes triggered by IRX3 and IRX5***

311 It is well established that cardiac transcription factors regulate the expression of other
312 TF-coding genes. For instance, *GATA4* activates *NKX2-5* expression and both *GATA4* and *NKX2-*
313 *5* activates *TBX5* expression (Bruneau, 2013; McCulley and Black, 2012; Nemer and Nemer,
314 2010; Waardenberg et al., 2014). However, the precise molecular bases of these regulations
315 were still to be uncovered. Using daily-generated transcriptomic data, we characterized the
316 course of expression of these major cardiac TFs showing that, in accordance with the
317 functional data, they are successively launched, starting with *GATA4* around day 5, followed
318 by *NKX2-5* two days later and finally by *TBX5* two days later too. Obviously, this raised the
319 question of how *GATA4* expression is, in the first place, launched. Using gene expression-
320 based network we identified IRX3 and IRX5 TFs as potential activators of *GATA4* expression.
321 Accordingly, the expression of these TFs was launched simultaneously about 1 day prior to
322 *GATA4* expression. These TFs are of growing interest as, while most studies were performed
323 in knockout mice showing that they play redundant roles in cardiac development leading to
324 embryonic lethality, and in postnatal electrophysiological function, their role in human cardiac
325 function now also emerges (Al Sayed et al., 2021; Bonnard et al., 2012). In this context, the

326 present study therefore further explored and specified the role of IRX TFs in the course of
327 human cardiac development.

328 ***Exploring the functional interplay between IRX3/IRX5 and GATA4, NKX2-5, TBX5***

329 It has been shown that GATA4, NKX2-5 and TBX5 act as multiprotein complex to
330 regulate cardiac gene expression (McCulley and Black, 2012). Here, we completed this
331 knowledge by showing that IRX3 and IRX5 can also physically bind to this TF regulatory
332 complex. Furthermore, all five TFs could physically and functionally interact on the promoter
333 of *SCN5A* that encodes the major cardiac sodium ion channel. Accordingly, *SCN5A* expression
334 gradually increases during hiPSC cardiac differentiation, paralleling the progressive expression
335 establishment of the five TFs. Some of the interactions between IRX TFs and GATA4, NKX2-5
336 or TBX5 have previously been published. For instance, physical and functional interactions
337 between *Irx3*, *Nkx2-5* and *Tbx5* have been shown in mice to regulate genes implicated in
338 ventricular conduction system establishment and maturation (Kimura et al., 2016).
339 Furthermore, our group has previously demonstrated physical and functional interactions
340 between IRX5 and GATA4 on *SCN5A* promoter (Al Sayed et al., 2021). In this study we further
341 detailed the complexity of the interactions between IRX3, IRX5, GATA4, NKX2-5 and TBX5, and
342 how these TF complex compositions impact the expression of a target gene.

343 ***Perspectives***

344 Altogether, this study provides a comprehensive dynamic blueprint of transcription
345 factors that control transcriptional regulation during human cardiac development as well as a
346 new methodological approach that may be applied to other research fields. These insights
347 may help to further understand both pathological cardiac development leading to CHD, as
348 well as physiological cardiac development, which is a prerequisite to emerging cardiac
349 regenerative therapy strategies (Rowton et al., 2021). Moreover, in recent years, transcription
350 factor regulation of cardiac functions was widely supported by GWAS, linking numerous
351 common genetic variations at loci harboring TF genes to cardiac diseases (Barc et al., 2022;
352 Roselli et al., 2018; Supplementary Table 1). Confronting the present knowledge to the one
353 obtained from cardiac differentiation of hiPSCs reprogrammed from patients carrying such
354 genetic variants may provide important information regarding their impact on cardiac

355 development and therefore may lead to new targets for treatment and clinical management
356 improvement.

357 ***Limitations of study***

358 We have identified several limitations in our study. First, the kinetic transcriptomic
359 analysis has been performed using bulk-based strategy, however, the use of single-cell
360 analysis instead would have provided us with a better overview of the cellular transcriptomic
361 diversity. Second, the limited sample size that has been used prevented us from identifying
362 the impact of the gender and ethnicity on transcriptomic regulation. Therefore, further studies
363 will have to investigate if the identified TF networks (1) are activated in a cell-specific manner,
364 and (2) whether they are specific to gender and/or ethnicity.

365 **Acknowledgments**

366 We are most grateful to the Genomics Core Facility GenoA, member of Biogenouest
367 and France Genomique and to the Bioinformatics Core Facility BiRD, member of Biogenouest
368 and Institut Français de Bioinformatique (IFB) (ANR-11-INBS-0013) for the use of their
369 resources and their technical support. We acknowledge the IBISA MicroPICell facility
370 (Biogenouest), member of the national infrastructure France-Bioimaging supported by the
371 French national research agency (ANR-10-INBS-04). We thank Jérôme Montnach for helping
372 with muscle contraction analyses, Elise Douillard and Stéphane Minvielle for technical help
373 with single-cell experiments, and Béatrice Guyomarch for helpful discussions on statistical
374 analyses. B.C. is laureate of fellowships from the *Fondation pour la Recherche Médicale* (FRM,
375 PBR201810007614) and the Fondation Genavie. R.C. is laureate of fellowship from the Fonds
376 Marion Elizabeth Brancher. This work was funded by grants from The National Research
377 Agency (HEART iPS ANR-15-CE14-0019-01), from the *Fédération Française de Cardiologie*, and
378 from International Incoming Fellowship FP7-PEOPLE-2012-IIF [PIIF-GA-2012-331436] to N.G,
379 and by a grant from The French Regional Council of Pays de la Loire (RFI project VaCaRMe).

380 **Author contributions**

381 Conceptualization, R.C., B.C., R.R., P.L., N.G. and G.L., Data curation, R.C., B.C. and A.G.,
382 Formal Analysis, R.C. and B.C., Funding acquisition, R.R., P.L., N.G. and G.L., Investigation, R.C.,

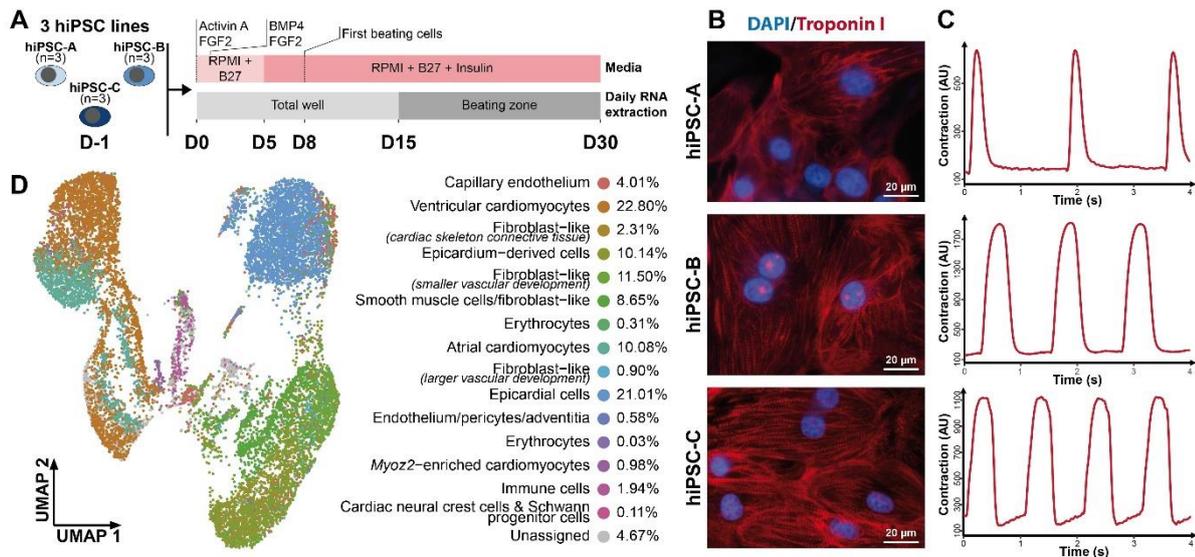
383 B.C., A.G. and V.F., Methodology, R.C., B.C., A.G., N.G. and G.L., Project administration, R.R.,
384 P.L., N.G. and G.L., Software, R.C., B.C. and P.O., Supervision, J.P., R.R., P.L., N.G. and G.L.,
385 Validation, R.C., B.C., A.G., V.F., J.P., R.R., P.L., N.G. and G.L., Visualization, R.C., B.C., Writing –
386 original draft, R.C., B.C., A.G., V.F., J.P., R.R., P.L., N.G. and G.L.

387 **Declaration of interests**

388 The authors declare no competing interests.

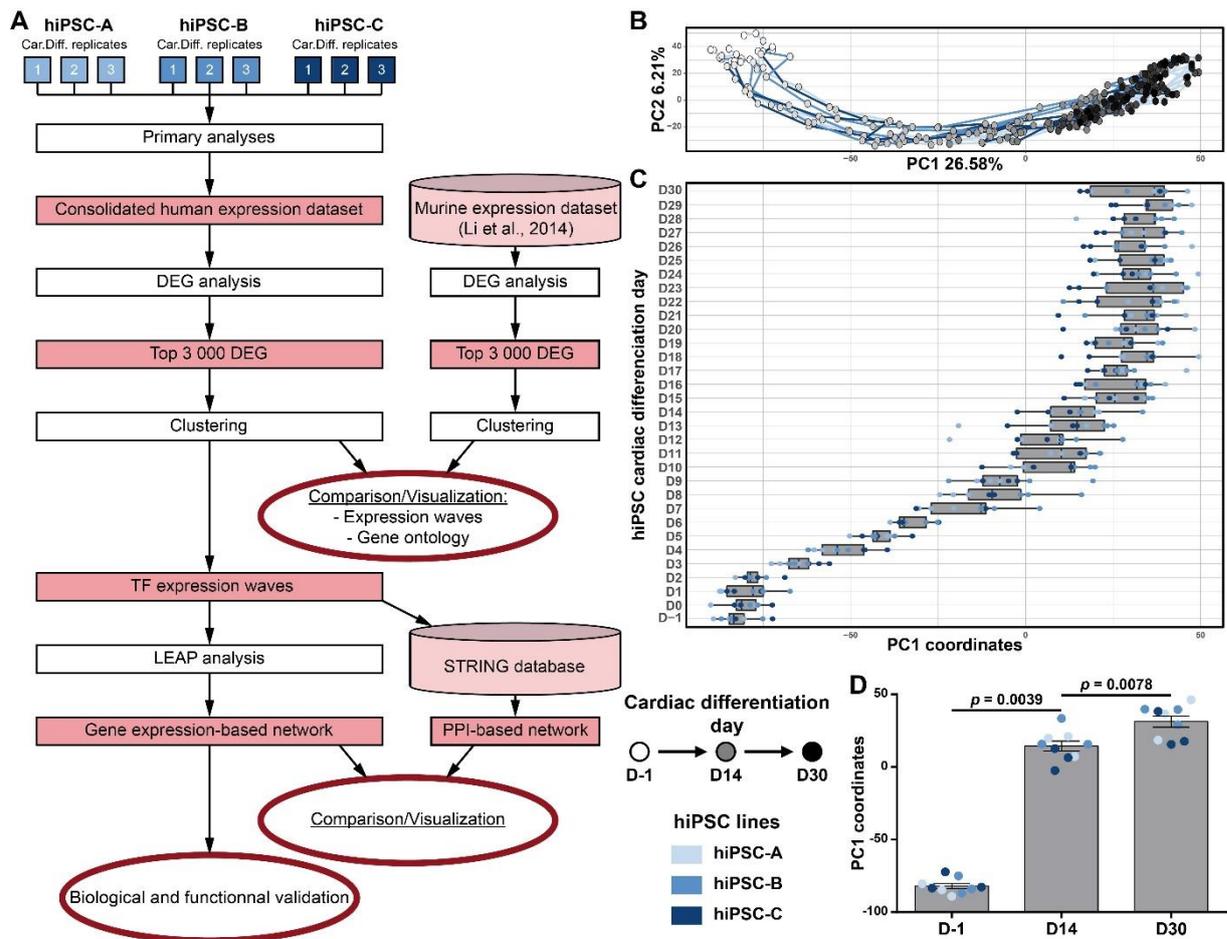
389 **Data and code availability**

390 The authors declare that all data supporting the findings of this study are available
391 within the article and its supplementary information files. Bulk and single-cell transcriptomic
392 data are available in ArrayExpress database at EMBL-EBI
393 (<https://www.ebi.ac.uk/arrayexpress>) under accession number E-MTAB-11822 and E-MTAB-
394 11817, respectively. All data will be publicly available as of the date of publication. All original
395 code has been deposited at Gitlab (<https://gitlab.univ-nantes.fr/E132534J/cardiff.git>) and will
396 be publicly available as of the date of publication. Any additional information required to
397 reanalyze the data reported in this work is available from the lead contact upon reasonable
398 request.

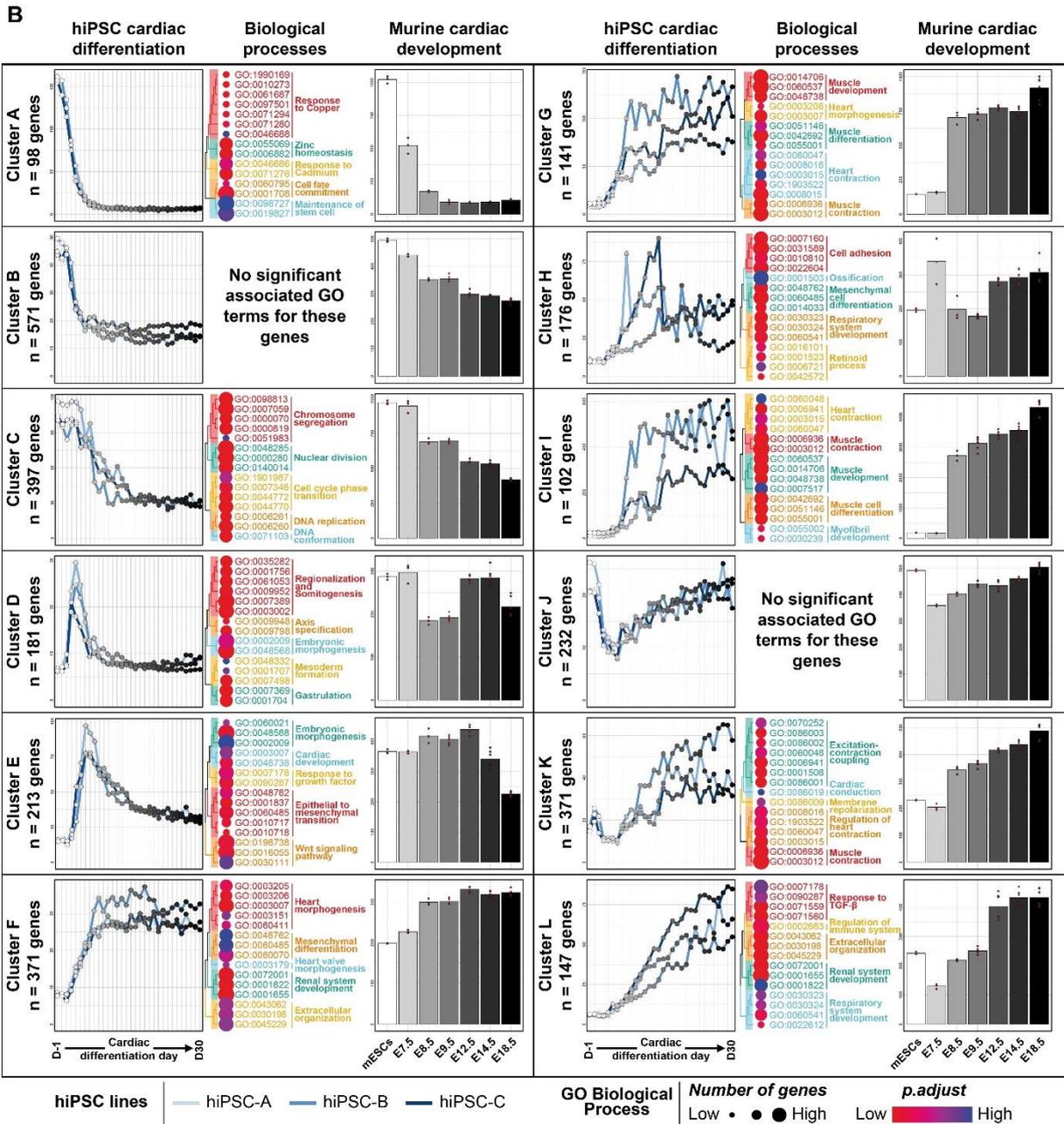
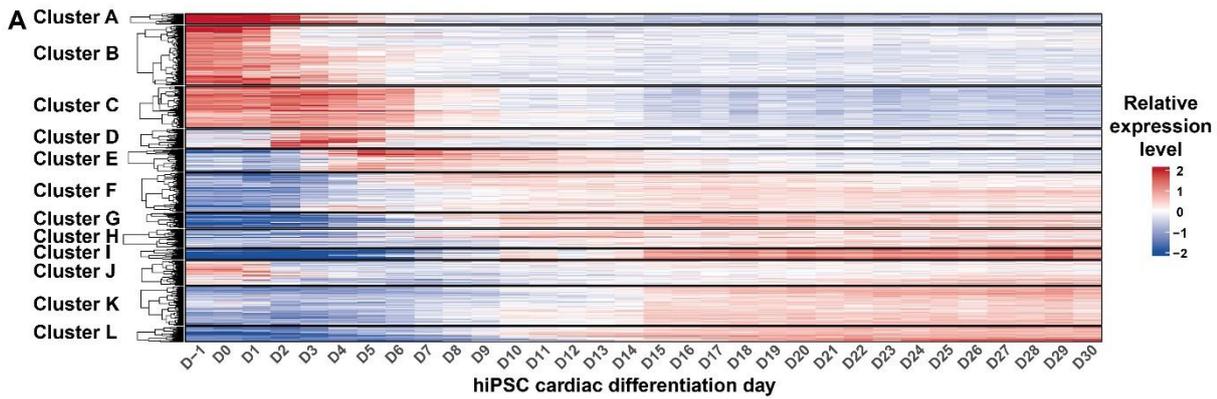


399 **Figure 1 – Transcriptomic and functional characterization of cardiac cells derived from**
 400 **hiPSCs.**

401 (A) Diagram illustrating the experimental design involving three distinct cardiac
 402 differentiations of three hiPSC lines reprogrammed from healthy donors. (B)
 403 Immunocytochemistry staining of troponin I (red) and DAPI (blue) at D30 of cardiac
 404 differentiation for all 3 hiPSC lines. (C) Representative contraction patterns captured by
 405 MUSCLEMOTION software on movies at D30 of cardiac differentiation for the 3 hiPSC lines.
 406 (D) UMAP displaying single-cell RNA-seq data at D30 of cardiac differentiation of hiPSC-A line.
 407 The color code indicates the different cell types identified. Cell population fractions are listed
 408 on the right. See also Figure S1.

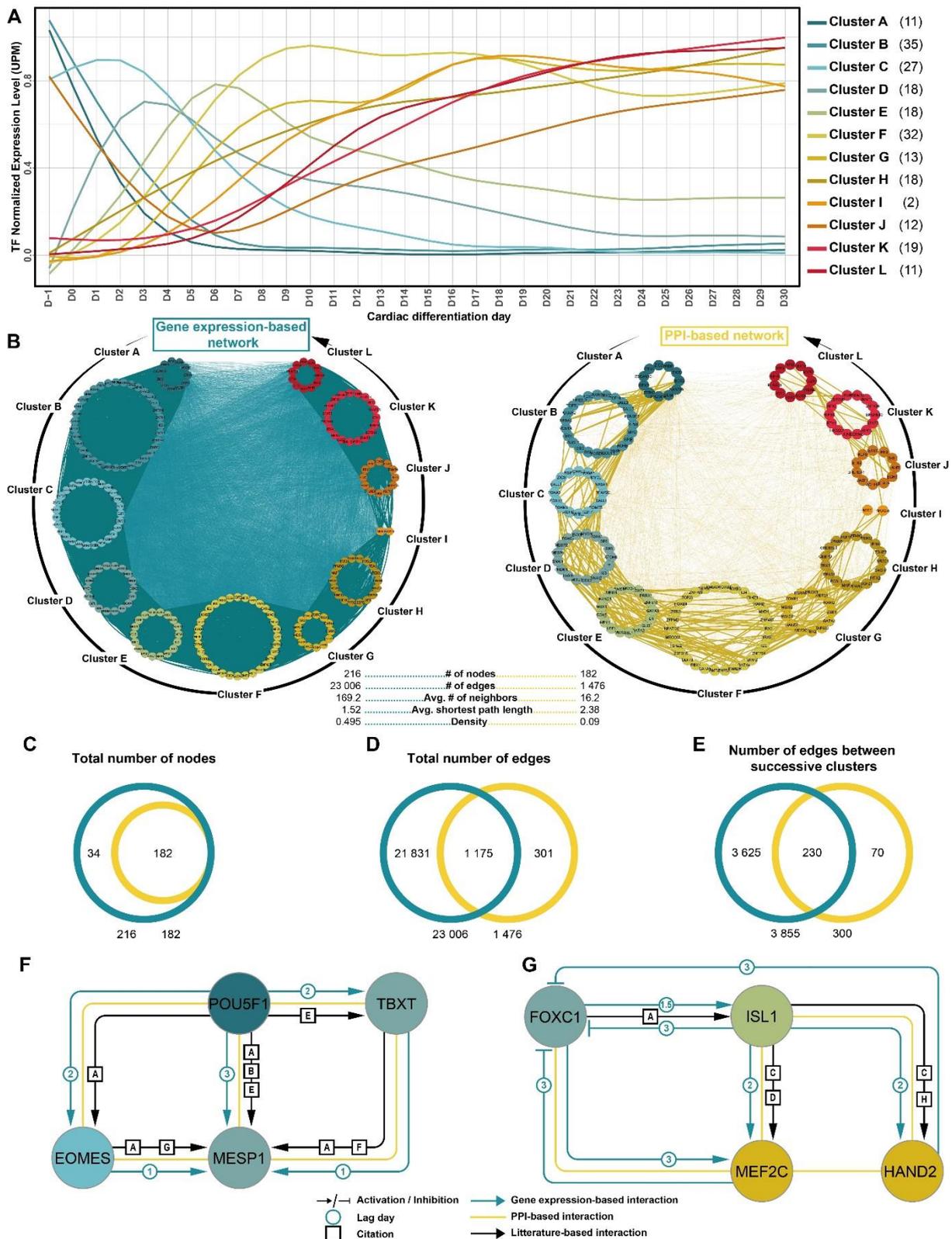


409 **Figure 2 – Transcriptomic time-course analysis of hiPSC cardiac differentiation.**
 410 (A) Methodological workflow. Steps are represented in white rectangle and outputs in red
 411 rectangles. (B) Global transcriptomic variations displayed with the first two components of the
 412 Principal Component Analysis. Three cardiac differentiations were studied for each of the
 413 three hiPSC lines. For each cardiac differentiation, a line connects the time-points in
 414 chronological order. (C) Boxplots displaying the distribution of PC1 coordinates of each
 415 replicates at each day (median +/- quartile). (D) Histogram comparing distribution of PC1
 416 coordinates at the beginning (D-1), the middle (D14) and the end (D30) of hiPSC cardiac
 417 differentiations (Mean +/- SEM; Wilcoxon matched-pairs signed rank test).



418 **Figure 3 – Expression profile and functional annotation of the top 3 000 differentially**
 419 **expressed genes (DEGs) during hiPSC cardiac differentiation, and comparison with murine**
 420 **cardiac development gene expression dataset.**

421 (A) Heatmap displaying DEG expression levels. The entire data set was used to aggregate the
422 genes into 12 clusters and the mean expression level of 9 replicates is represented. (B) For
423 each cluster, average gene expression level during hiPSC cardiac differentiation (left panel for
424 each cluster) and of their orthologs during murine cardiac development (mESCs to E18.5 stage,
425 right panel for each cluster) are shown. Replicate gene expression levels were averaged for
426 each hiPSC line (n=3 per hiPSC line and per timepoint) and for murine data (n=3 to n=6 per
427 timepoint, depending on the stage). The 15 most significantly related GO terms are displayed
428 for each cluster on the middle panel. See also Figure S2.



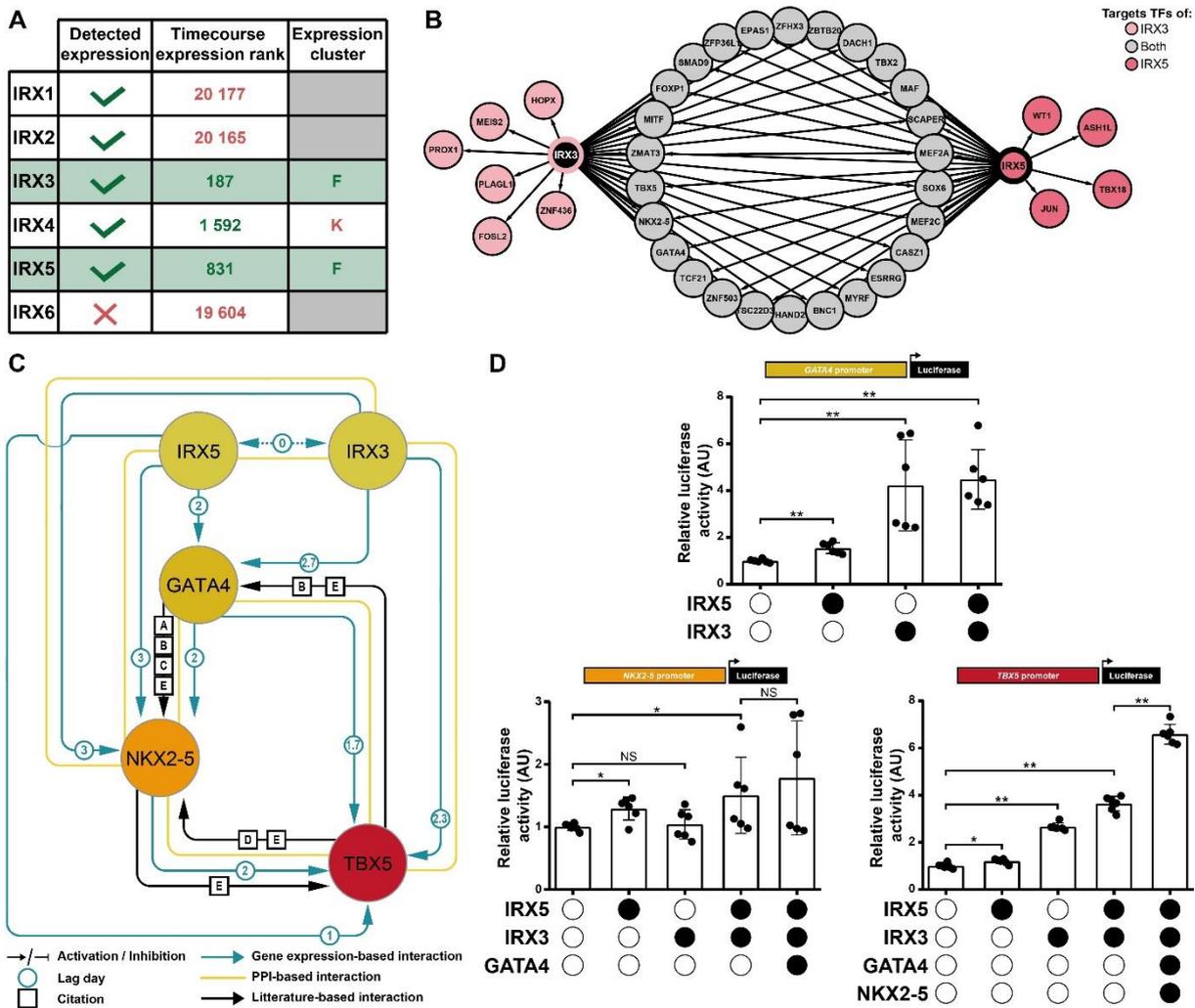
429 **Figure 4 – Inferred TF regulatory network governing hiPSC cardiac differentiation.**

430 (A) Normalized gene expression of the 216 TFs (identified within the top 3 000 differentially

431 expressed genes during hiPSC cardiac differentiation) were quantified and averaged in each

432 gene cluster. UPM: UMI per million. The number in () indicates the TF number per cluster. (B)

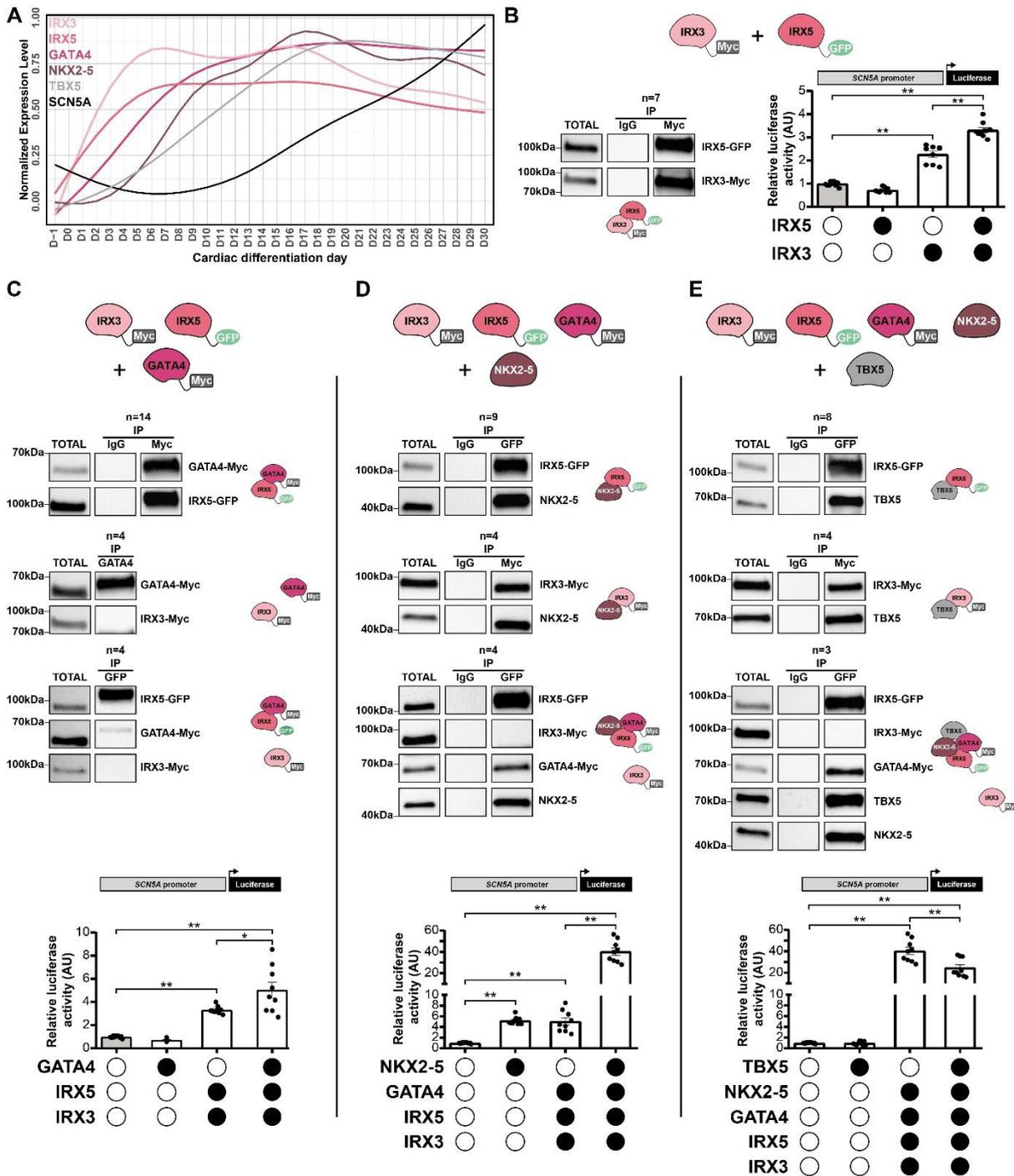
433 Graphical representation of gene expression-based network and protein-protein interaction-
434 based network (LEAP- and STRING- based method, respectively) of the same TFs as in A.
435 Interactions between TFs of successive clusters are shown using bold lines. (C-E) Comparative
436 quantitative analysis between both networks. (F-G) Examples of two literature-based sub-
437 networks. Interactions uncovered in gene expression-based network are shown in blue, in PPI-
438 based network, in yellow, and by literature curation, in black. Node colors correspond to the
439 one of their corresponding gene cluster (as in A). Paper PMID associated with literature-based
440 links: [A] 33803261; [B] 23417899; [C] 34901033; [D] 15253934; [E] 17011492; [F] 21632880;
441 [G] 22402664; [H] 22438573.



442 **Figure 5 – Exploration of the inferred IIGNT sub-network.**

443 (A) Table of TF selection criteria for the IRX genes family. Timecourse expression rank is the
444 output of the timecourse package and illustrates the variation in gene expression during
445 directed cardiac differentiation (a lower number indicating a higher variation). The Expression
446 cluster column refers at the expression cluster in which each TF ranged as in Figure 4A.
447 Selection and exclusion criteria are indicated in green and red respectively. (B) Potential target
448 TFs of IRX3 and/or IRX5 identified in the G-to-L clusters based on the gene expression-based
449 network. (C) Gene expression-based network of IRX3, IRX5, GATA4, NKX2-5 and TBX5 TFs.
450 Node colors represent their corresponding clusters as in Figure 4A: IRX3 and IRX5 – cluster F;
451 GATA4 – cluster G; NKX2-5 – cluster I; TBX5 – cluster L. Lag is shown in days. References to
452 literature-based links: [A] Book chapter DOI: 10.1016/B978-0-12-381332-9.00027-X.; [B]
453 PMID: 23457256; [C] PMID: 22449847; [D] PMID: 32450132; [E] PMID: 25280899. (D) Graphs
454 displaying activity levels of luciferase that is under the control of GATA4 (-1800_TSS_+200),

455 NKX2-5 (-2000bp_Start codon) and TBX5 (-1800_TSS_+200) promoter constructs. Mean +/-
456 SD; * and **: $p < 0.05$ and $p < 0.01$, respectively (Mann-Whitney test).



457 **Figure 6 – Physical and functional interactions of IRX3/IRX5/GATA4/NKX2-5/TBX5**
 458 **multiprotein complex.**

459 (A) Normalized mean expression level overtime of IRX3, IRX5, GATA4, NKX2-5, TBX5 and of
 460 SCN5A ion channel genes. UPM: UMI per million. (B-E) Co-immunoprecipitation and luciferase
 461 results associated to the transfection of (B) IRX3 and/or IRX5, (C) IRX3, IRX5 and/or GATA4, (D)
 462 IRX3, IRX5, GATA4 and/or NKX2-5, (E) IRX3, IRX5, GATA4, NKX2-5 and/or TBX5. Immunoblots
 463 representative of the various co-immunoprecipitations and the schematic illustration of the

464 results. Graphs display activity levels of luciferase that is under the control of the -2109/+1072
465 region of human SCN5A promoter, in the various transfection conditions. Mean +/- SEM; *
466 and **: $p < 0.05$ and $p < 0.0001$, respectively (Mann-Whitney test). See also Figure S3.

467 **Materials and Methods**

468 ***Reprogramming and maintenance of hiPSCs***

469 All cell lines, from 3 healthy donors, were previously characterized. The hiPSC-A (C2a
470 in Si-Tayeb et al., 2010) line was generated using lentivirus method while hiPSC-B (IRX5-Wt in
471 Canac et al., 2022; RRID:CVCL_B5QD) and hiPSC-C (WT8288 in Girardeau et al., 2021;
472 RRID:CVCL_B5Q5) lines were generated using Sendai virus method. hiPSC lines were
473 maintained at 37°C, 5% CO₂, 21% O₂ in StemMACSTM iPS Brew XF Medium (Miltenyi Biotec)
474 on culture plates coated with Matrigel® hESC-Qualified Matrix (0.05 mg/mL, Corning). At 75%
475 confluency, cells were passaged using Gentle Cell Dissociation Reagent (STEMCELL™
476 Technologies).

477 ***Cardiac differentiation of hiPSCs***

478 Directed cardiac differentiations of hiPSCs were performed using the established
479 matrix sandwich method (Figure 1A; Es-Salah-Lamoureux et al., 2016). Briefly, when hiPSCs
480 reached 90% confluency, an overlay of Growth Factor Reduced Matrigel (0.033 mg/ml, BD
481 Corning) was added. Differentiation was initiated 24h later by culturing the cells in RPMI1640
482 medium (Life Technologies) supplemented with B27 (without insulin, Life Technologies), 2 mM
483 L-glutamine (Life Technologies), 1% NEAA (Life Technologies), 100 ng/mL Activin A (Miltenyi
484 Biotec), 1X Pen/Strep (Life Technologies) and 10 ng/mL FGF2 for 24 hours. On the next day,
485 the medium was replaced by RPMI1640 medium supplemented with B27 without insulin, 2
486 mM L-glutamine, 1% NEAA, 10 ng/mL BMP4 (Miltenyi Biotec), 1X Pen/Strep and 5 ng/mL FGF2
487 for 4 days. By day 5, cells were cultured in RPMI1640 medium supplemented with B27
488 complete (Life Technologies), 2 mM L-glutamine, 1X Pen/Strep and 1% NEAA and changed
489 every two days until day 30. Specifically, for video analysis and immunofluorescence staining,
490 glucose starvation was performed to obtain purified cardiomyocyte population: at day 10 the
491 medium was replaced by Depletion medium (RPMI 1640 medium without glucose (Life
492 Technologies) supplemented with B27 complete, and 1X Pen/Strep) for 3 days. Cells were
493 dissociated at day 13 with 10X TrypLE solution (Life Technologies) and replated in CMs
494 medium (RPMI1640 medium supplemented with B27 complete, 2 mM L-glutamine, 1X
495 Pen/Strep, 1% NEAA) supplemented with Y-27632 Rho-kinase inhibitor (STEMCELL™

496 Technologies). On day 14, the medium was replaced by Depletion medium for 3 days. From
497 day 17, cells were maintained in CMs medium.

498 ***Bulk transcriptomics***

499 *RNA extraction and sequencing*

500 For each hiPSC line, samples were harvested daily from D-1 to D30 of the cardiac
501 differentiation protocol from three independent cardiac differentiations. Total RNA were
502 extracted using the NucleoSpin RNA kit (MACHEREY-NAGEL) and their quality assessed by
503 NanoDrop™ 1000 Spectrophotometer (Thermo Fisher Scientific). From D-1 to D14 samples,
504 all cells were collected while, from D15 to D30, to obtain samples enriched with
505 cardiomyocytes, only spontaneously beating cell clusters were collected following mechanical
506 isolation using a needle. 3'RNA libraries were prepared by GenoBiRD core facility according to
507 their published method (Charpentier et al., 2021) and sequenced on 8 individual runs on a
508 NovaSeq 6000 or HiSeq 2500 Sequencing System (Illumina).

509 *Primary analysis of bulk transcriptomic data*

510 Demultiplexing, alignment on GRCh38 reference genome and counting steps were
511 conducted on each sequencing run with the Snakemake pipeline developed by the GenoBiRD
512 core facility (Charpentier et al., 2021). Normalized and log-transformed expression matrices
513 were generated using the *multiplates* function correcting potential batch effects by treating
514 cardiac differentiation time points as replicates.

515 *PCA*

516 Principal Component Analysis (PCA) was performed with the R package *FactoMineR*
517 (Lê et al., 2008; RRID:SCR_014602) on the entire mean-centered and log-transformed matrix.

518 *Time-course gene expression analysis*

519 Genes with significant expression variation between the different cardiac
520 differentiation time-points (indicated as Differentially Expressed Genes; DEG) were identified
521 by multivariate empirical Bayes statistics using the R package *timecourse* (Tai, 2022;
522 RRID:SCR_000077) applied to the entire log-transformed matrix. We selected the top 3 000

523 DEG based on their highest Hotelling \tilde{T}^2 statistics. The same method was used to select genes
524 with significant expression variation during murine cardiac development from a published
525 transcriptomic dataset (Li et al., 2014). When necessary, human and murine orthologous gene
526 names were identified using the R package *biomaRt* (Durinck et al., 2009; RRID:SCR_019214)
527 and Ensembl databases.

528 Clustering and heatmap

529 DEG were grouped into clusters, based on their expression level variation across the
530 288 samples, using the R function *k-means* set on 2 000 iterations, and visualized with the R
531 package *ComplexHeatmap* (Gu et al., 2016; RRID:SCR_017270).

532 Gene Ontology analyses

533 Gene Ontology (GO) analysis was performed using the R package *ClusterProfiler* (Wu
534 et al., 2021; RRID:SCR_016884), based on GO Biological Process terms from
535 org.Hs.eg.db_3.14.0 and org.Mm.eg.db_3.14.0 databases for human and mouse annotations,
536 as appropriate. Significantly enriched (bonferroni-corrected p-value < 0.05) biological
537 processes, as compared to reference transcriptome, and with a Gene Set Size (GSSize)
538 between 10 and 500, were considered for further analysis. The 15 GO terms with the lowest
539 corrected p-value were visualized with treeplot.

540 Network construction and analysis

541 For each hiPSC line, the gene regulatory network was inferred using the R package
542 *LEAP* (Lag-based Expression Association for Pseudotime-series; Specht and Li, 2017), based on
543 the average from the log-transform data of triplicate cardiac differentiations. Cardiac
544 differentiation time points were used to rank samples as required by the *LEAP* tool. The
545 *max_lag_prop* parameter was set to 1/10, meaning that, at most, 3-day windows were used
546 to calculate the maximum absolute correlation (MAC) score. Only links with a significant MAC
547 score (determined by a permutation test; *p-value* < 0.05) and related to a non-null time delay
548 were considered. Links with a positive correlation score were interpreted as activation
549 relationships and those with a negative correlation score as repression relationships. STRING
550 software (Szklarczyk et al., 2019) was used to obtain information on physical and functional

551 interactions between proteins of interest, with a *minimum required interaction score* of 0.4.
552 Nodes without any interaction were excluded. STRING-based or LEAP-based interactions were
553 processed using Cytoscape 3.9.1 for network reconstruction (Shannon et al., 2003;
554 RRID:SCR_003032). Networks parameters were obtained using the *Analyze network* function.

555 ***Single-cell transcriptomic***

556 *Single-cell RNA-seq data generation*

557 Cells at D30 of hiPSC-A cardiac differentiation were harvested from three distinct
558 beating wells, dissociated, using the Multi Tissue Dissociation Kit 3 (Miltenyi Biotec), and
559 pooled. This experiment was performed in duplicates. Cell suspensions were filtered on a 40
560 µm Flowmi® Cell Strainer, counted and cell viability was assessed (viability was 92% for the
561 first experiment and 94% for the second). For each replicate, single-cell droplet libraries were
562 generated from 16 000 cells with the Chromium Single Cell 3' GEM, Library & Gel Bead Kit v3
563 (10X Genomics). After qPCR quantification, libraries were pooled and sequenced on a single
564 run, on a NovaSeq 6000 Sequencing System (Illumina), providing a read depth of >20,000 read
565 pairs per cell, according to manufacturer's instructions.

566 *Primary analysis of single-cell transcriptomic data*

567 Data were processed using *cellranger* 4.0.0 (10X Genomics). First, demultiplexing of
568 *raw base call* files into FASTQ files was accomplished using *cellranger mkfastq* function.
569 Second, alignment on GRCh38 reference genome, filtering and counting steps were
570 performed separately on each replicate with *cellranger count* function. Lastly, aggregation
571 with normalization of duplicates was performed using *cellranger aggr* function.

572 *Secondary analysis of single-cell transcriptomic data*

573 The gene expression matrix was analyzed using the R package *Seurat* (Stuart et al.,
574 2019; RRID:SCR_016341). Doublets were identified and removed using the R package
575 *DoubletFinder* (McGinnis et al., 2019; RRID:SCR_018771), assuming a 7.5% doublet formation
576 rate. Also, only cells with 200 to 5 000 detected features and with <25% reads aligned to
577 mitochondrial genes were selected for further analysis. After normalization, unwanted
578 sources of intercellular variations such as number of detected genes or differences between

579 cell cycle phases were regressed using the *ScaleData* function. A principal component analysis
580 was then performed using the 2 000 most variable genes according to the
581 *FindVariableFeatures* function and the first 10 components were used to calculate the UMAP.
582 Cell-type labelling was performed using published single-cell RNA-seq data from a human fetal
583 heart as a reference (Asp et al., 2019). Cell-type labels from reference were automatically
584 transferred after cell-to-cell matching at the individual cell level using the R package *CellID*
585 (Cortal et al., 2021).

586 ***Muscle motion***

587 hiPSC-CMs were filmed after glucose starvation at D30 in routine culture condition
588 (37°C, 5% CO₂), without electrical stimulation, using Nikon A1 RSI confocal microscope with
589 X20 Dry N.A 0.75 objective. MUSCLEMOTION software (*Gaussian Blur: No; Speed Window: 5;*
590 *Noise Reduction: Yes; Automatic Reference Frame Detection: Yes; Transient analysis: Yes; Sala*
591 *et al., 2018*) was used to obtain contraction traces from 120fps videos. Contraction profiles
592 were analyzed using homemade R pipeline.

593 ***HEK293 cell culture and transfection***

594 HEK293 cells were maintained at 37°C, 5% CO₂, in DMEM media with 10% FBS, 5% L-
595 Glutamine and 5% Pen/Strep. Cells were plated in 24-well plate or 6-well plate and transfected
596 next day using FuGENE[®] 6 (Promega, E2691). For luciferase assay, cells were transfected with
597 a total of 2µg of plasmid including: (1) pGL2-Renilla plasmid, (2) plasmid containing Firefly
598 luciferase gene upstream promoter of interest and (3) expression plasmids coding for proteins
599 of interest (Table 1). DNA quantities were equalized in each condition using empty pcDNA3.1
600 plasmid. Media was changed 24h post transfection, and cell lyses performed 48h post
601 transfection. For co-immunoprecipitation, cells were transfected only with expression
602 plasmids prior lysis 24h post transfection.

603 ***Co-immunoprecipitation***

604 *Protein sample extraction and quantification*

605 Previously transfected HEK293 cells were lysed (4°C, 15min, with rotation) in lysis
606 buffer: 1% TritonX-100, 100mM NaCl, 50mM Tris-HCl, 1mM EGTA, 1mM Na₃VO₄, 50mM NaF,

607 1mM phenylmethylsulfonyl fluoride, protease inhibitors cocktail (Sigma-Aldrich, P8340), and
 608 centrifuged at 15 000g (4°C, 15 min). Protein quantification was carried out using Pierce™ BCA
 609 Protein Assay Kit (Thermo Fisher, 23225).

610 *Bead-antibody complexes preparation*

611 Co-Immunoprecipitation was performed using Dynabeads® Protein G (Invitrogen,
 612 10004D) and DynaMag™-2 Magnet (Invitrogen, 12321D). First, 12.5µL of beads were
 613 conjugated (Room temperature (RT), 40min, with rotation) with 2µg of antibody (Table 1). The
 614 bead-antibody complexes were cross-linked (RT, 30min, with rotation) using 5,4mg/ml
 615 dimethyl pimelimidate (ThermoScientific, 21667). The cross-linking was quenched with 50mM
 616 Tris pH7.5 (RT, 15min, with rotation). Beads were washed using (1) PBS 1X, (2) 0.1M citrate
 617 pH3.1, (3) Na-phosphate solutions, then incubated in PBS 0.5% NaDOC (RT, 15min, with
 618 rotation) and were finally washed with lysis buffer.

619 *Immunoprecipitation and western blotting analysis*

620 Bead-antibody complexes were incubated with 1mg protein samples (4°C, 2h, with
 621 rotation). Supernatant was then discarded and beads were washed 3 times with lysis buffer.
 622 Beads-protein complexes were then heated (50°C, 10min) in NuPAGE™ LDS Sample Buffer (4X)
 623 (Invitrogen, NP0008). Samples were magnetized prior supernatants collection and incubated
 624 (70°C, 10min) in NuPAGE® Sample Reducing Agent 10X (Invitrogen, NP0009). Finally, samples
 625 were loaded onto a 4–15% precast polyacrylamide gel (Biorad, 4568083) together with 10µg
 626 of total protein used as control. Revelation was performed using corresponding antibody
 627 (Table 1) with ECL Clarity Max (Biorad, 1705062). Images were acquired with ChemiDoc
 628 camera (Biorad) and analysed using Image Lab Software (Biorad).

629 **Table 1. Plasmids and antibodies references**

Plasmid Name	Sequence/Reference	Supplier
Nkx2.5 promoter - FireflyLuc	-2000bp_Start codon	Vectorbuilder
GATA4 promoter - FireflyLuc	-1800_TSS_+200	Vectorbuilder
Tbx5 promoter - FireflyLuc	-1800_TSS_+200	Vectorbuilder
SCN5A promoter - FireflyLuc	-2109_TSS_+1072	Adapted from Al Sayed et al., 2021
pGL2 Renilla luciferase		Promega
IRX5	RG234228	Origene
IRX3	RG205722	Origene
GATA4	RC210945	Origene

Tbx5	SC120046	Origene	
Nkx2.5	SC122678	Origene	
pcDNA3.1		Invitrogen	
Antibody	Reference	RRID	Supplier
anti-GFP	TA150041	AB_2622256	Origene
anti-Myc Tag	05-724	AB_309938	Merck Millipore
anti-IRX5	sc-81102	AB_1124818	Santa Cruz
anti-IRX3	sc-166877	AB_10609525	Santa Cruz
anti-GATA4	sc-25310	AB_627667	Santa Cruz
anti-Tbx5	sc-515536		Santa Cruz
anti-Nkx2.5	sc-8697	AB_650280	Santa Cruz
anti-Troponin I	sc-15368	AB_793465	Santa Cruz
Mouse IgG Isotype Control	02-6502	AB_2532951	Thermo Fisher Scientific

630 ***Luciferase assay***

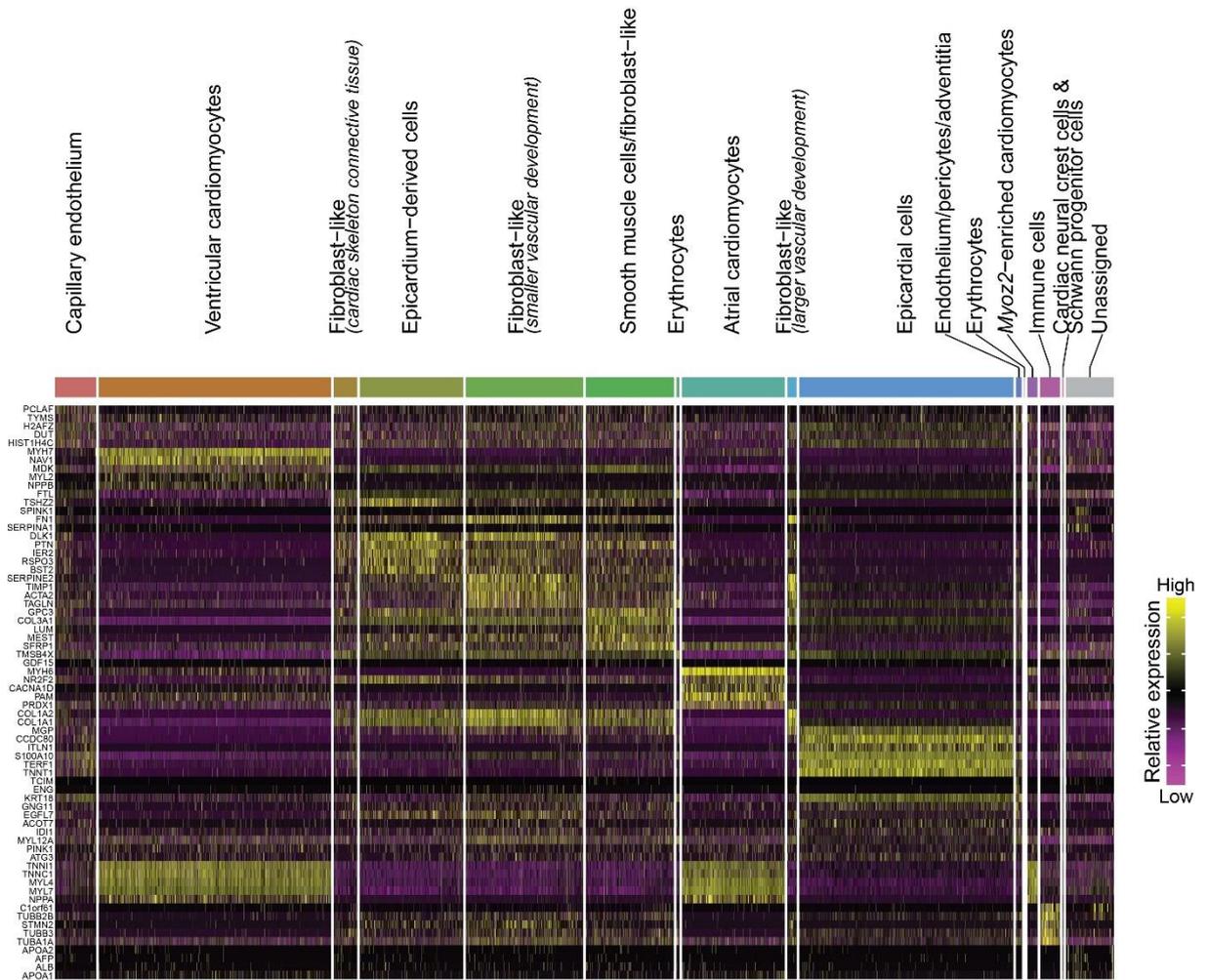
631 Cells were lysed according to the manufacturer recommendations and luciferase
632 activity was measured using Dual Luciferase reporter assay system (Promega, E1910) with
633 Varioskan™ LUX microplate reader (Thermofisher). Mann-Whitney statistical tests were
634 performed with Prism software (v8.0.1).

635 ***Immunofluorescence***

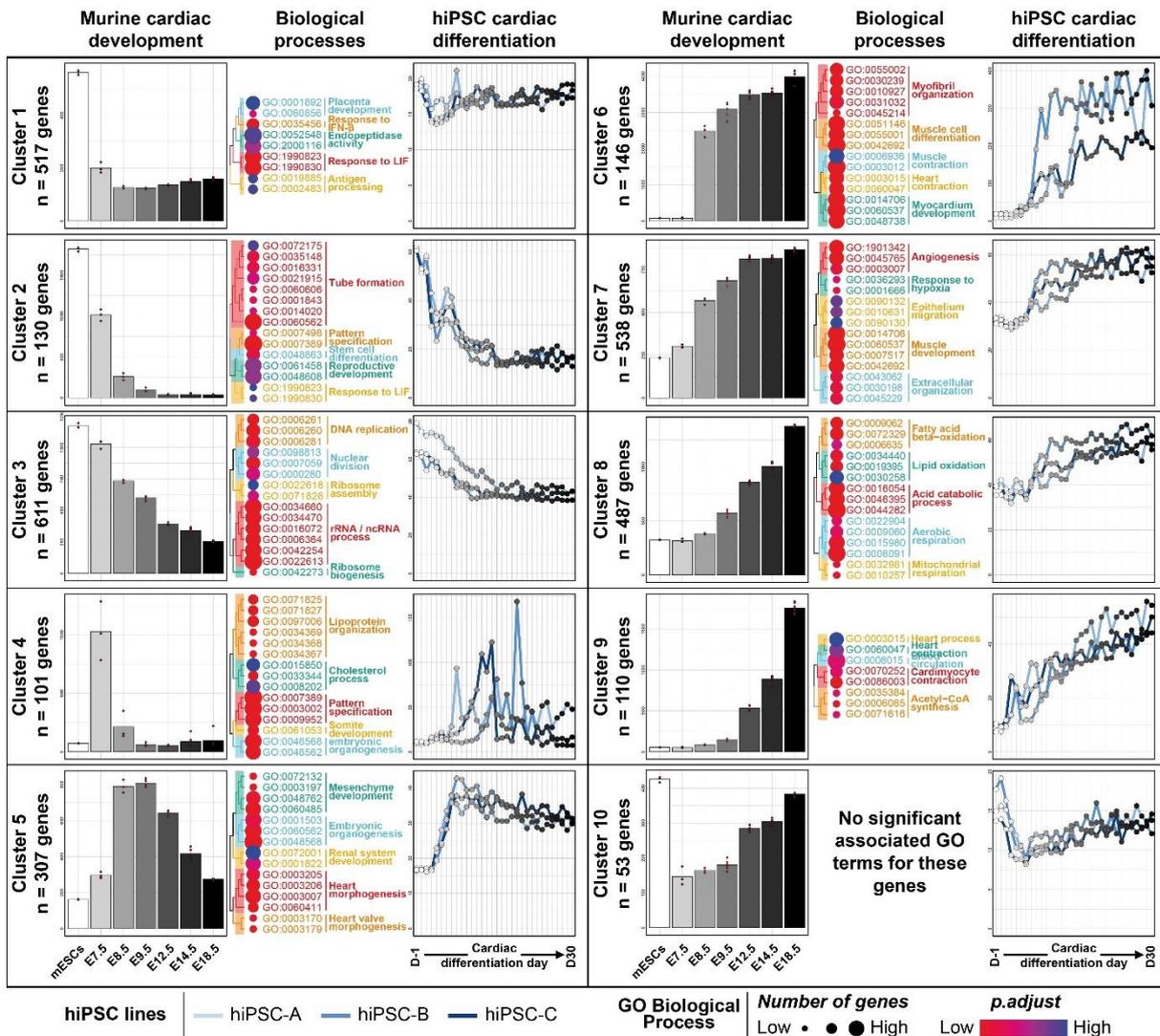
636 Cells were fixed with 4% paraformaldehyde for 15min at room temperature (RT) in
637 Matrigel®-coated μ -Slide 8 Well (IBIDI) prior permeabilization with 0.1% PBS-BSA 1% Saponin
638 (RT, 15min) and blocking with 3% PBS-BSA (RT, 30min). Cells were then incubated with primary
639 antibodies (dilution 1/250) in PBS 0.1% BSA 0.1% Saponin solution (4°C, overnight). Finally,
640 cells were washed and incubated with secondary antibodies and DAPI (RT, 1h) and stored in
641 0.5% paraformaldehyde (4°C). Images were acquired using an inverted epifluorescence
642 microscope (Zeiss Axiovert 200 M).

643 ***TF and cardiac phenotypes association***

644 The association between cardiac phenotypes and transcription factors was performed
645 using the DisGeNET (v7.0; Piñero et al., 2020) and NHGRI-EBI GWAS Catalog (Buniello et al.,
646 2019) databases, filtering on cardiovascular traits, which were then manually validated.

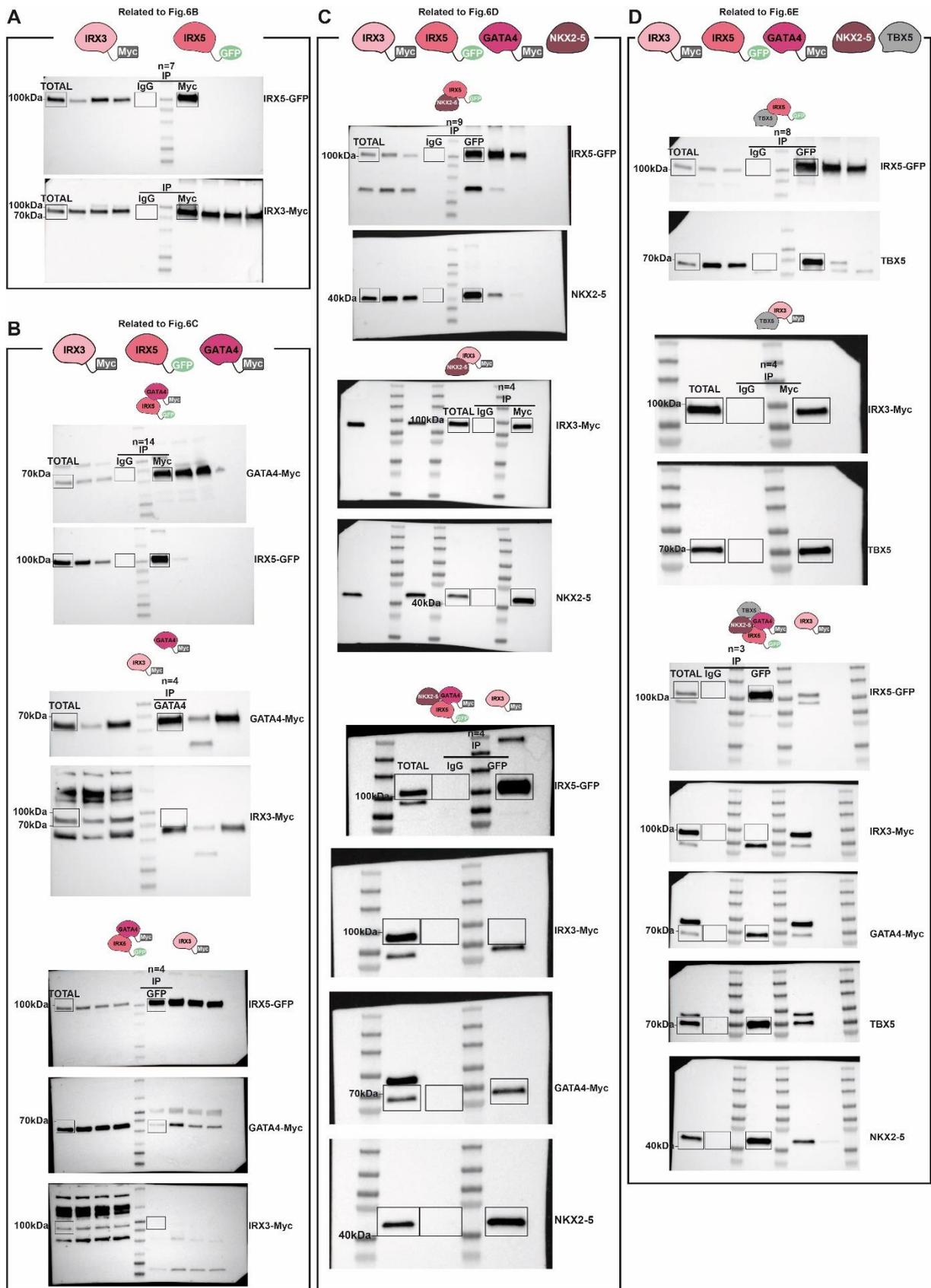


647 **Figure S1 – Transcriptomic signatures of cell types generated by hiPSC cardiac**
 648 **differentiation.**
 649 Heatmap displaying expression levels of the top 5 markers of each cell population found in
 650 Figure 1D. Markers were identified using the FindAllMarkers function from the R package
 651 Seurat.



652 **Figure S2 – Expression profile and functional annotation of the top 3 000 differentially**
 653 **expressed genes during murine cardiac development, and comparison with hiPSC cardiac**
 654 **differentiation gene expression dataset.**

655 For each cluster, average gene expression level during murine cardiac development (left panel
 656 for each cluster) and their human orthologs during hiPSC cardiac differentiation (right panel
 657 for each cluster) are shown. Replicates of gene expression levels were averaged for murine
 658 data (n=3 to n=6 per timepoint, depending on the stage) and hiPSC cardiac differentiation (n=3
 659 per hiPSC line and per timepoint). The 15 most significantly related GO terms are displayed on
 660 the middle panel for each cluster.



661 Figure S3 – Original western blots related to Figure 6.

662 **References**

- 663 Al Sayed, Z.R., Canac, R., Cimarosti, B., Bonnard, C., Gourraud, J.-B., Hamamy, H., Kayserili, H.,
664 Girardeau, A., Jouni, M., Jacob, N., et al. (2021). Human model of IRX5 mutations reveals key
665 role for this transcription factor in ventricular conduction. *Cardiovasc. Res.* 117, 2092–2107.
666 <https://doi.org/10.1093/cvr/cvaa259>.
- 667 Ang, Y.-S., Rivas, R.N., Ribeiro, A.J.S., Srivas, R., Rivera, J., Stone, N.R., Pratt, K., Mohamed,
668 T.M.A., Fu, J.-D., Spencer, C.I., et al. (2016). Disease Model of GATA4 Mutation Reveals
669 Transcription Factor Cooperativity in Human Cardiogenesis. *Cell* 167, 1734-1749.e22.
670 <https://doi.org/10.1016/j.cell.2016.11.033>.
- 671 Anzai, T., Yamagata, T., and Uosaki, H. (2020). Comparative Transcriptome Landscape of
672 Mouse and Human Hearts. *Front. Cell Dev. Biol.* 8, 268.
673 <https://doi.org/10.3389/fcell.2020.00268>.
- 674 Arnolds, D.E., Liu, F., Fahrenbach, J.P., Kim, G.H., Schillinger, K.J., Smemo, S., McNally, E.M.,
675 Nobrega, M.A., Patel, V.V., and Moskowitz, I.P. (2012). TBX5 drives Scn5a expression to
676 regulate cardiac conduction system function. *J. Clin. Invest.* 122, 2509–2518.
677 <https://doi.org/10.1172/JCI62617>.
- 678 Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wårdell, E., Custodio, J.,
679 Reimegård, J., Salmén, F., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and
680 Cell Atlas of the Developing Human Heart. *Cell* 179, 1647-1660.e19.
681 <https://doi.org/10.1016/j.cell.2019.11.025>.
- 682 Barc, J., Tadros, R., Glinge, C., Chiang, D.Y., Jouni, M., Simonet, F., Jurgens, S.J., Baudic, M.,
683 Nicastro, M., Potet, F., et al. (2022). Genome-wide association analyses identify new Brugada
684 syndrome risk loci and highlight a new mechanism of sodium channel regulation in disease
685 susceptibility. *Nat. Genet.* 54, 232–239. <https://doi.org/10.1038/s41588-021-01007-6>.
- 686 Bonnard, C., Strobl, A.C., Shboul, M., Lee, H., Merriman, B., Nelson, S.F., Ababneh, O.H., Uz, E.,
687 Güran, T., Kayserili, H., et al. (2012). Mutations in IRX5 impair craniofacial development and
688 germ cell migration via SDF1. *Nat. Genet.* 44, 709–713. <https://doi.org/10.1038/ng.2259>.
- 689 Briggs, L.E., Takeda, M., Cuadra, A.E., Wakimoto, H., Marks, M.H., Walker, A.J., Seki, T., Oh,
690 S.P., Lu, J.T., Sumners, C., et al. (2008). Perinatal Loss of Nkx2-5 Results in Rapid Conduction
691 and Contraction Defects. *Circ. Res.* 103, 580–590.
692 <https://doi.org/10.1161/CIRCRESAHA.108.171835>.
- 693 Bruneau, B.G. (2013). Signaling and transcriptional networks in heart development and
694 regeneration. *Cold Spring Harb. Perspect. Biol.* 5, a008292.
695 <https://doi.org/10.1101/cshperspect.a008292>.
- 696 Bruneau, B.G., Bao, Z.Z., Fatkin, D., Xavier-Neto, J., Georgakopoulos, D., Maguire, C.T., Berul,
697 C.I., Kass, D.A., Kuroski-de Bold, M.L., de Bold, A.J., et al. (2001). Cardiomyopathy in Irx4-
698 deficient mice is preceded by abnormal ventricular gene expression. *Mol. Cell. Biol.* 21, 1730–
699 1736. <https://doi.org/10.1128/MCB.21.5.1730-1736.2001>.

- 700 Buniello, A., MacArthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C.,
701 McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog
702 of published genome-wide association studies, targeted arrays and summary statistics 2019.
703 *Nucleic Acids Res.* 47, D1005–D1012. <https://doi.org/10.1093/nar/gky1120>.
- 704 Canac, R., Caillaud, A., Cimarosti, B., Girardeau, A., Hamamy, H., Reversade, B., Bonnard, C., Al
705 Sayed, Z.R., David, L., Poschmann, J., et al. (2022). Generation of three human induced
706 pluripotent stem cell lines with IRX5 knockout and knockin genetic editions using CRISPR-Cas9
707 system. *Stem Cell Res.* 58, 102627. <https://doi.org/10.1016/j.scr.2021.102627>.
- 708 Charpentier, E., Cornec, M., Dumont, S., Meistermann, D., Bordron, P., David, L., Redon, R.,
709 Bonnaud, S., and Bihouée, A. (2021). 3'RNA sequencing for robust and low-cost gene
710 expression profiling.
- 711 Cortal, A., Martignetti, L., Six, E., and Rausell, A. (2021). Gene signature extraction and cell
712 identity recognition at the single-cell level with Cell-ID. *Nat. Biotechnol.* 39, 1095–1102.
713 <https://doi.org/10.1038/s41587-021-00896-6>.
- 714 Costantini, D.L., Arruda, E.P., Agarwal, P., Kim, K.-H., Zhu, Y., Zhu, W., Lebel, M., Cheng, C.W.,
715 Park, C.Y., Pierce, S.A., et al. (2005). The Homeodomain Transcription Factor *Irx5* Establishes
716 the Mouse Cardiac Ventricular Repolarization Gradient. *Cell* 123, 347–358.
717 <https://doi.org/10.1016/j.cell.2005.08.004>.
- 718 DeLaughter, D.M., Bick, A.G., Wakimoto, H., McKean, D., Gorham, J.M., Kathiriya, I.S., Hinson,
719 J.T., Homsy, J., Gray, J., Pu, W., et al. (2016). Single-Cell Resolution of Temporal Gene
720 Expression during Heart Development. *Dev. Cell* 39, 480–490.
721 <https://doi.org/10.1016/j.devcel.2016.10.001>.
- 722 Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the
723 integration of genomic datasets with the R/Bioconductor package *biomaRt*. *Nat. Protoc.* 4,
724 1184–1191. <https://doi.org/10.1038/nprot.2009.97>.
- 725 Durocher, D., Charron, F., Warren, R., Schwartz, R.J., and Nemer, M. (1997). The cardiac
726 transcription factors *Nkx2-5* and *GATA-4* are mutual cofactors. *EMBO J.* 16, 5687–5696.
727 <https://doi.org/10.1093/emboj/16.18.5687>.
- 728 Es-Salah-Lamoureux, Z., Jouni, M., Malak, O.A., Belbachir, N., Al Sayed, Z.R., Gandon-Renard,
729 M., Lamirault, G., Gauthier, C., Baró, I., Charpentier, F., et al. (2016). HIV-Tat induces a
730 decrease in IKr and IKs via reduction in phosphatidylinositol-(4,5)-bisphosphate availability. *J.*
731 *Mol. Cell. Cardiol.* 99, 1–13. <https://doi.org/10.1016/j.yjmcc.2016.08.022>.
- 732 Gaborit, N., Sakuma, R., Wylie, J.N., Kim, K.-H., Zhang, S.-S., Hui, C.-C., and Bruneau, B.G.
733 (2012). Cooperative and antagonistic roles for *Irx3* and *Irx5* in cardiac morphogenesis and
734 postnatal physiology. *Dev. Camb. Engl.* 139, 4007–4019. <https://doi.org/10.1242/dev.081703>.
- 735 Girardeau, A., Atticus, D., Canac, R., Cimarosti, B., Caillaud, A., Chariou, C., Simonet, F., Cariou,
736 B., Charpentier, F., Gourraud, J.-B., et al. (2021). Generation of human induced pluripotent
737 stem cell lines from four unrelated healthy control donors carrying European genetic
738 background. *Stem Cell Res.* 59, 102647. <https://doi.org/10.1016/j.scr.2021.102647>.

- 739 Gonzalez-Teran, B., Pittman, M., Felix, F., Thomas, R., Richmond-Buccola, D., Hüttenhain, R.,
740 Choudhary, K., Moroni, E., Costa, M.W., Huang, Y., et al. (2022). Transcription factor protein
741 interactomes reveal genetic determinants in heart disease. *Cell* 185, 794-814.e30.
742 <https://doi.org/10.1016/j.cell.2022.01.021>.
- 743 Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations
744 in multidimensional genomic data. *Bioinforma. Oxf. Engl.* 32, 2847–2849.
745 <https://doi.org/10.1093/bioinformatics/btw313>.
- 746 He, W., Jia, Y., and Takimoto, K. (2009). Interaction between transcription factors Iroquois
747 proteins 4 and 5 controls cardiac potassium channel Kv4.2 gene transcription. *Cardiovasc. Res.*
748 81, 64–71. <https://doi.org/10.1093/cvr/cvn259>.
- 749 Hofbauer, P., Jahnel, S.M., Papai, N., Giesshammer, M., Deyett, A., Schmidt, C., Penc, M.,
750 Tavernini, K., Grdseloff, N., Meledeth, C., et al. (2021). Cardioids reveal self-organizing
751 principles of human cardiogenesis. *Cell* <https://doi.org/10.1016/j.cell.2021.04.034>.
- 752 Kathiriya, I.S., Rao, K.S., Iacono, G., Devine, W.P., Blair, A.P., Hota, S.K., Lai, M.H., Garay, B.I.,
753 Thomas, R., Gong, H.Z., et al. (2021). Modeling Human TBX5 Haploinsufficiency Predicts
754 Regulatory Networks for Congenital Heart Disease. *Dev. Cell* 56, 292-309.e9.
755 <https://doi.org/10.1016/j.devcel.2020.11.020>.
- 756 Kimura, Y., Aiba, T., Sasano, T., Furukawa, T., Kusano, K., and Shimizu, W. (2016). IRX3 variant
757 as a modifier of Brugada syndrome with frequent ventricular fibrillation. *Hear. Case Rep.* 2,
758 465–468. <https://doi.org/10.1016/j.hrcr.2016.06.002>.
- 759 Koizumi, A., Sasano, T., Kimura, W., Miyamoto, Y., Aiba, T., Ishikawa, T., Nogami, A., Fukamizu,
760 S., Sakurada, H., Takahashi, Y., et al. (2016). Genetic defects in a His-Purkinje system
761 transcription factor, IRX3, cause lethal cardiac arrhythmias. *Eur. Heart J.* 37, 1469–1475.
762 <https://doi.org/10.1093/eurheartj/ehv449>.
- 763 Krishnan, A., Samtani, R., Dhanantwari, P., Lee, E., Yamada, S., Shiota, K., Donofrio, M.T.,
764 Leatherbury, L., and Lo, C.W. (2014). A detailed comparison of mouse and human cardiac
765 development. *Pediatr. Res.* 76, 500–507. <https://doi.org/10.1038/pr.2014.128>.
- 766 Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J.*
767 *Stat. Softw.* 25, 1–18. .
- 768 Li, X., Martinez-Fernandez, A., Hartjes, K.A., Kocher, J.-P.A., Olson, T.M., Terzic, A., and Nelson,
769 T.J. (2014). Transcriptional atlas of cardiogenesis maps congenital heart disease interactome.
770 *Physiol. Genomics* 46, 482–495. <https://doi.org/10.1152/physiolgenomics.00015.2014>.
- 771 Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C.L., Lindberg, E.L., Kanda,
772 M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. *Nature* 588,
773 466–472. <https://doi.org/10.1038/s41586-020-2797-4>.
- 774 Luna-Zurita, L., Stirnimann, C.U., Glatt, S., Kaynak, B.L., Thomas, S., Baudin, F., Samee, M.A.H.,
775 He, D., Small, E.M., Mileikovsky, M., et al. (2016). Complex Interdependence Regulates

- 776 Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. *Cell* 164, 999–
777 1014. <https://doi.org/10.1016/j.cell.2016.01.004>.
- 778 McCulley, D.J., and Black, B.L. (2012). Transcription factor pathways and congenital heart
779 disease. *Curr. Top. Dev. Biol.* 100, 253–277. [https://doi.org/10.1016/B978-0-12-387786-
780 4.00008-7](https://doi.org/10.1016/B978-0-12-387786-780-4.00008-7).
- 781 McGinnis, C.S., Murrow, L.M., and Gartner, Z.J. (2019). DoubletFinder: Doublet Detection in
782 Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst.* 8, 329–337.e4.
783 <https://doi.org/10.1016/j.cels.2019.03.003>.
- 784 Meilhac, S.M. (2018). Cardiac growth I: Cardiomyocyte proliferation. *ESC Textb. Cardiovasc.*
785 *Dev.* 75. .
- 786 Meilhac, S.M., and Buckingham, M.E. (2018). The deployment of cell lineages that form the
787 mammalian heart. *Nat. Rev. Cardiol.* 15, 705–724. [https://doi.org/10.1038/s41569-018-0086-
788 9](https://doi.org/10.1038/s41569-018-0086-788-9).
- 789 Nemer, G., and Nemer, M. (2010). GATA4 in Heart Development and Disease. *Heart Dev.*
790 *Regen.* 599–616. <https://doi.org/10.1016/B978-0-12-381332-9.00027-X>.
- 791 Ng, A.H.M., Khoshakhlagh, P., Rojo Arias, J.E., Pasquini, G., Wang, K., Swiersy, A., Shipman,
792 S.L., Appleton, E., Kiaee, K., Kohman, R.E., et al. (2021). A comprehensive library of human
793 transcription factors for cell fate engineering. *Nat. Biotechnol.* 39, 510–519.
794 <https://doi.org/10.1038/s41587-020-0742-6>.
- 795 Olson, E.N., and Srivastava, D. (1996). Molecular pathways controlling heart development.
796 *Science* 272, 671–676. <https://doi.org/10.1126/science.272.5262.671>.
- 797 Piñero, J., Ramírez-Anguita, J.M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., and
798 Furlong, L.I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update.
799 *Nucleic Acids Res.* 48, D845–D855. <https://doi.org/10.1093/nar/gkz1021>.
- 800 Postma, A.V., Christoffels, V.M., and Bezzina, C.R. (2011). Developmental aspects of cardiac
801 arrhythmogenesis. *Cardiovasc. Res.* 91, 243–251. <https://doi.org/10.1093/cvr/cvr134>.
- 802 Roselli, C., Chaffin, M.D., Weng, L.-C., Aeschbacher, S., Ahlberg, G., Albert, C.M., Almgren, P.,
803 Alonso, A., Anderson, C.D., Aragam, K.G., et al. (2018). Multi-ethnic genome-wide association
804 study for atrial fibrillation. *Nat. Genet.* 50, 1225–1233. [https://doi.org/10.1038/s41588-018-
805 0133-9](https://doi.org/10.1038/s41588-018-805-0133-9).
- 806 Rowton, M., Guzzetta, A., Rydeen, A.B., and Moskowitz, I.P. (2021). Control of cardiomyocyte
807 differentiation timing by intercellular signaling pathways. *Semin. Cell Dev. Biol.* 118, 94–106.
808 <https://doi.org/10.1016/j.semcdb.2021.06.002>.
- 809 Sala, L., van Meer, B.J., Tertoolen, L.G.J., Bakkens, J., Bellin, M., Davis, R.P., Denning, C., Dieben,
810 M.A.E., Eschenhagen, T., Giacomelli, E., et al. (2018). MUSCLEMOTION: A Versatile Open
811 Software Tool to Quantify Cardiomyocyte and Cardiac Muscle Contraction In Vitro and In Vivo.
812 *Circ. Res.* 122, e5–e16. <https://doi.org/10.1161/CIRCRESAHA.117.312067>.

813 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski,
814 B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of
815 biomolecular interaction networks. *Genome Res.* 13, 2498–2504.
816 <https://doi.org/10.1101/gr.1239303>.

817 Si-Tayeb, K., Noto, F.K., Nagaoka, M., Li, J., Battle, M.A., Duris, C., North, P.E., Dalton, S., and
818 Duncan, S.A. (2010). Highly efficient generation of human hepatocyte-like cells from induced
819 pluripotent stem cells. *Hepatology*. Baltimore, Md 51, 297–305. <https://doi.org/10.1002/hep.23354>.

820 Specht, A.T., and Li, J. (2017). LEAP: constructing gene co-expression networks for single-cell
821 RNA-sequencing data using pseudotime ordering. *Bioinformatics*. Oxford, Engl. 33, 764–766.
822 <https://doi.org/10.1093/bioinformatics/btw729>.

823 Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., Hao, Y.,
824 Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data.
825 *Cell* 177, 1888-1902.e21. <https://doi.org/10.1016/j.cell.2019.05.031>.

826 Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M.,
827 Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association
828 networks with increased coverage, supporting functional discovery in genome-wide
829 experimental datasets. *Nucleic Acids Res.* 47, D607–D613.
830 <https://doi.org/10.1093/nar/gky1131>.

831 Tai, Y.C. (2022). timecourse: Statistical Analysis for Developmental Microarray Time Course
832 Data (Bioconductor version: Release (3.14)).

833 Tarradas, A., Pinsach-Abuin, M.L., Mackintosh, C., Llorà-Batlle, O., Pérez-Serra, A., Batlle, M.,
834 Pérez-Villa, F., Zimmer, T., Garcia-Bassets, I., Brugada, R., et al. (2017). Transcriptional
835 regulation of the sodium channel gene (SCN5A) by GATA4 in human heart. *J. Mol. Cell. Cardiol.*
836 102, 74–82. <https://doi.org/10.1016/j.yjmcc.2016.10.013>.

837 Vincent, S.D., and Buckingham, M.E. (2010). How to make a heart: the origin and regulation
838 of cardiac progenitor cells. *Curr. Top. Dev. Biol.* 90, 1–41. [https://doi.org/10.1016/S0070-2153\(10\)90001-X](https://doi.org/10.1016/S0070-2153(10)90001-X).

840 Waardenberg, A.J., Ramialison, M., Bouveret, R., and Harvey, R.P. (2014). Genetic Networks
841 Governing Heart Development. *Cold Spring Harb. Perspect. Med.* 4, a013839.
842 <https://doi.org/10.1101/cshperspect.a013839>.

843 Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie,
844 J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and Coordinated Epigenetic Regulation of
845 Developmental Transitions in the Cardiac Lineage. *Cell* 151, 206–220.
846 <https://doi.org/10.1016/j.cell.2012.07.035>.

847 Wang, Y., Yi, N., Hu, Y., Zhou, X., Jiang, H., Lin, Q., Chen, R., Liu, H., Gu, Y., Tong, C., et al. (2020).
848 Molecular Signatures and Networks of Cardiomyocyte Differentiation in Humans and Mice.
849 *Mol. Ther. Nucleic Acids* 21, 696–711. <https://doi.org/10.1016/j.omtn.2020.07.011>.

- 850 Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al.
 851 (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. N.*
 852 *Y. N* 2, 100141. <https://doi.org/10.1016/j.xinn.2021.100141>.
- 853 Yoshida, Y., and Yamanaka, S. (2017). Induced Pluripotent Stem Cells 10 Years Later: For
 854 Cardiac Applications. *Circ. Res.* 120, 1958–1968.
 855 <https://doi.org/10.1161/CIRCRESAHA.117.311080>.
- 856 Zhang, S.-S., Kim, K.-H., Rosen, A., Smyth, J.W., Sakuma, R., Delgado-Olguín, P., Davis, M., Chi,
 857 N.C., Puvionindran, V., Gaborit, N., et al. (2011). Iroquois homeobox gene 3 establishes fast
 858 conduction in the cardiac His–Purkinje network. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13576–
 859 13581. <https://doi.org/10.1073/pnas.1106911108>.

IV.2. Projet 2 : Fonctions cardiaques spécifiques et redondantes d'IRX3 et d'IRX5

IV.2.A. Contexte

Les six facteurs de transcription de la famille Iroquois homeobox sont tous exprimés dans le cœur et y ont des actions spécifiques, mais également communes. La forte homologie de séquence entre les différents membres est en accord avec cette redondance fonctionnelle (Kim et al., 2012). Des études des rôles individuels et conjoints de membres de cette famille ont été réalisées pour les facteurs de transcription *Irx3* et *Irx5*, chez la souris. Elles ont démontré, qu'ils agissent dans des compartiments tissulaires cardiaques communs, en coopération et en redondance fonctionnelle (Gaborit et al., 2012). En effet, la souris *Irx5-KO* est viable, fertile et ne présente qu'une susceptibilité accrue aux arythmies due à la déstabilisation du gradient ventriculaire transmural du potentiel d'action (Costantini et al., 2005). De la même manière, la souris *Irx3-KO* est viable, fertile et ne présente qu'un ralentissement de la conduction ventriculaire (Zhang et al., 2011). Cependant, la double délétion d'IRX3 et IRX5 est létale au stade embryonnaire (E14.5), et les souris présentent des défauts cardiaques structuraux majeurs, liés à la perte des compensations de l'un envers l'autre, retrouvées chez les souris simple KO. Chez l'humain, les rôles d'IRX3 et IRX5 seuls ou en couple coopératif et fonctionnellement redondant, n'ont que peu été étudiés.

Le but du second projet de cette thèse a été d'identifier sans a priori, les fonctions spécifiques et les fonctions redondantes d'IRX3 et d'IRX5 dans un contexte cardiaque humain.

Etat d'avancement : Ce projet a débuté par la création et la publication de nouvelles lignées de cellules iPS pour l'étude d'IRX5. Le travail a ensuite principalement porté sur l'identification de l'ensemble des fonctions cardiomyocytaires d'IRX5, avec un focus sur deux d'entre elles (la repolarisation cardiaque et la fibrose). Des approfondissements de certaines analyses et la réalisation d'un travail similaire sur des modèles simple KO pour IRX3 et double KO pour IRX3 et IRX5 seront nécessaires pour finaliser le projet ([voir partie IV.2.D Perspectives pour la poursuite du projet 2](#)).

Article 2

Generation of three human induced pluripotent stem cell lines with IRX5 knockout and knockin genetic editions using CRISPR-Cas9 system

Robin Canac, Amandine Caillaud, Bastien Cimarosti, Aurore Girardeau, Hanan Hamamy, Bruno Reversade, Carine Bonnard, Zeina R. Al Sayed, Laurent David, Jeremie Poschmann, Patricia Lemarchand, Guillaume Lamirault, Nathalie Gaborit

DOI : 10.1016/j.scr.2021.102627 ; PMID : 34929443

Publié dans *Stem Cell Research*, 2022

- Génération et validation d'une nouvelle lignée de cellules iPS KO pour IRX5 (IRX5^{KO})
- Génération et validation d'une nouvelle lignée de cellules iPS exprimant la protéine IRX5 couplée à un tag HA (IRX5-HA)
- Validation du contrôle isogénique associé à ces deux lignées génétiquement modifiées



Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Lab Resource: Genetically-Modified Multiple Cell Lines



Generation of three human induced pluripotent stem cell lines with IRX5 knockout and knockin genetic editions using CRISPR-Cas9 system

Robin Canac^a, Amandine Caillaud^a, Bastien Cimarosti^a, Aurore Girardeau^a, Hanan Hamamy^b, Bruno Reversade^{c,d,e,f,g}, Carine Bonnard^h, Zeina R. Al Sayed^a, Laurent David^{i,j}, Jeremie Poschmann^j, Patricia Lemarchand^a, Guillaume Lamirault^{a,1}, Nathalie Gaborit^{a,1,*}

^a Université de Nantes, CHU Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France^b Department of Genetic Medicine and Development, Geneva University, Geneva 1211, Switzerland^c Genome Institute of Singapore, A*STAR, Singapore, Singapore^d Medical Genetics Department, Koç University School of Medicine (KUSOM), Istanbul, Turkey^e Department of Paediatrics, National, University of Singapore, Singapore, Singapore^f Institute of Molecular and Cell Biology, A*STAR, Singapore, Singapore^g Reproductive Biology Laboratory, Amsterdam Zuidoost, the Netherlands^h Skin Research Institute of Singapore, A*STAR, Singapore, Singaporeⁱ Université de Nantes, CHU Nantes, INSERM, CNRS, SFR Santé, FED 4203, INSERM UMS 016, CNRS, UMS 3556, F-44000 Nantes, France^j Université de Nantes, CHU Nantes, INSERM, CRTI, UMR 1064, ITUN, F-44000 Nantes, France

A B S T R A C T

Studies on animal models have shown that *Irx5* is an important regulator of cardiac development and that it regulates ventricular electrical repolarization gradient in the adult heart. Mutations in *IRX5* have also been linked in humans to cardiac conduction defects. In order to fully characterize the role of *IRX5* during cardiac development and in cardiomyocyte function, we generated three genetically-modified human induced pluripotent stem cell lines: two knockout lines (heterozygous and homozygous) and a knockin HA-tagged line (homozygous).

1. Resource table

Unique stem cell lines identifier	1. ITXi002-A 2. ITXi002-A-1 3. ITXi002-A-2 4. ITXi002-A-3
Alternative name(s) of stem cell lines	1. IRX5-Wt (ITXi002-A) 2. IRX5-Het (ITXi002-A-1) 3. IRX5-KO (ITXi002-A-2) 4. IRX5-KI (ITXi002-A-3)
Institution	l'institut du thorax (INSERM UMR1087, CNRS UMR6291)
Contact information of the reported cell line distributor	Nathalie Gaborit, nathalie.gaborit@univ-nantes.fr
Type of cell lines	iPSC
Origin	Human
Additional origin info (applicable for human ESC or iPSC)	IRX5-Wt (ITXi002-A) • Age: 15 • Male
Cell Source	Skin Fibroblasts

(continued on next column)

(continued)

Unique stem cell lines identifier	1. ITXi002-A 2. ITXi002-A-1 3. ITXi002-A-2 4. ITXi002-A-3
Method of reprogramming	Sendai virus (<i>OCT4</i> , <i>KLF4</i> , <i>SOX2</i> and <i>c-MYC</i>)
Clonality	Clonal
Evidence of the reprogramming transgene loss (including genomic copy if applicable)	RT/q-PCR
Cell culture system used	StemMACS™ iPSC Brew XF Medium on Matrigel® hESC-Qualified Matrix
Type of Genetic Modification	CRISPR/Cas9
Associated disease	N/A
Gene/locus	IRX5/16q12.2
Method of modification/site-specific nuclease used	CRISPR/Cas9
Site-specific nuclease (SSN) delivery method	Electroporation of RNP

(continued on next page)

* Corresponding author at: l'institut du thorax Inserm UMR 1087, CNRS, UMR 6291, IRS-UN, 8 quai Moncousu, 44007 Nantes cedex 1, France.

E-mail address: nathalie.gaborit@univ-nantes.fr (N. Gaborit).¹ Contributed equally to this work.<https://doi.org/10.1016/j.scr.2021.102627>

Received 26 November 2021; Accepted 12 December 2021

Available online 13 December 2021

1873-5061/© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>.

(continued)

Unique stem cell lines identifier	1. ITXi002-A 2. ITXi002-A-1 3. ITXi002-A-2 4. ITXi002-A-3
All genetic material introduced into the cells	IRX5-KI: single-stranded oligo DNA nucleotide (ssODN)
Analysis of the nuclease-targeted allele status	Sequencing of the targeted alleles
Method of the off-target nuclease activity surveillance	N/A
Name of transgene	N/A
Eukaryotic selective agent resistance (including inducible/gene expressing cell-specific)	N/A
Inducible/constitutive system details	N/A
Date archived/stock date	IRX5-Wt: 04-05-2013
Cell line repository/bank	IRX5-Wt: https://hpscereg.eu/cell-line/ITXi002-A IRX5-Het: https://hpscereg.eu/cell-line/ITXi002-A-1 IRX5-KO: https://hpscereg.eu/cell-line/ITXi002-A-2 IRX5-KI: https://hpscereg.eu/cell-line/ITXi002-A-3
Ethical/GMO work approvals	A signed informed consent has been received. Authorization MESRI Authorization IE 2009-505(1) has been obtained from competent ethics review board and approval authority (French Health Ministry and French ethical committee) allowing the generation of these lines.
Addgene/public access repository recombinant DNA sources' disclaimers (if applicable)	N/A

2. Resource utility

The three generated hiPSC lines provide a (1) heterozygous or (2) homozygous deletion of *IRX5* gene and (3) an HA-tagged *IRX5* protein which are a useful resource to investigate *IRX5* transcription factor function in any human cell type, including mature and differentiating cardiomyocytes.

3. Resource details

IRX5 transcription factor (TF) of the Iroquois homeobox family has been described to participate to murine cardiac development and cardiomyocyte functions, especially the regulation of cardiac ventricular electrical repolarization gradient (Costantini et al., 2005; Gaborit et al., 2012; Kim et al., 2012). Several mutations in *IRX5* have been found in patients with Hamamy syndrome presenting a severe clinical phenotype, with cardiac structural and electrical defects (Bonnard et al., 2012; Al Sayed et al., 2021).

We previously generated a control human induced pluripotent stem cell (hiPSC) line, IRX5-Wt, from skin fibroblasts collected from a healthy control individual (Al Sayed et al., 2021), that we used to perform various genetic modifications of *IRX5* gene. Using CRISPR-Cas9 technology we first generated two knockout hiPSC lines: heterozygous (IRX5-Het) and homozygous (IRX5-KO). For that, we electroporated the IRX5-Wt hiPSC line with ribonucleoproteins (RNP) that combined the Cas9 protein, a fluorescent transactivator-ATTO488, the gRNA1 and gRNA2 (targeting respectively *IRX5* first and third exons), in order to delete the entire gene. Dissociated fluorescent hiPSC cells were FACS-

sorted, replated by wells to allow clonal expansion and genotyped to validate the *IRX5* deletions (Fig. 1A). *IRX5* being more expressed in cardiomyocytes than in hiPSCs, using western-blot on cardiomyocytes derived from IRX5-Wt, IRX5-Het and IRX5-KO hiPSC lines, we confirmed the reduction or complete absence of *IRX5* protein expression (Fig. 1B). Second, to generate the hiPSC line expressing HA-tagged *IRX5* protein (IRX5-KI) IRX5-Wt cells were transfected with an RNP combining Cas9 protein, a fluorescent transactivator-ATTO488 and the gRNA2, together with a fluorescent ssODN-Alexa660 that carried the HA tag sequence. The double fluorescent hiPSC clones sorted by FACS and genotyped by PCR, were sequenced to confirm HA tag insertion (Fig. 1C). The four isogenic hiPSC lines were cultured on Matrigel®-coated plates with StemMACS™ iPS-Brew XF medium. They were tested negative for mycoplasma contamination by luminescence analysis (Supplementary Fig. 1E). They all showed typical colony morphology assessed by brightfield microscopy (Fig. 1D-Left panels). They also expressed pluripotency markers OCT3/4 and TRA-1-60 as evaluated by immunofluorescence experiments (Fig. 1D-right panels). Pluripotency status of all four cell lines was also quantitatively assessed by RT-qPCR on *NANOG* and *POU5F1* markers (Fig. 1E). Regarding specifically the control line, IRX5-Wt, the absence of Sendai virus persistence following reprogramming process was proved (Fig. 1F). We then validated the ability of the three genetically modified cell lines and their isogenic control to differentiate into derivatives of the three germ layers by RT-qPCR analysis after 20 days of spontaneously differentiating embryoid bodies (Fig. 1G). Furthermore, no major chromosomal aberrations were found between the IRX5-Wt isogenic control and (1) its corresponding parental fibroblasts (Supplemental Figure 2A) and (2) its CRISPR-Cas9 resulting cell lines, IRX5-Het, IRX5-KO and IRX5-KI (Fig. 1H-J). Finally, evidence of genetic identity between the control parental fibroblasts and the four hiPSC lines was shown by single nucleotide polymorphism (SNP) analysis (Supplementary Fig. 2B). Detailed informations for all lines are provided in Table 1 and Table 2.

4. Materials and methods

4.1. hiPSC generation and culture

Following Sendai virus reprogramming and CRISPR-based editing, hiPSC isogenic lines were maintained at 37 °C, 5% CO₂, 21% O₂ in StemMACS™ iPS Brew XF Medium (Miltenyi Biotec) on Matrigel® hESC-Qualified Matrix (0.05 mg/mL, Corning). At 75% confluency, cells were passaged as clumps using Gentle Cell Dissociation Reagent (STEMCELL Technologies).

4.2. CRISPR-Cas9 genome editing

Two gRNAs were used to delete *IRX5* gene in IRX5-Wt hiPSC line to generate IRX5-Het and IRX5-KO lines. Regarding IRX5-KI, one gRNA and a fluorescent HA-tag ssODN-Alexa660 were used to introduce HA-tag before *IRX5* STOP codon. Cas9/gRNA/transactivator/±ssODN were delivered by electroporation with P3 Primary Cell 4D-Nucleofector™ X kit, CA137 pulse program (Lonza). Cells were FACS-sorted 48 h post-electroporation (BD FACSMelody™ cell sorter). Single or double positive cells were replated as single-cells and maintained in iPS Brew medium until they reached 80% confluency, prior duplication for genotyping and banking.

4.3. Immunofluorescence analysis

Cells were fixed with 4% paraformaldehyde for 15 min at room

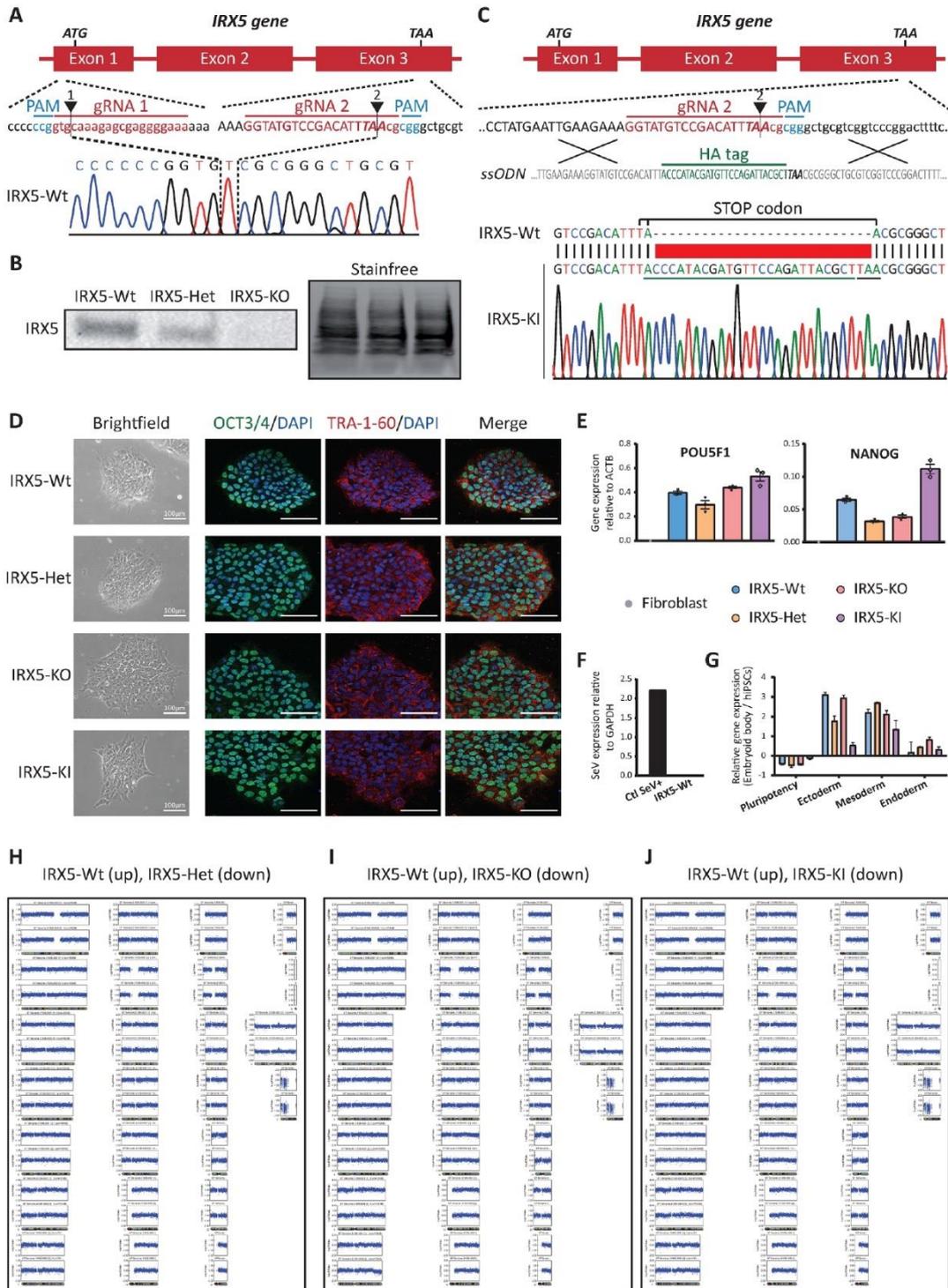


Fig. 1.

Table 1
Characterization and validation.

Classification	Test	Result	Data
Morphology	Photography	Typical morphological aspect of hiPSCs and colonies	Fig. 1D, left panels
Pluripotency status evidence for the described cell line	Qualitative analysis (immunofluorescence)	Assess staining of pluripotency markers: OCT3/4 and TRA-1-60	Fig. 1D, right panels
	Quantitative analysis (RT-qPCR)	Assess transcript expression of pluripotency markers: POU5F1 and NANOG	Fig. 1E
Karyotype	Molecular karyotyping (SNP array; >300000 markers)	IRX5-Wt: 46,XY IRX5-Het: 46,XY IRX5-KO: 46,XY IRX5-KI: 46,XY	Fig. 1H-J and Supplementary Fig. 2A
Verification of the absence of random plasmid integration events	PCR/Southern	N/A	N/A
Parental and modified cell line genetic identity evidence	CNV analysis	SNP array to compare IRX5-Wt to parental fibroblast and CRISPR-edited lines to IRX5-Wt line. Matching genotypes confirmed the origins of iPSC lines	Supplementary Fig. 2B and archived with the journal
Mutation analysis	PCR across the edited site and sequencing	169 tested sites Complete IRX5 deletion validated for IRX5-Het and IRX5-KO cell lines. Validated insertion of the HA-tag at the end of <i>IRX5</i> 's ORF for the IRX5-KI cell line.	Archived with the journal Fig. 1A, C and Supplementary Fig. 1A-D
Off-target nuclease analysis-	Western blot	No IRX5 protein expression in IRX5-KO cell line	Fig. 1B
	PCR across top 5/10 predicted top likely off-target sites, whole genome/exome sequencing	N/A	N/A
Specific pathogen-free status	Mycoplasma	Mycoplasma testing by luminescence: no contamination for all cell lines	Supplementary Fig. 1E
Multilineage differentiation potential	Embryoid body formation and differentiation followed by RT-qPCR analysis	Expression of genes in embryoid bodies: Ectoderm (<i>PAX6</i>); Mesoderm (<i>HAND1</i>); Endoderm (<i>SOX17</i>)	Fig. 1G
Donor screening (OPTIONAL)	HIV 1 + 2 Hepatitis B, Hepatitis C	N/A	N/A
Genotype - additional	Blood group genotyping	N/A	N/A
histocompatibility info (OPTIONAL)	HLA tissue typing	N/A	N/A

temperature (RT) in Matrigel®-coated μ -Slide 8 Well (IBIDI) prior permeabilisation with 0.5% PBS-Triton (15 min) and blocking with 3% PBS-BSA (RT, 30 min). Cells were then incubated with primary antibodies (4 °C, overnight). Finally, cells were washed and incubated with secondary antibodies and DAPI (RT, 1 h) and conserved in 0.5% paraformaldehyde (4 °C). Images were acquired by confocal microscopy (MicroPICell core facility).

4.4. Embryoid bodies (EB) generation

HiPSC colonies were detached with Dispase (STEMCELL Technologies) and replated in Costar® 6 Well Plate Low Attachment (Corning). Medium was changed the next day and unsupervised EB differentiation was started a day after, by incubating cells in EB medium: 78% (v/v) KnockOut™ DMEM/F-12 (Gibco) + 1% (v/v) GlutaMAX™ Supplement (Gibco) + 20% (v/v) KnockOut™ Serum Replacement (Gibco) + 1X MEM Non-Essential Amino Acids Solution + 0.1 mM 2-Mercaptoethanol (Gibco). The EB medium was changed daily during 21 days before RT-qPCR analysis.

4.5. SNP and CNV analysis

To confirm isogenic hiPSC line genome integrity we compared their single nucleotide polymorphisms (SNP). Genomic DNA was extracted from somatic and hiPSC samples using the QIAGEN QiaAmp kit. 200 ng of gDNA were outsourced to Integragen Company (Evry, France) for karyotype analysis using HumanCore-24-v1 SNP arrays. Analysis was performed with GenomeStudio software. Chromosome abnormalities

were visually determined on logR ratios values, comparing parental cells with hiPSC-derived samples. SNP data were used to compute CNV and validate cell line parentality.

4.6. RT-qPCR

Total RNAs were extracted using NucleoSpin RNA kit (MACHEREY-NAGEL) and reverse transcribed using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems™). PCR amplification was performed using (1) FAM-labeled TaqMan probes (Applied Biosystems) for pluripotency and differentiation potential, and (2) SYBR-green technology (Applied Biosystems) for SeV loss. Data were normalized to ACTB or GAPDH.

4.7. Western blot

60 μ g proteins were denatured using NuPAGE Sample Reducing Agent and NuPAGE LDS Sample Buffer and ran in 4–15% precast polyacrylamide gel (Bio-Rad). Then, proteins were transferred on nitrocellulose membrane prior blocking (5% milk) and antibodies incubations. Revelation was performed using chemiluminescence camera (Bio-Rad), and the Stainfree technology was used as loading control.

4.8. Mycoplasmas test

Mycoplasma contamination was assessed with MycoAlert™ Mycoplasma Detection Kit (Lonza).

Table 2
Reagents details.

Antibodies and stains used for immunocytochemistry/flow-cytometry			
	Antibody	Dilution	Company Cat # and RRID
Pluripotency Markers	Rat anti-OCT3/4	1:200	Thermo Fisher Scientific Cat# 14-5841-80, RRID:AB_914299
	Mouse anti-TRA-1-60	1:200	Thermo Fisher Scientific Cat# 14-8863-80, RRID:AB_891,612
Nuclear stain	DAPI	1 µg/mL	Merck Cat# D9564
Secondary antibodies	Goat Anti-Mouse IgG (H + L) Antibody, Alexa Fluor 568 Conjugated	1:1000	Molecular Probes Cat# A-11004, RRID: AB_2534072
	Goat Anti-Rat IgG (H + L) Antibody, Alexa Fluor 488 Conjugated	1:1000	Molecular Probes Cat# A-11006, RRID: AB_141373
	Goat anti-mouse IgG-HRP antibody	1:10 000	Santa Cruz Biotechnology Cat# sc-2055, RRID: AB_631738
Mutation analysis - WB	Mouse anti-IRX5	1:250	Santa Cruz Biotechnology Cat# sc-81102, RRID: AB_1124818
Site-specific nuclease			
Nuclease information	Alt-R® S.p. HiFi Cas9 Nuclease V3		Integrated DNA Technologies (IDT) Cat# 1,081,061
Delivery method	Nucleofection		P3 Primary Cell 4D-Nucleofector™ X, Lonza Cat# V4XP-3024
Selection/enrichment strategy	FACS sorting using transactivator-associated fluorescence with (if present) ssODN-associated fluorescence		<u>Transactivator</u> : Alt-R® CRISPR-Cas9 tracrRNA ATTO™ 488, <u>IDT ssODN</u> : Single-stranded DNA oligo Alexa Fluor® 660, IDT
Primers and Oligonucleotides used in this study			
	Target	Forward/Reverse primer (5'-3')	
Sendai Virus Detection (primer)	SeV	Forward: 5'-GGA TCA CTA GGT GAT ATC GAG C-3' Reverse: 5'-ACC AGA CAA GAG TTT AAG AGA TAT GTA TC-3'	
House-Keeping Genes (primer)	<i>GAPDH</i>	Forward: 5'-AAT CCC ATC ACC ATC TTC CA-3' Reverse: 5'-TGG ACT CCA CGA CGT ACT CA-3'	
Pluripotency Markers (TaqMan probe)	<i>NANOG</i>	Thermo Fisher Scientific Cat# Hs02387400	
Pluripotency Markers (TaqMan probe)	<i>POU5F1</i>	Thermo Fisher Scientific Cat# Hs04260367	
Ectoderm Marker (TaqMan probe)	<i>PAX6</i>	Thermo Fisher Scientific Cat# Hs01088114	
Endoderm Marker (TaqMan probe)	<i>SOX17</i>	Thermo Fisher Scientific Cat# Hs00751752	
Mesoderm Marker (TaqMan probe)	<i>HAND1</i>	Thermo Fisher Scientific Cat# Hs02330376	
House-Keeping Genes (TaqMan probe)	<i>ACTB</i>	Thermo Fisher Scientific Cat# Hs01060665	
IRX5 gDNA PCR	<i>IRX5</i>	For1: 5'-ACAAGGCGTCAGAATCCTCA-3' For2: 5'-GGGCCCTCGGTTATCCATTG-3' ForA: 5'-TCTATCCCGGCTACACGAAC-3' Rev1: 5'-GTCCGTCCCGACTTTTCTA-3' RevB: 5'-TGCTTAACCAACCGAAAAGAAA-3' RevB: 5'-TGCTTAACCAACCGAAAAGAAA-3' ForA: 5'-TCTATCCCGGCTACACGAAC-3'	
Sequencing	<i>IRX5</i> deletion HA tag sequence insertion	5'-GTGCAAAGAGCGAGGGGAAA-3' 5'-GGTATGTCGACATTTAACG-3' Human, GRCh38.p13, Chr16: 54,930,865-54,934,485 5'-TATGAATTGAAGAAAGGTATGTCGACATTTACCCATACGA TGTTCCAGATTACGCTTAACGGGGTGGTGGTCCGTCGGGACTTTTCTAATTTATAAAA-3'	
crRNA IRX5-Exon1 (gRNA 1)	<i>IRX5</i>		
crRNA IRX5-Exon3 (gRNA 2)	<i>IRX5</i>		
Genomic target sequence	<i>IRX5</i> gene		
ssODN HA tag	<i>IRX5</i>		

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the iPS core facility of Nantes University, for their technical support. This work was funded by grants from The French National Research Agency (HEART iPS ANR-15-CE14-0019-01) and La Fédération Française de Cardiologie. NG was laureate of fellowships from Fondation Lefoulon-Delalande and Marie Curie International Incoming Fellowship FP7-PEOPLE-2012-IF [PIIF-GA-2012-331436]. R.C. is laureate of fellowship from the Fonds Marion

Elizabeth Brancher.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scr.2021.102627>.

References

- Al Sayed, Z.R., Canac, R., Cimaresti, B., Bonnard, C., Gourraud, J.B., Hamamy, H., Kayserili, H., Girardeau, A., Jouni, M., Jacob, N., Gaignerie, A., Chariou, C., David, L., Forest, V., Marionneau, C., Charpentier, F., Loussouarn, G., Lamirault, G., Reversade, B., Zibara, K., Lemarchand, P., Gaborit, N. 2021. Human model of IRX5 mutations reveals key role for this transcription factor in ventricular conduction. *Cardiovasc Res.* 117, 2092-2107.
- Bonnard, C., Strobl, A.C., Shboul, M., Lee, H., Merriman, B., Nelson, S.F., Ababneh, O.H., Uz, E., Güran, T., Kayserili, H., Hamamy, H., Reversade, B., 2012. Mutations in IRX5

R. Canac et al.

Stem Cell Research 58 (2022) 102627

impair craniofacial development and germ cell migration via SDF1. *Nat Genet.* 44 (6), 709–713.

Costantini, D.L., Arruda, E.P., Agarwal, P., Kim, K.-H., Zhu, Y., Zhu, W., Lebel, M., Cheng, C.W., Park, C.Y., Pierce, S.A., Guerchicoff, A., Pollevick, G.D., Chan, T.Y., Kabir, M.G., Cheng, S.H., Husain, M., Antzelevitch, C., Srivastava, D., Gross, G.J., Hui, C.-C., Backx, P.H., Bruneau, B.G., 2005. The homeodomain transcription factor *Irx5* establishes the mouse cardiac ventricular repolarization gradient. *Cell* 123 (2), 347–358.

Gaborit, N., Sakuma, R., Wylie, J.N., Kim, K.H., Zhang, S.S., Hui, C.C., Bruneau, B.G., 2012. Cooperative and antagonistic roles for *Irx3* and *Irx5* in cardiac morphogenesis and postnatal physiology. *Development* 139, 4007–4019.

Kim, K.-H., Rosen, A., Bruneau, B.G., Hui, C.-C., Backx, P.H., 2012. Iroquois homeodomain transcription factors in heart development and function. *Circ. Res.* 110 (11), 1513–1524.

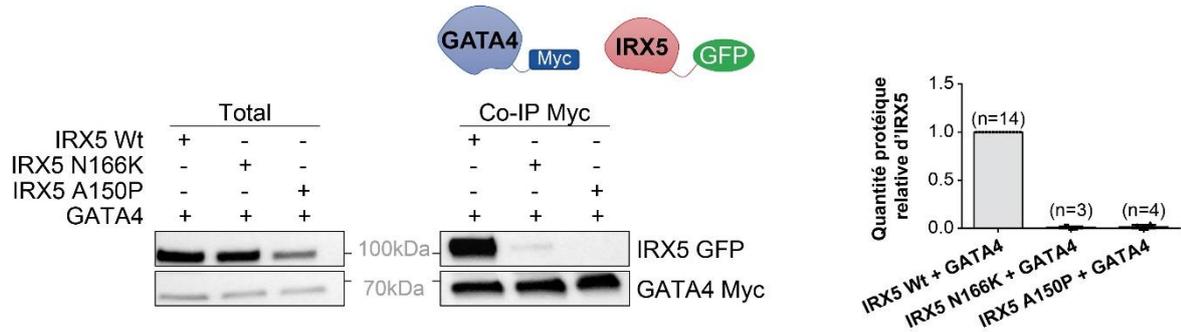
IV.2.C. Résultats complémentaires

Les résultats présentés ci-après apportent des informations complémentaires sur la structure protéique d'IRX5, ainsi que sur ses mécanismes d'action sur l'ADN et les fonctions biologiques qu'il régule dans la différenciation cardiomyocytaire.

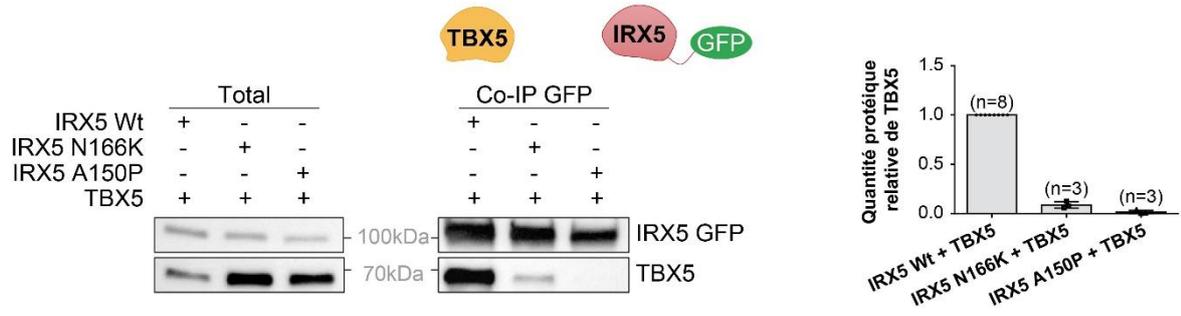
L'homéodomaine d'IRX5 est important pour les interactions avec ses partenaires

Dans le but de mieux comprendre le rôle des différents domaines d'IRX5 dans ses interactions protéiques avec ses partenaires (identifiés dans l'article 1, [IV.1.A - Article 1](#)), des formes d'IRX5 porteuses des mutations Hamamy (A150P et N166K) ont été utilisées dans des expériences de transfection suivie de co-immunoprécipitation (Co-IP) sur des cellules humaines HEK293, pour étudier les interactions entre IRX5 et GATA4, TBX5 et NKX2-5 (Figure 1A-C). Les deux mutations provoquent, de manière similaire, une diminution ou une perte totale d'interaction entre IRX5 et ses partenaires précédemment identifiés. Les mutations étant localisées dans deux hélices α différentes de l'homéodomaine, connu pour son importance dans les interactions à l'ADN, ces résultats suggèrent un rôle jusque-là non décrit de l'homéodomaine dans les interactions protéiques d'IRX5. Afin d'étudier plus précisément l'impact des différentes hélices de l'homéodomaine, ainsi que les régions centrale et C-terminale d'IRX5, différentes formes tronquées de la protéine ont été générées (Figure 1D). Ces protéines IRX5 tronquées ont ensuite été utilisées dans des expériences de Co-IP après transfection de HEK293, pour étudier l'effet des délétions sur les interactions physiques entre IRX5 et GATA4, TBX5 ou NKX2-5. Bien que les délétions des régions centrale (Mid) et C-terminale (C-ter) n'induisent qu'une diminution légère de l'interaction entre IRX5 et ses partenaires, les délétions individuelles de chacune des trois hélices ($\Delta H1$, $\Delta H2$, $\Delta H3$) de l'homéodomaine d'IRX5 provoquent une perte totale d'interaction (Figure 1E-G), démontrant le rôle majeur de l'homéodomaine d'IRX5 dans ses interactions protéiques.

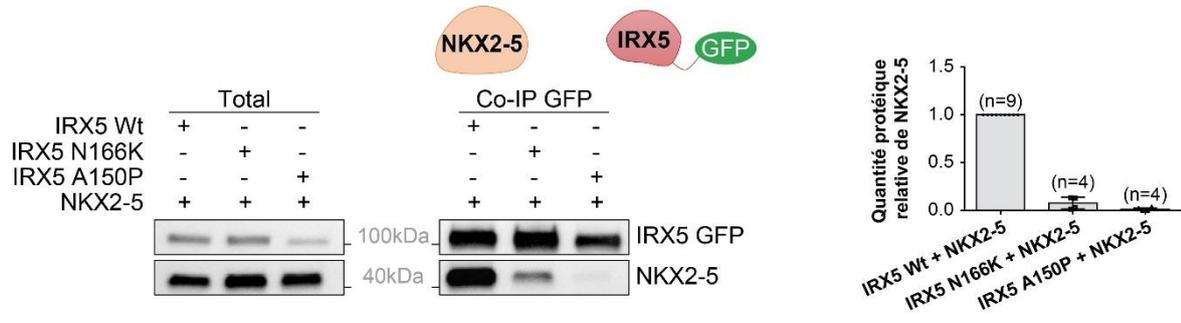
A



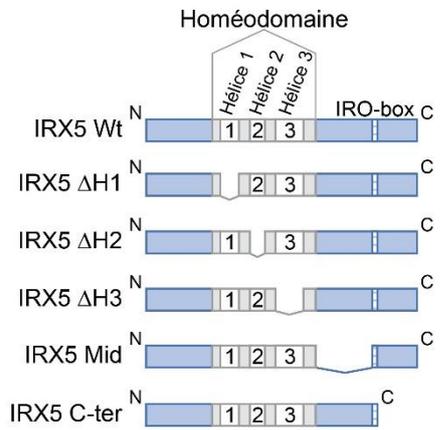
B



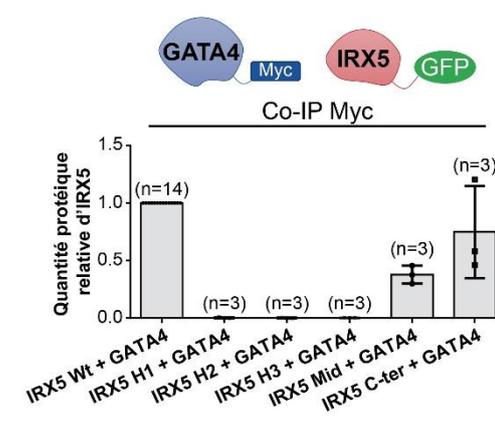
C



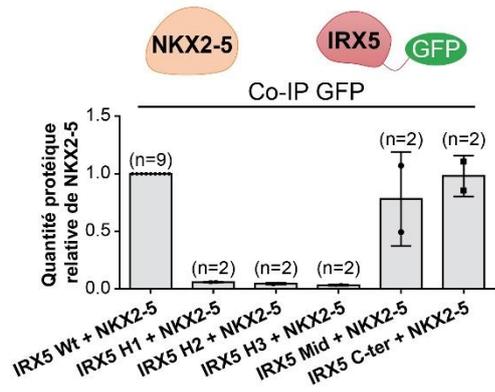
D



E



F



G

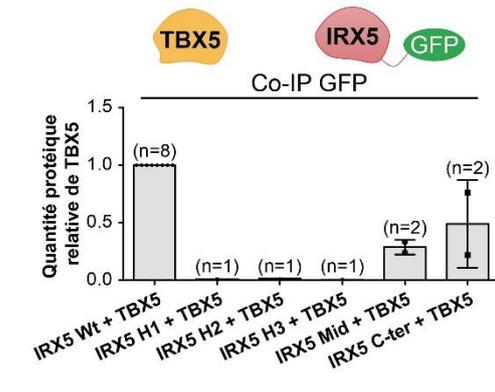


Figure 1 : L'homéodomaine d'IRX5 a un rôle central dans les interactions protéiques

A-C. Co-immunoprécipitations après surexpression des protéines d'intérêt dans des cellules HEK293. Un exemple représentatif de membrane de western-blot est présenté (panels de gauche), ainsi que la quantification associée (panels de droite). **D.** Schéma des constructions protéiques d'IRX5 utilisées pour les co-immunoprécipitations présentées par la suite. **E-G.** Co-immunoprécipitations des constructions protéiques d'IRX5, après surexpression dans des cellules HEK293. Seules les quantifications sont présentées.

IRX5 se fixe à distance de gènes cardiaques dans les cardiomyocytes dérivés d'iPS

Afin d'identifier les régions génomiques occupées par IRX5 dans les cardiomyocytes, des expériences de ChIP-Seq ont été réalisées sur 6 réplicats biologiques de cardiomyocytes dérivés d'iPS en fin de différenciation (J30). IRX5 a été localisé dans 2 232 régions du génome (correspondant aux pics IRX5) et associé à 2 253 gènes par le logiciel *GREAT* (voir partie III.2.B.c Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 1). IRX5 semble occuper des régions à distance de ces gènes, majoritairement à plus de 50kb en amont ou en aval de leurs *Transcription Start Site* ou TSS (Figure 2A). Une analyse d'annotation fonctionnelle a ensuite permis d'associer les pics IRX5 à des fonctions biologiques spécifiques : celles de développement et de différenciation cardiaque. Ces résultats suggèrent donc qu'IRX5 exerce des rôles spécifiques dans les processus développementaux cardiaques humains (Figure 2B).

IRX5 semble orienter l'action du complexe GATA4/TBX5/NKX2-5 dans l'activité électrique cardiomyocytaire

Les résultats présentés précédemment (IV.1.A - Article 1) montrant une interaction physique et fonctionnelle entre IRX5 et le complexe GATA4/TBX5/NKX2-5, nous avons ensuite collecté et intégré à nos données, des données publiques de ChIP-Seq qui avaient été obtenues pour ces 3 facteurs de transcription dans un contexte cellulaire comparable au notre : des cellules souches humaines différenciées en cardiomyocytes (Anderson et al., 2018; Ang et al., 2016). La liste des 2 253 gènes occupés par IRX5 a alors été comparée aux listes de gènes occupés respectivement par GATA4, TBX5 et NKX2-5 (Figure 2C). Les membres du complexe GATA4/TBX5/NKX2-5 occupent des régions génomiques à proximité de 2 990 gènes qui sont, en partie (848 gènes), également occupés par IRX5. L'analyse d'annotation fonctionnelle globale de ces 2 990 gènes révèle un enrichissement de différentes fonctions biologiques, dont la conduction électrique et la contraction cardiaque (Figure 2D). Afin d'identifier, parmi ces fonctions retrouvées, celles qui sont spécifiquement enrichies pour les 848 gènes également co-occupés par IRX5, une nouvelle analyse d'annotation fonctionnelle a été réalisée (Figure 2E). De manière intéressante, la fonction de transport des canaux ioniques

a ainsi été retrouvée enrichie pour les 848 gènes occupés par les quatre facteurs de transcription, suggérant qu'IRX5 participe à la spécialisation du complexe GATA4/TBX5/NKX2-5 dans la régulation de l'activité électrique cardiomyocytaire.

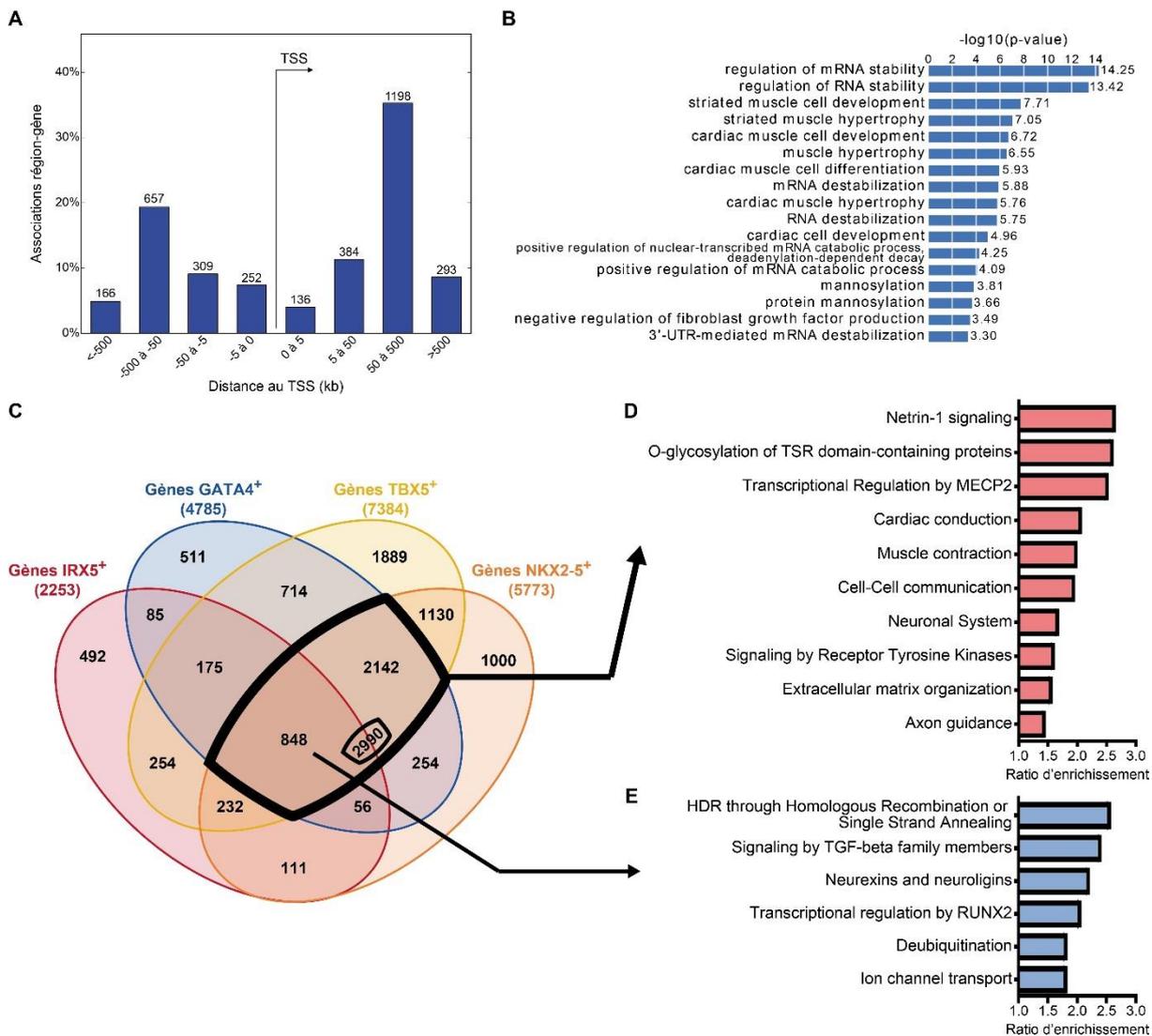


Figure 2 : Etude des régions génomiques occupées par IRX5 dans les cardiomyocytes dérivés d'iPS à J30

A. Distribution des pics ChIP-Seq d'IRX5 par rapport aux *Transcription Start Site* (TSS) des gènes les plus proches associés par le logiciel *GREAT*. **B.** Annotation fonctionnelle des gènes associés aux pics ChIP-Seq d'IRX5 par le logiciel *GREAT*. **C.** Comparaison des gènes occupés (ayant au moins un pic ChIP-Seq associé par *GREAT*) par IRX5, GATA4, TBX5 et NKX2-5 dans des cardiomyocytes en fin de différenciation cardiaque de cellules souches. Données ChIP-Seq GATA4, TBX5 et NKX2-5 : Ang et al., 2016 ; Anderson et al., 2018. **D.** Annotation fonctionnelle des 2 990 gènes occupés par GATA4, TBX5 et NKX2-5, avec ou sans IRX5. Analyse réalisée avec le logiciel *WebGestalt*. Background pour l'analyse : Génome. **E.** Annotation fonctionnelle des 848 gènes occupés par les quatre facteurs de transcription. Analyse réalisée avec le logiciel *WebGestalt*. Background pour l'analyse : liste des 2 990 gènes GATA4⁺, TBX5⁺ et NKX2-5⁺ ± IRX5⁺.

Impact de l'absence d'IRX5 sur les profils d'expression de gènes au cours de la différenciation cardiaque

Pour étudier plus spécifiquement les voies moléculaires par lesquelles IRX5 régule des mécanismes biologiques non seulement dans les cardiomyocytes obtenus en fin de différenciation mais aussi au cours de la différenciation cardiaque, nous avons mis à profit la lignée iPS IRX5^{KO} générée précédemment. Des échantillons d'ARN ont ainsi été collectés (1) au stade iPS (Jour -1, J-1), (2) au stade cardiomyocyte (Jour 30 ou J30) ainsi (3) qu'à deux stades cardiomyocytaires intermédiaires de la différenciation cardiaque (Jour 12 ou J12 et Jour 20 ou J20), et ce, pour quatre différenciations d'iPS humaines contrôles et IRX5^{KO} respectivement (Figure 3A). Le stade J12 a été sélectionné sur la base des données d'expression d'IRX5 en cinétique sur lignées contrôles (IV.1.A - Article 1), car il correspond à une phase du processus de différenciation où le niveau d'expression d'IRX5 est le plus élevé (Figure 3B). Le stade J20 a lui été sélectionné afin d'inclure un stade intermédiaire entre J12 et J30 qui assure la possibilité de pouvoir faire des analyses en cinétique. L'absence effective d'expression d'IRX5 à ces quatre stades étudiés dans les cellules IRX5^{KO} par rapport aux contrôles a, comme attendu, été observée dans les données de 3'SRP et confirmée par qPCR (Figure 3C-D). L'analyse en composante principale (PCA) de l'ensemble de ces données de transcriptomique en cinétique, a permis de séparer les échantillons selon le jour de prélèvement (par la Composante 1), ainsi que par la condition contrôle ou IRX5^{KO} (par la Composante 3), suggérant que l'absence d'IRX5 est associée à une signature transcriptomique particulière à chaque jour de différenciation étudié (Figure 3E). La ségrégation entre les échantillons contrôle ou IRX5^{KO} à J-1 sur la PCA suggère que l'absence d'IRX5 au stade iPS provoque déjà des variations d'expression de gènes spécifiques, alors que son expression évaluée par 3'SRP dans les échantillons contrôles, semblait être nulle à J-1. Une analyse par qPCR a donc été réalisée pour investiguer si l'expression d'IRX5 est effectivement nulle au stade iPS, ou bien trop faible pour être détectée par 3'SRP (Figure 3F). Une gamme de quantité d'ADNc a été utilisée sur trois échantillons J-1 contrôles, car en considérant l'aspect exponentiel de la PCR, si IRX5 était exprimé, le fait de doubler la quantité d'ADNc devait provoquer la perte d'environ un Ct de qPCR. Dans le cas inverse, si IRX5 n'était pas exprimé, le doublement de la quantité d'ADNc ne devait pas avoir d'impact sur la valeur de Ct obtenue. Une relation linéaire inverse entre le nombre de Ct et la quantité d'ADNc a bien été obtenue, démontrant donc qu'IRX5 est exprimé, bien que faiblement, dès le stade iPS, et que la technique de 3'SRP ne permet pas de détecter

ce niveau d'expression. Il est intéressant de mentionner qu'il s'agit du premier constat d'expression d'IRX5 au stade de pluripotence. Il sera intéressant à l'avenir d'investiguer les voies moléculaires régulées par IRX5 à ce stade.

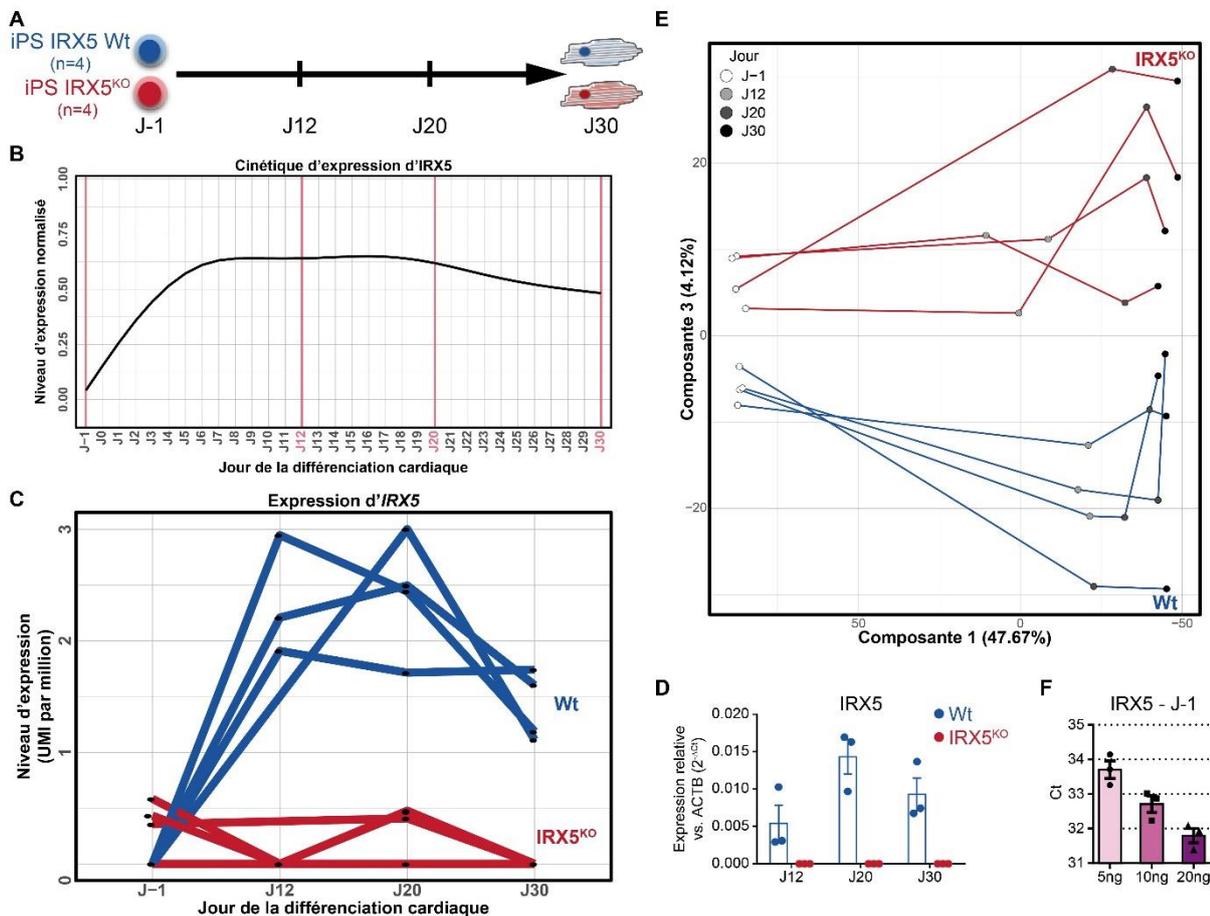


Figure 3 : Analyse de l'expression d'IRX5 à J-1, J12, J20 et J30 de la différenciation cardiaque des iPS

A. Représentation schématique des cinétiques d'expression étudiées. **B.** Niveau d'expression d'IRX5 à chaque jour de la différenciation cardiaque des iPS contrôles (données de 3'SRP issues de l'Article 1), mettant en évidence le plateau d'expression maximal d'IRX5 à J12 et sa diminution d'expression à J20, avant la fin de la différenciation à J30. **C.** Variations d'expression d'IRX5 aux quatre jours étudiés de la différenciation cardiaque des iPS Wt et IRX5^{KO}, obtenue par 3'SRP. Unité en *Unique Molecular Identifier* (UMI) par million. **D.** Expression relative d'IRX5 à J12, J20 et J30 des différenciations Wt et IRX5^{KO}, quantifiée par qPCR. Contrôle positif : *ACTB*. **E.** Analyse en composante principale des données de transcriptomique 3'SRP Wt vs. IRX5^{KO}. Ne sont présentées que les composantes 1 et 3, séparant le mieux les échantillons Wt et IRX5^{KO}. Un échantillon contrôle et un échantillon IRX5^{KO} à J12 ont été identifiés et retirés de la suite des analyses lors du contrôle qualité de séquençage, à cause d'un nombre de *reads* trop faible. **F.** Nombre de Ct obtenus par qPCR ciblée sur *IRX5*, en fonction de la quantité d'ADNc iPS Wt (J-1) utilisée.

Une analyse d'expression différentielle de gènes a été réalisée à chaque jour étudié, pour identifier les fonctions et mécanismes biologiques régulés par IRX5 (Figure 4). Les analyses PCA réalisées séparément sur les données transcriptomiques à chacun des jours étudiés, montrent que les échantillons Wt et IRX5^{KO} se clusterisent séparément (Figure 4A) comme observé sur la PCA globale (Figure 3E). Des gènes dont l'expression varie

significativement entre les cellules contrôles et IRX5^{KO} ont été identifiés à chaque jour étudié de la différenciation cardiaque des iPS (Figure 4B). De manière intéressante, plus de gènes différentiels ont été retrouvés à J12 qu'à J20 ou J30, suggérant qu'IRX5 a un impact plus important dans le développement cardiomyocytaire plutôt qu'en fin de différenciation. L'analyse d'annotation fonctionnelle des gènes différentiels à chaque stade a permis de déterminer les fonctions et mécanismes biologiques impactés par l'absence d'IRX5 (Figure 4C). Certaines fonctions biologiques sont retrouvées associées aux gènes différentiels de plusieurs jours, comme les fonctions associées au développement et la contraction cardiaque. D'autres fonctions modulées par IRX5 sont spécifiques de certains jours étudiés, comme la morphogenèse précoce à J-1, la migration cellulaire à J12, ou le métabolisme énergétique à J20. Les *Gene Ontology* (GO) catégories associées à l'organisation de la matrice extracellulaire, fortement associée aux mécanismes de fibrose, sont retrouvées dès J12 et jusqu'au J30, où elles représentent une part importante des annotations, suggérant un rôle d'IRX5 dans ces mécanismes.

La suite du travail s'est focalisée sur deux fonctions biologiques identifiées comme pouvant être régulées par IRX5 : la repolarisation électrique (car montrée chez la souris comme étant régulée par IRX5) et la fibrose cardiaque (identifiée dans nos analyses d'annotation fonctionnelle comme étant altérée dans les différenciations cardiaques IRX5^{KO}).

Investigation du rôle d'IRX5 dans la repolarisation électrique

La repolarisation électrique des cardiomyocytes garantit la relaxation cellulaire, nécessaire à une nouvelle activation des cardiomyocytes. Ces cycles d'activation et de relaxation électrique des cellules sont à la base de la contraction coordonnée de l'ensemble des cardiomyocytes (voir partie 1.1 Physiologie cardiaque). L'analyse de l'expression de génique de l'ensemble des canaux potassiques cardiaques repolarisants a été réalisée à J30 entre les cardiomyocytes Wt IRX5^{KO} (non présentée), et de manière intéressante, la plus forte variation d'expression observée est associée au gène codant une sous-unité du courant I_{to} , *KCNIP2* (augmentation d'environ 12 fois dans les cellules KO). Ce courant potassique I_{to} est à la base de la première phase de repolarisation du potentiel d'action (juste après la phase de dépolarisation rapide). Des anomalies de ce courant et donc de cette phase de repolarisation sont connues pour être une source majeure d'hétérogénéité et donc d'induction d'arythmies sévères. Chez la souris, ce courant est généré par la combinaison de deux sous-unités : une sous-unité α , *Kcnd2* (formant le pore du canal) et une sous-unité régulatrice *Kcnip2*. Il a d'ailleurs été précédemment démontré qu'*Irx5* impacte I_{to} via la régulation négative de *Kcnd2* dans la paroi ventriculaire murine (voir Figure 22). Chez l'humain, le courant I_{to} n'est pas généré par les mêmes sous-unités, suggérant qu'IRX5 pourrait avoir un rôle différent chez l'humain dans la régulation de ce courant. C'est la combinaison de la sous-unité α , *KCND3* et de la sous-unité régulatrice *KCNIP2* qui génère I_{to} chez l'humain. Nous avons donc analysé le profil d'expression des gènes codant ces sous-unités dans nos données de transcriptomique en cinétique et par qPCR (Figure 5A-C). De manière intéressante, la délétion d'IRX5 ne provoque pas d'altération de l'expression de *KCND3*, car il ne semble pas exprimé. Cependant, la délétion d'IRX5 entraîne une augmentation significative de l'expression de *KCNIP2*. Ces données sont en accord avec des données de la littérature montrant que dans la paroi ventriculaire du cœur humain (étude sur échantillons tissulaires humains sains), c'est *KCNIP2* et non pas *KCND3* qui présente un gradient d'expression inverse de celui d'IRX5 (Gaborit et al., 2010; Rosati et al., 2001).

Enfin, en cohérence avec ce résultat, une fixation d'IRX5 a été retrouvée à environ 7kb en aval du gène *KCNIP2* (Figure 5D), dans une zone ouverte et active (pic d'ATAC-seq encadré de pics d'H3K27Ac) qui présentent aussi une fixation de TBX5. Cette région correspond également au promoteur du gène *OGA*, mais dont l'expression n'est pas modifiée par la délétion d'IRX5. Ceci

suggère que cette zone pourrait être une région régulatrice ou *enhancer* de *KCNIP2* dans les cardiomyocytes humains.

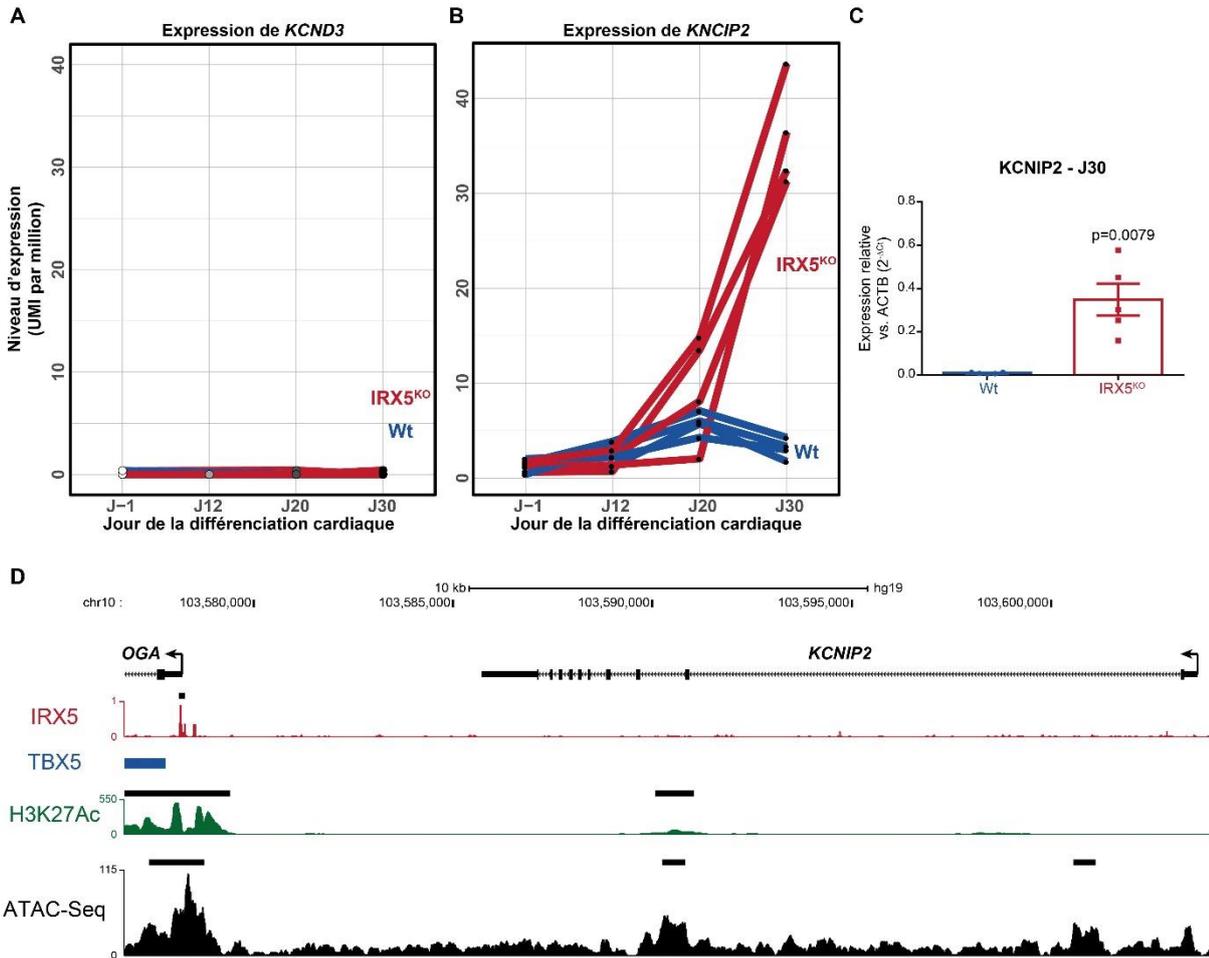


Figure 5 : IRX5 impacte l'expression de KCNIP2 dans la différenciation cardiaque humaine

A-B. Variations d'expression de *KCND3* et *KCNIP2* aux quatre jours étudiés de la différenciation cardiaque des iPS Wt et IRX5^{KO}, obtenue par 3'SRP. **C.** Expression relative de *KCNIP2* à J30 des différenciations Wt et IRX5^{KO}, quantifiée par qPCR. Contrôle positif : *ACTB*. Test statistique : Mann-Whitney. **D.** Représentation des intensités de signaux et pics de ChIP-Seq IRX5 (en rouge), TBX5 (en bleu) et H3K27Ac (en vert) dans la région génomique de *KCNIP2*. Les régions ouvertes de chromatine sont également indiquées, par les signaux (en noir) et pics significatifs (rectangles noirs) d'ATAC-Seq Wt à J30.

Investigation du rôle d'IRX5 dans la fibrose

Dans le but d'élucider le rôle d'IRX5 dans les mécanismes d'établissement de la fibrose cardiaque, le modèle de fibroblastes cardiaques dérivés d'iPS (iPS-CFs) a été mis en place. Les fibroblastes cardiaques obtenus en fin de différenciation (18 jours) ont ensuite été validés par imagerie en fond clair (*brightfield*) et immunofluorescence, de par leur forme caractéristique (cellules plates, en forme de fuseau) et la forte expression d'un marqueur spécifique, la protéine du cytosquelette vimentine (Figure 6A). Par ailleurs, l'expression protéique forte de

la vimentine dans les iPS-CFs (Figure 6B) et l'absence d'expression du marqueur cardiomyocytaire TBX5, ont été validées par western blot (Figure 6C). L'absence d'expression du marqueur de myofibroblaste, α -SMA, a également été validée par cytométrie en flux (données non présentées), validant le phénotype quiescent des fibroblastes obtenus. Enfin, l'expression protéique d'IRX5 semble être nulle dans les iPS-CFs (Figure 6D), suggérant que les impacts potentiels d'IRX5 dans les mécanismes fibrotiques passent par les cardiomyocytes.

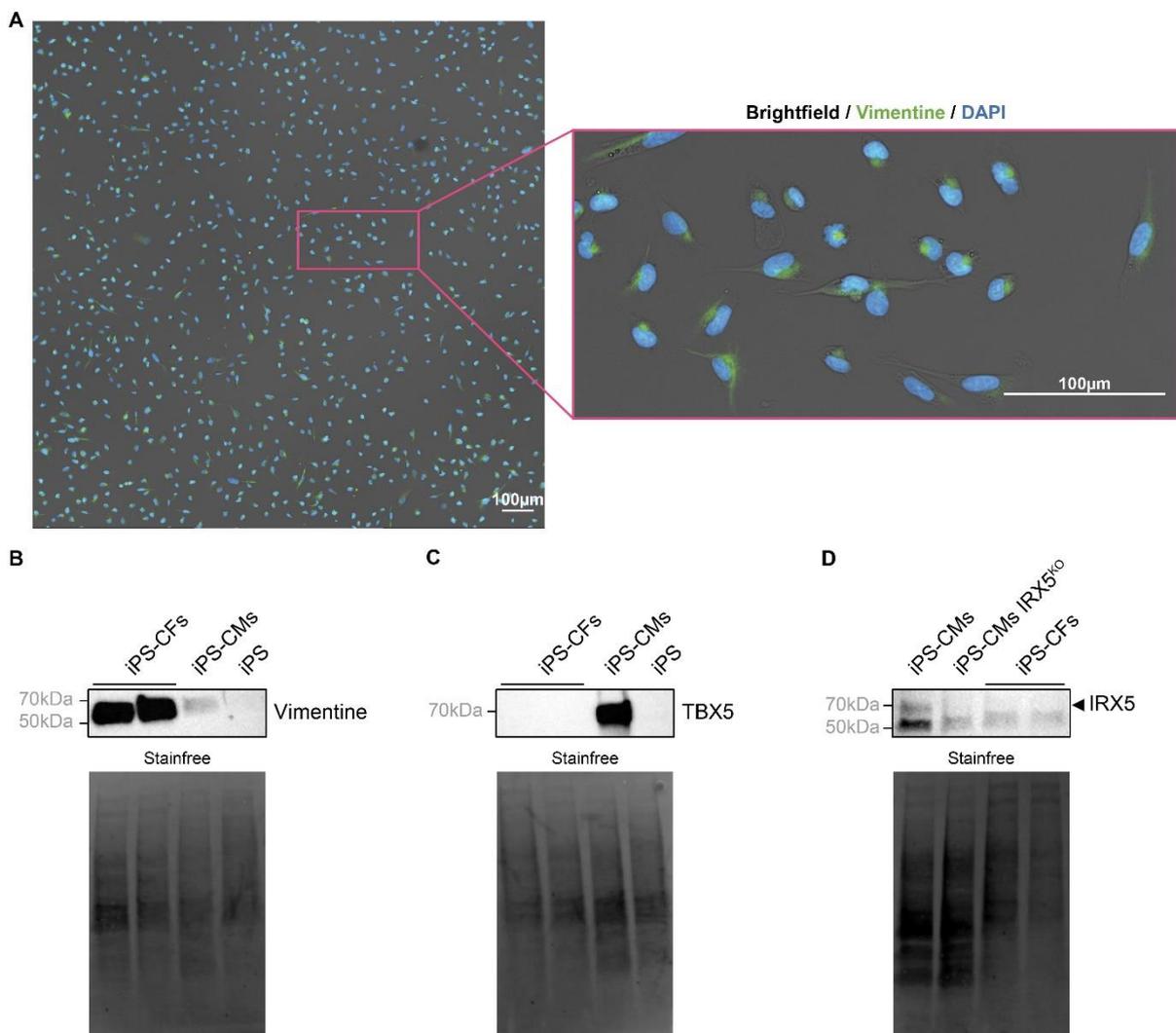


Figure 6 : Modèle de fibroblastes cardiaques dérivés d'iPS et expression d'IRX5

A. Image combinée d'immunofluorescence et fond clair de fibroblastes cardiaques dérivés d'iPS (iPS-CFs) en fin de différenciation. Immunomarquage vimentine (vert) et DAPI (bleu). Echelle : 100 μ m. **B.** Expression protéique de la vimentine dans les iPS-CFs en fin de différenciation. Les échantillons de cardiomyocytes dérivés d'iPS (iPS-CMs, J30) et de cellules souches iPS servent de contrôles négatifs. **C.** Expression protéique de TBX5 dans les iPS-CFs en fin de différenciation. Les échantillons iPS-CMs (J30) et iPS servent de contrôles positifs et négatifs respectivement. **D.** Expression protéique d'IRX5 dans les iPS-CFs en fin de différenciation. Les échantillons iPS-CMs Wt et IRX5^{KO} (J30) servent de contrôles positifs et négatifs respectivement. La bande d'intérêt (haute) est indiquée par la flèche noire.

Le modèle ayant été mis en place au laboratoire, il pourra à l'avenir être utilisé dans le cadre de co-cultures avec des cardiomyocytes Wt ou IRX5^{KO} pour identifier un rôle potentiel d'IRX5 dans le processus de fibrose cardiaque ([voir IV.2.D Perspectives pour la poursuite du projet 2](#)).

IRX5 impacte la fixation à l'ADN de GATA4 à J12 mais pas à J30

Afin d'identifier si les différences de profils d'expression de gènes au cours de la différenciation cardiaque Wt ou IRX5^{KO} étaient dû à l'implication variable de facteurs de transcription spécifiques, des expériences d'ATAC-Seq ont été réalisées à J12 et à J30 de deux différenciations contrôles et trois différenciations IRX5^{KO}. Dans l'objectif d'identifier des variations potentielles d'occupation de motifs de facteurs de transcription dans les régions ouvertes de chromatine, du fait de l'absence d'IRX5, les échantillons d'ATAC-Seq ont été bioinformatiquement combinés par jour et par condition (contrôle ou IRX5^{KO}) afin de réaliser une étude de *footprinting* ([voir partie I.4.C.c Techniques d'analyse](#)). En cohérence avec le nombre plus important de gènes différentiels à J12 par rapport à J30, des dynamiques de fixation de facteurs de transcription significativement impactées du fait de l'absence d'IRX5, ont été retrouvées à J12, mais pas à J30 (Figure 7). En effet à J12, 8 facteurs de transcription (des familles KLF et SP) ont été retrouvés avec une fixation à l'ADN plus importante dans les échantillons contrôles comparés aux IRX5^{KO}, et à l'inverse, 57 facteurs de transcription (principalement des familles GATA, FOX, ZNF ou MEF2) ont été retrouvés avec une fixation à l'ADN plus grande dans les échantillons IRX5^{KO} comparés aux contrôles. De manière intéressante, GATA4 spécifiquement, se fixe de manière plus importante à l'ADN dans les échantillons IRX5^{KO} suggérant qu'IRX5 a un impact sur sa fixation à l'ADN. Ces données sont en relevance directe avec une action d'IRX5 en complexe avec GATA4 ([IV.1.A - Article 1](#)).

Analyse préliminaire de ChIP-Seq à J12 et J30

Pour compléter l'étude des mécanismes pouvant expliquer l'obtention d'une plus grande proportion de gènes différentiels entre les cellules contrôles et IRX5^{KO} à J12 par rapport à J30, une nouvelle analyse de ChIP-Seq d'IRX5, d'IRX3 et de la marque H3K27Ac a été réalisée. Pour cela, des échantillons à J12 et J30 de deux cinétiques contrôles et IRX5^{KO} respectivement, qui avaient également été utilisées pour l'analyse d'ATAC-Seq, ont été utilisés.

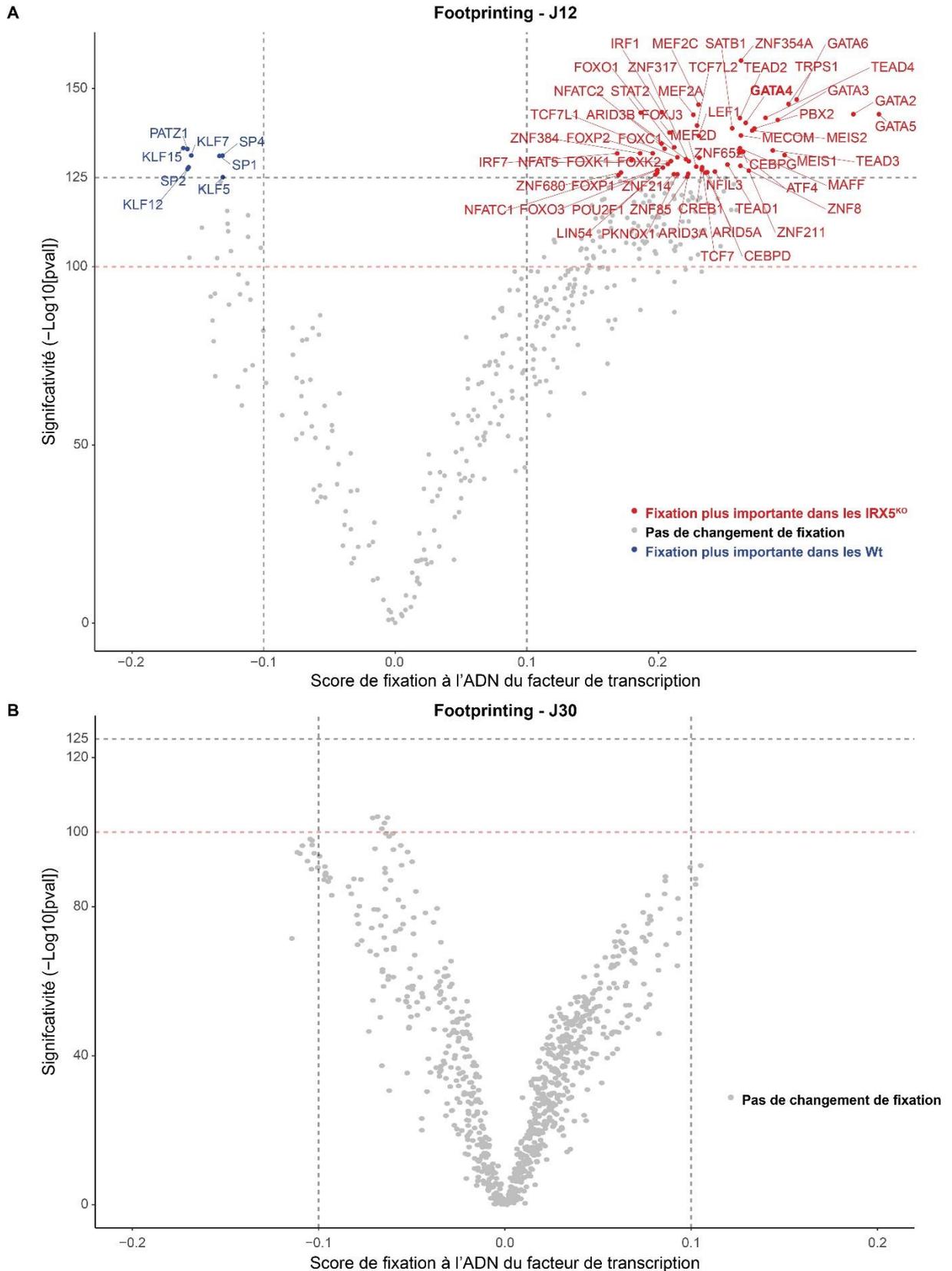


Figure 7 : Dynamique de fixation à l'ADN des facteurs de transcription à J12 et J30 entre les conditions contrôle et IRX5^{KO}

Pour identifier les variations de fixation des facteurs de transcription à l'ADN significatives, les seuils suivants ont été utilisés : $-\log_{10}(pval)$: 125 ; Score de fixation : 0.1.

Les échantillons de ChIP-Seq IRX3 et IRX5 n'ont cependant pas passé les contrôles qualités (non présentés), du fait du nombre très faible de pics identifiés et d'un manque de reproductibilité entre les réplicats. À l'inverse, les échantillons de ChIP-Seq H3K27Ac ont tous passé les contrôles qualité, avec un enrichissement typique en double-pic autour des TSS des gènes (qui est associé aux nucléosomes -1 et +1) à J12 et à J30 (Figure 8A-B). De plus, de manière cohérente avec les liens connus entre la marque H3K27Ac et l'activation de la transcription, les comparaisons à J12 et J30 entre les données de transcriptome et les résultats de ChIP-Seq H3K27Ac ont bien permis d'observer que la majorité des gènes exprimés avait des marques H3K27Ac associées, et à l'inverse, la majorité des gènes non exprimés n'avait pas de marques H3K27Ac associées que ce soit pour les cardiomyocytes Wt ou IRX5^{KO} (Figure 8C-D). L'ensemble de ces résultats sur la marque H3K27Ac valide la qualité des échantillons biologiques de départ, ainsi que l'expérience de ChIP-Seq réalisée. Les modifications épigénétiques associées à l'action des facteurs de transcription, régulent finement l'expression des gènes. Les données obtenues pour les marques H3K27Ac suggèrent de

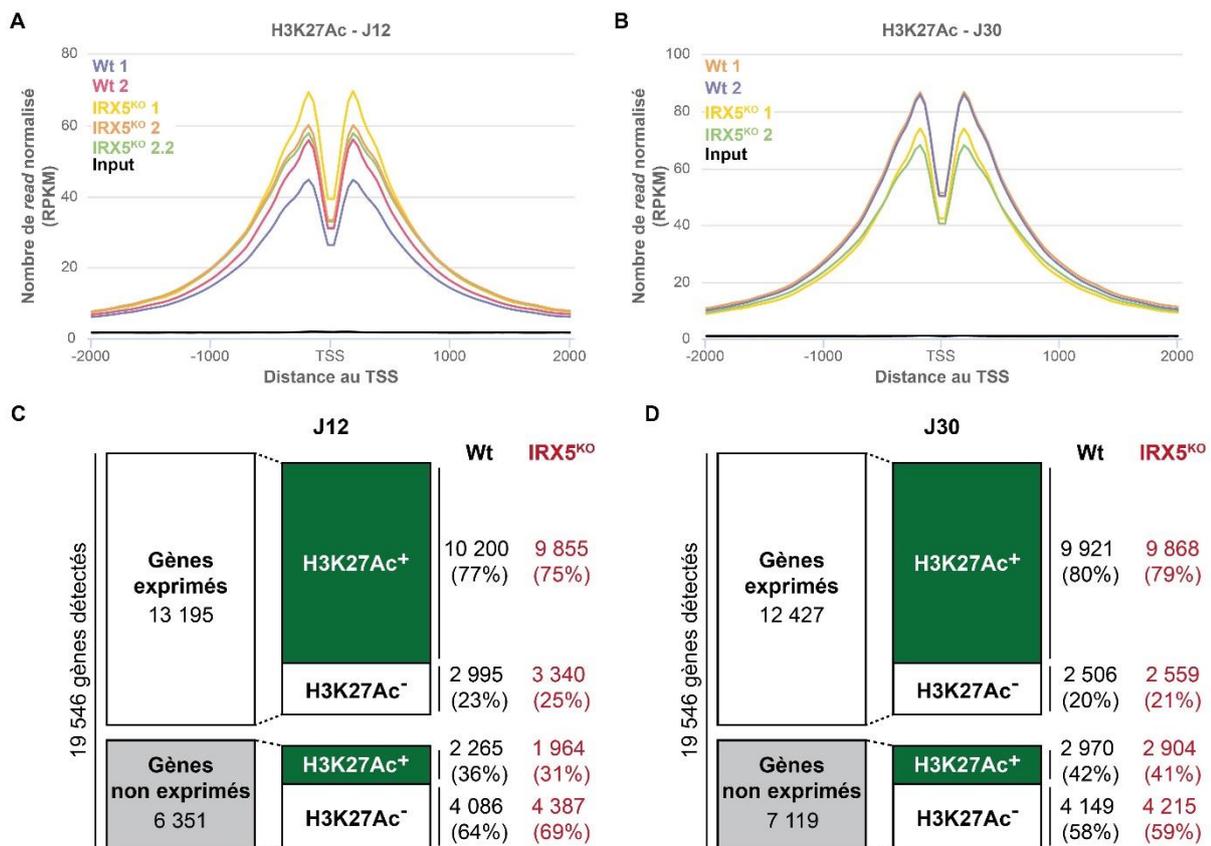


Figure 8 : Validation des ChIP-Seq H3K27Ac à J12 et J30 des différenciations cardiaques d'iPS

A-B. TSS-plot représentant la distribution des reads autour du *Transcription Start Site* (TSS) des gènes. Unité en *Reads Per Kilobase per Million mapped reads* (RPKM). C-D. Comparaison entre les gènes exprimés ou non exprimés et les gènes arborant des marques H3K27Ac dans les échantillons Wt et IRX5^{KO}.

manière consistante que l'absence d'IRX5 semble être associée à une diminution du pourcentage de gènes présentant des marques d'acétylation et en cohérence, à une augmentation du pourcentage de gènes non acétylés. Si ces données et tendances préliminaires sont par la suite confirmées, ces différences globales (à l'échelle du génome) devront à l'avenir être investiguées plus finement, d'une part en les comparant aux variations transcriptomiques associées à la délétion d'IRX5, et d'autre part, en étudiant la possibilité que l'absence d'IRX5 ait entraîné des changements de marques épigénétiques (méthylation, phosphorylation, ...) sur les gènes concernés.

IV.2.D. Perspectives pour la poursuite du projet 2

Ce projet avait pour but d'identifier sans a priori, les fonctions spécifiques et redondantes d'IRX3 et d'IRX5 dans un contexte cardiaque humain. Les modèles cellulaires et résultats générés pour ce projet ont principalement été focalisés sur l'étude d'IRX5 dans un premier temps. Par conséquent, la réalisation d'un travail similaire sur IRX3 sera nécessaire pour compléter ce travail. Pour cela, de nouveaux modèles cellulaires KO pour IRX3 et double-KO IRX3/IRX5 devront être générés et permettront d'identifier les fonctions spécifiques à IRX3, ainsi que les celles redondantes entre IRX3 et IRX5.

Dans le cadre de l'étude du rôle des différents domaines d'IRX5 dans ses interactions protéiques, l'homéodomaine d'IRX5 a été montré comme spécifiquement impliqué, lui accordant ainsi une fonction supplémentaire à celle déjà connue pour l'interaction à l'ADN (Kim et al., 2012). Les régions N-terminale et IRO-box sont encore à analyser dans le contexte des interactions protéiques d'IRX5 et de nouvelles constructions d'IRX5 suivies d'expériences de Co-IP devront être réalisées. Les résultats obtenus permettront notamment de confirmer le rôle supposé de l'IRO-box dans les interactions protéiques d'IRX5, qui n'a pas encore été spécifiquement analysé ([voir partie I.4.B Les facteurs de transcription Iroquois homeobox](#)) (Barrios and Campuzano, 2016; Cavodeassi et al., 2001).

Pour l'étude de la localisation génomique d'IRX5, deux expériences de CHIP-Seq ont été réalisées. Bien que la première expérience de CHIP-Seq ait permis d'identifier 2 232 régions du génome occupées par IRX5, la seconde a montré des problèmes de reproductibilité ne permettant pas de poursuivre l'analyse. L'utilisation d'anticorps différents pour IRX5 ([voir](#)

partie III.2.D.c Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 2) entre les deux expériences, bien que tous validés par western blot, est certainement impliquée dans les différences de résultats obtenus entre les deux expériences. Ce résultat souligne la nécessité d'utilisation d'anticorps adaptés, constituant la principale limite de la technique de ChIP-Seq (voir partie I.4.C.c Techniques d'analyse). L'autre limite étant le niveau assez faible d'expression d'IRX5. L'obtention de bons résultats de ChIP-Seq H3K27Ac (dont l'anticorps est bien validé pour cette application), réalisé en parallèle de la seconde expérience de ChIP-Seq IRX5, a permis la validation de l'intégrité des échantillons et de la qualité du séquençage, renforçant l'hypothèse d'un problème d'anticorps pour les ChIP-Seq IRX5. Pour pallier ce problème de reproductibilité lié aux anticorps IRX5, dans le cas de la première expérience de ChIP-Seq, les 6 réplicats biologiques ont été intégrés avec l'outil *NGS-Integrator*, qui est adapté pour améliorer le rapport signal/bruit de fond, notamment dans le cas d'analyse de ChIP-Seq (voir partie III.2.B.c Chromatin Immunoprecipitation Sequencing (ChIP-Seq) – Expérience 1). De plus, afin de confirmer et approfondir les résultats de ChIP-Seq obtenus, tout en s'affranchissant des contraintes liées à la spécificité des anticorps, un modèle iPS exprimant la protéine IRX5 taguée avec la séquence HA a été généré (IV.2.B - Article 2). En tirant parti de la grande spécificité des anticorps HA et de leur validation dans le cadre d'applications comme le ChIP-Seq, de nouvelles différenciations cardiaques combinées à des expériences de ChIP-Seq IRX5-HA devront être réalisées. La génération d'une autre lignée iPS exprimant la protéine IRX3 taguée avec une autre séquence permettra également l'identification des régions génomiques occupées par IRX3, sans les contraintes de spécificité d'anticorps. Enfin, l'utilisation d'anticorps spécifiques de séquences tag offre également la possibilité de réaliser des expériences de ChIP-Seq en *single-cell*, qui permettraient notamment l'étude des régions occupées par IRX3 et IRX5 dans les différentes sous-populations de cellules cardiaques.

L'analyse des marques épigénétiques dans l'étude des fonctions et mécanismes d'action d'IRX5 semble pertinente, du fait de son association démontrée chez la souris avec des partenaires recrutant des modificateurs épigénétiques (mBop recrute HDAC ; Costantini et al., 2005). Les données de ChIP-Seq H3K27Ac ont permis de valider la qualité des échantillons et la technique de ChIP-Seq, mais suggèrent également un potentiel impact de la perte d'IRX5 sur le dépôt de marques H3K27Ac dans certaines régions du génome. Ce résultat est en cohérence avec l'association d'IRX5 avec des régulateurs épigénétiques, mais nécessite

confirmation par des analyses spécifiques des régions présentant différemment la marque H3K27Ac, entre les cellules Wt et IRX5^{KO}. On peut également faire l'hypothèse qu'IRX5 pourrait avoir un impact sur le dépôt d'autres marques épigénétiques et altérer l'expression des gènes par un mécanisme plus complexe. Ainsi, ces données seront également à compléter avec celles d'autres marques d'histones, afin d'obtenir une vision complète du paysage épigénétique associés aux régions génomiques occupées par IRX5 et celle impactées par sa délétion.

L'étude de l'impact d'IRX5 sur la dynamique de fixation à l'ADN des autres facteurs de transcription, réalisée par analyse de *footprinting*, a permis l'identification d'une fixation plus importante de GATA4 dans les échantillons IRX5^{KO}. Ce résultat suggère un rôle d'IRX5 dans la fixation de GATA4 à l'ADN, ce qui est concordant avec leur action en complexe multiprotéique démontrée lors du projet 1. De plus, la fixation accrue de GATA4 à l'ADN pourrait refléter un mécanisme de compensation pour la perte d'IRX5. Afin de valider l'établissement d'un tel mécanisme, l'analyse des variations d'intensités de fixation de GATA4 en présence ou non d'IRX5 sera à effectuer, grâce à la génération de données de CHIP-Seq GATA4 au cours de différenciations cardiaques d'iPS contrôles et IRX5^{KO}.

La régulation de la production de matrice extracellulaire et l'établissement de la fibrose cardiaque étant fortement associés aux fibroblastes cardiaques ([voir partie I.1 Physiologie cardiaque](#)), l'implication d'IRX5 dans ces mécanismes de fibrose sera à poursuivre en utilisant le modèle iPS-CFs récemment mis en place. IRX5 ne semblant pas être exprimé dans les fibroblastes, l'analyse des surnageants de co-cultures directes entre des iPS-CMs contrôles ou IRX5^{KO} et des iPS-CFs contrôles, permettra d'identifier si des protéines associées aux mécanismes de fibrose cardiaque (TGF- β , collagène, métalloprotéase, ...) sont différemment sécrétées du fait de la délétion d'IRX5 dans les cardiomyocytes, démontrant ainsi l'impact d'IRX5 dans les mécanismes de fibrose. Pour investiguer plus en détail cet impact potentiel d'IRX5, les co-cultures pourront être réalisées dans un contexte stable ou stimulé par un stress pharmacologique (stimulation adrénergique, à l'angiotensine, ...) et permettre ainsi l'identification du rôle d'IRX5 sur les mécanismes de fibrose dans ces différents contextes. Afin d'analyser par quelles voies les cardiomyocytes et les fibroblastes communiquent entre eux pour déclencher ces mécanismes fibrotiques potentiellement impactés par IRX5, la

comparaison des résultats d'expériences de co-cultures directes et indirectes permettrait d'identifier si les interactions se font par contacts physiques ou par effet paracrine.

V

Discussion générale

Ce travail de thèse, à travers les deux projets principaux qui ont été menés, a permis d'apporter de nouvelles connaissances sur les facteurs de transcription cardiaques et en particulier sur les facteurs de transcription Iroquois homeobox (Figure 29-Figure 30).

La réalisation d'une étude transcriptomique en cinétique au cours de la différenciation cardiaque de cellules iPS humaines, a ainsi permis (1) de valider les iPS comme un modèle pertinent d'étude d'expression génique du développement cardiaque humain, et (2) d'identifier les vagues d'expression de gènes qui s'enchainent chronologiquement au cours de la différenciation (Figure 29). Un focus spécifique sur les facteurs de transcription exprimés dans ces vagues a ensuite été réalisé, dans le but de mettre à jour le réseau global de régulation, tenant compte du temps, entre les facteurs de transcription les plus impliqués dans la différenciation cardiaque (n=216). De nombreuses relations, jusqu'alors inconnues, d'activation et d'inhibition ont pu être identifiées entre les facteurs de transcription. Cette étude nous a permis de positionner les facteurs de transcription IRX3 et IRX5 dans le réseau global des facteurs de transcription impliqués dans la différenciation cardiaque. De nouveaux mécanismes de régulation ont alors pu être découverts et validés biologiquement entre IRX3, IRX5 et trois facteurs de transcription centraux du développement et de la fonction cardiaque, GATA4, TBX5 et NKX2-5. D'une part, IRX3 et IRX5 activent successivement leur expression au cours de la différenciation cardiaque, et d'autre part, ils interagissent physiquement avec eux sous la forme de complexes multiprotéiques fonctionnels, pour l'activation de l'expression d'un gène cible, *SCN5A*.

La réalisation d'une seconde étude spécifique d'IRX5 a permis d'apporter de nouvelles informations sur ses domaines d'interaction protéique, en particulier son homéodomaine, ainsi que sur sa localisation génomique (Figure 30). L'analyse des gènes occupés par IRX5 et par GATA4, TBX5 et NKX2-5 a renforcé le lien fonctionnel d'interaction identifié précédemment entre ces facteurs de transcription, et a également identifié un rôle potentiel d'IRX5 dans l'orientation de l'action régulatrice du complexe GATA4/TBX5/NKX2-5, pour la régulation de l'activité électrique des cardiomyocytes. Enfin, la réalisation d'un nouveau modèle iPS humain IRX5^{KO} a permis l'identification de mécanismes potentiellement régulés par IRX5, comme ceux associés à l'activité électrique et la fibrose cardiaque. IRX5 ayant un mécanisme d'action supposé faisant intervenir des modificateurs d'histones (Costantini et al., 2005), une analyse multi-omique a également été débutée afin d'identifier les impacts de la

perte d'IRX5 sur le « paysage » épigénétique, ainsi que sur l'activité d'autres facteurs de transcription qui pourraient être à l'origine de mécanismes de compensation.

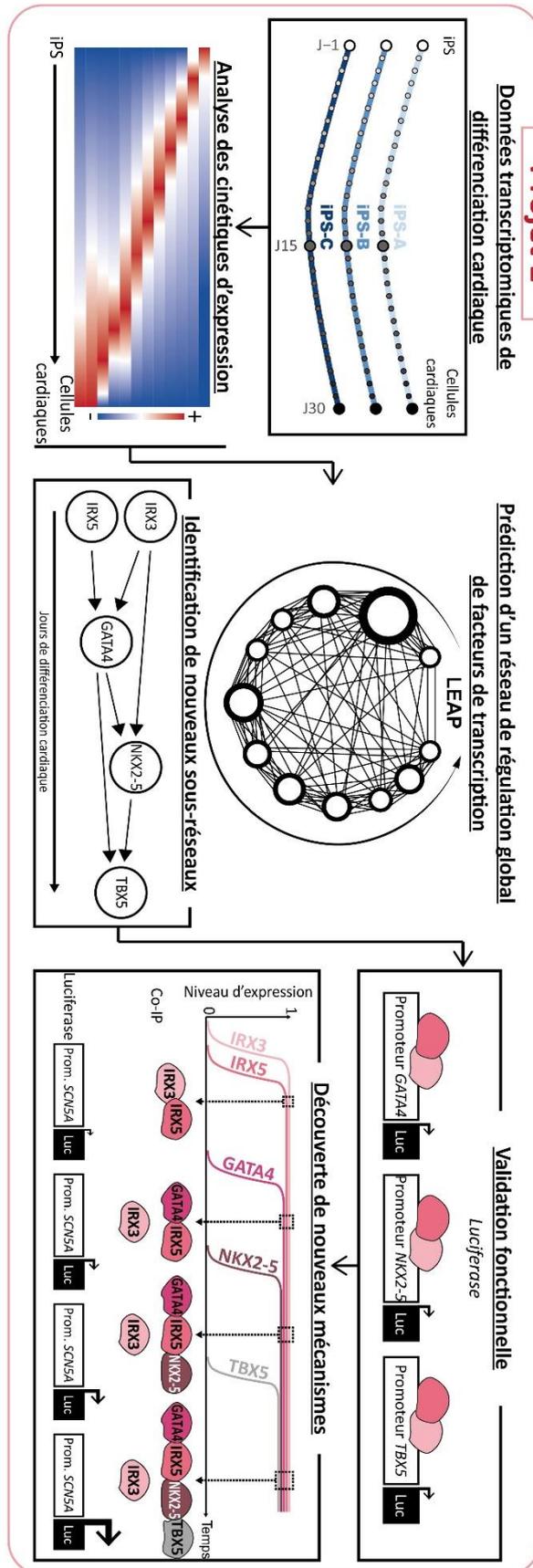


Figure 29 : Récapitulatif des résultats du projet 1

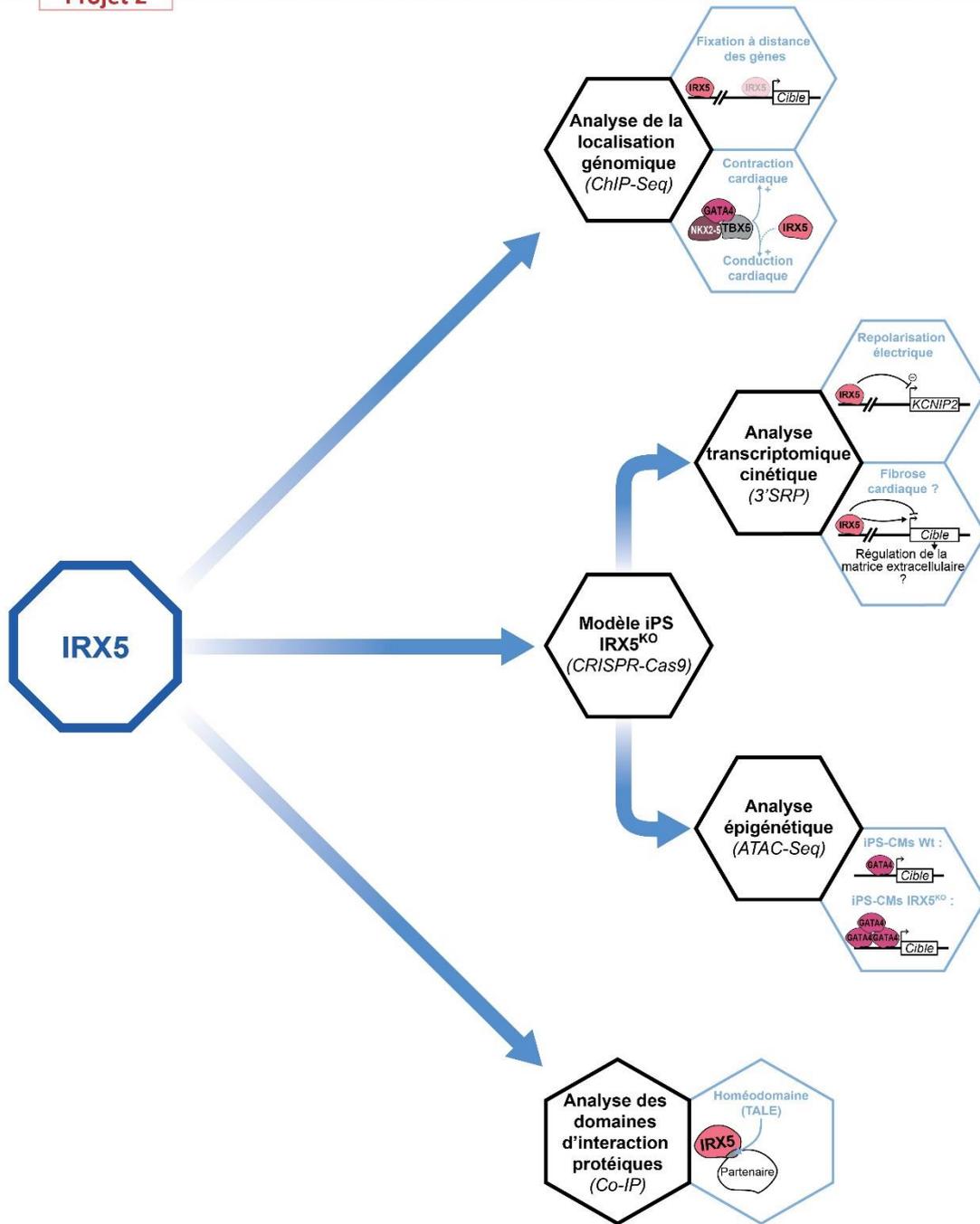


Figure 30 : Récapitulatif des résultats du projet 2

V.1.A. Intégration de la famille Iroquois dans le réseau global de facteurs de transcription du développement cardiaque

Les facteurs de transcription de la famille Iroquois homeobox et leurs relations avec d'autres facteurs de transcription cardiaques n'ont jusqu'à présent été que peu investigués. Pour IRX3, il a été montré chez la souris, qu'il pouvait coopérer avec les facteurs de

transcription Nkx2-5 et Tbx5 dans le développement et la maturation du système de conduction ventriculaire (Kim et al., 2016). Pour IRX5, il a été montré qu'il pouvait potentiellement interagir avec Irx4 dans la régulation du gradient transmural du potentiel d'action (He et al., 2009). La construction d'un réseau global de facteurs de transcription au cours de la différenciation cardiaque d'iPS humaines (IV.1.A - Article 1) nous a permis de placer IRX3 et IRX5 dans un contexte de régulation intégré entre les facteurs de transcription. L'expression d'IRX3 et IRX5 dans le premier cluster de gènes ayant une expression maintenue jusqu'à la fin de la différenciation, suggère des rôles de ces facteurs de transcription plus précoces et plus vastes dans le développement cardiaque, que celui démontré jusqu'à présent dans la littérature sur la régulation de l'activité électrique cardiaque. A l'aide du réseau global de facteurs de transcription que nous avons construit, nous avons pu investiguer les rôles d'IRX3 et IRX5 sur l'expression de facteurs de transcription majeurs du développement et de la fonction cardiaque : GATA4, TBX5 et NKX2-5, et avons également mis en évidence qu'ils pouvaient agir en complexes multiprotéiques pour la régulation de *SCN5A* (Figure 31). D'autres cibles potentielles d'IRX3 et IRX5 auraient également pu être investiguées, comme par exemple le facteur de transcription FOXP1, impliqué dans la prolifération des cardiomyocytes et dont le KO provoque une létalité embryonnaire chez la souris due à un amincissement du myocarde ventriculaire (Golson and Kaestner, 2016). FOXP1 a d'une part, été identifié par notre analyse globale comme une cible d'IRX3 et IRX5 (IV.1.A - Article 1), et d'autre part, sa fixation à l'ADN semble modifiée dans les iPS-CMs IRX5^{KO} (données d'ATAC-seq, Projet 2). Ces résultats suggèrent un impact potentiel d'IRX3 et IRX5 dans la prolifération des cardiomyocytes ventriculaires, en accord avec leur impact supposé dans le cycle cellulaire (Barrios et al., 2015), qui passerait par la régulation de FOXP1, mais qui reste à démontrer formellement.

La découverte des effets activateurs d'IRX3 et IRX5 sur l'expression de GATA4, TBX5 et NKX2-5, qui sont exprimés dès la formation du croissant cardiaque, renforce l'idée d'une action précoce insoupçonnée pour IRX3 et IRX5 dans le développement cardiaque. L'expression d'Irx5 dans le développement cardiaque de la souris a été montrée comme débutant à E9 dans le tube cardiaque en looping (Christoffels et al., 2000). Cependant dans cette étude, l'analyse de l'expression cardiaque d'Irx5 a été réalisée par hybridation *in situ*, permettant de marquer des ARN spécifiques dans des coupes de tissus, uniquement sur des

échantillons après l'établissement du croissant cardiaque. D'autres études, focalisées sur la neurogenèse ont identifié par hybridation *in situ*, une expression d'Irx3 dès E6.5 pendant la gastrulation principalement dans des cellules dérivées du trophoctoderme (précurseurs du placenta) puis dans l'ectoderme à E7.5 (Bosse et al., 1997). Une expression d'Irx5 a été montrée uniquement à partir de E8 sans indication de localisation dans l'embryon (Cohen et al., 2000). Ces études n'ont pas analysé spécifiquement les dérivés de l'épiblaste qui donneront notamment le mésoderme, de plus, la technique d'hybridation *in situ* basée sur l'imagerie, a pour limite la difficulté d'identification de gènes faiblement exprimés, avec peu de copies d'ARN (Jensen, 2014). Ces résultats suggèrent une expression précoce potentielle d'IRX3 et IRX5 dans le développement, mais qui n'a pas été démontrée dans les progéniteurs cardiaques. Nos données sont pourtant en accord avec une expression très précoce d'IRX3 et IRX5 dans la différenciation cardiomyocytaire. En effet, l'apparition de l'expression d'IRX3 et IRX5 est observée dans le modèle de différenciation cardiaque d'iPS (IV.1.A - Article 1), aux alentours des jours 3-4 de la différenciation, un ou deux jours respectivement après l'expression des facteurs de transcription EOMES et MESP1 (marqueur des précurseurs du mésoderme cardiogénique). Plus spécifiquement pour IRX5, de manière concordante avec ces résultats, dans le réseau global de facteurs de transcription que nous avons généré (IV.1.A - Article 1), IRX5 serait une cible activée par le facteur de transcription ID1 (avec un lag de 1 jour), qui est déjà connu pour activer l'expression de MESP1 et spécifier le mésoderme cardiaque (Cunningham et al., 2017). Par ailleurs, nous avons même montré par qPCR une expression (bien que très faible) d'IRX5 très précocement dans la différenciation cardiaque des iPS, dès le stade de pluripotence (voir IV.2.C Résultats complémentaires). L'ensemble de ces résultats suggère ainsi qu'IRX3 et IRX5 pourraient avoir une expression et des impacts plus précoces dans le développement cardiaque que ceux identifiés chez la souris, potentiellement dès les étapes de gastrulation, dans les progéniteurs mésodermiques avant et pendant la formation du croissant cardiaque (Figure 31).

L'expression précoce potentielle des facteurs de transcription IRX3 et IRX5 pose la question de leur interaction fonctionnelle précoce avec d'autres facteurs de transcription cardiaques. Il est en effet suggéré que les facteurs de transcription des familles HOX (gènes HOX) et TALE (gènes MEIS et PBX) pourraient agir en coopération dans le développement cardiaque, notamment pendant la gastrulation (Lescroart and Zaffran, 2018). De manière intéressante, IRX3 et IRX5, qui sont aussi des facteurs de transcription TALE, ont été retrouvés

dans le même cluster d'expression que des facteurs de transcription HOX (HOXB2, HOXB3 et HOXB4) dans le réseau global présenté précédemment (IV.1.A - Article 1). Cette co-expression dans le réseau global de facteurs de transcription suggère la possibilité d'interactions physiques et/ou fonctionnelles entre IRX3, IRX5 et des facteurs de transcription de la famille HOX. Il a été montré que les facteurs de transcription HOX sont impliqués dans le développement craniofacial (Thesleff, 1995; Zaffran et al., 2018). Il serait donc intéressant d'étudier les interactions potentielles entre IRX et HOX, en lien avec le phénotype des patients Hamamy qui ont, en plus des défauts cardiaques, des malformations craniofaciales importantes (Bonnard et al., 2012).

Pour pouvoir valider et approfondir l'étude de l'expression précoce d'IRX3 et IRX5, ainsi que leurs rôles tout au long du développement cardiaque, la combinaison d'expériences de *lineage tracing* et de *single-cell* RNA-Seq pourrait être entreprise. Le *lineage tracing* désigne un ensemble de méthodes permettant de suivre le destin de cellules individuelles et des cellules qui en sont issues, avec une perturbation minimale de leur fonction physiologique (Wu et al., 2019). Dans le cadre de l'étude d'IRX3 et IRX5 dans le développement cardiaque précoce, le *lineage tracing* permettrait d'identifier par imagerie, les types cellulaires les exprimant, ainsi que leur devenir au cours du développement cardiaque embryonnaire. La combinaison de cette approche, avec des approches de *single-cell* RNA-Seq permettrait d'obtenir des informations sur le transcriptome des différents progéniteurs exprimant IRX3 et IRX5, et ainsi de réaliser un *lineage tracing* informatique, basé sur l'expression de ces gènes (Kester and van Oudenaarden, 2018). Cette analyse permettrait d'obtenir les transitions transcriptomiques s'effectuant entre les différents progéniteurs au cours du temps, et ainsi d'identifier les rôles et mécanismes d'action spécifiques d'IRX3 et IRX5 au cours de la cardiogénèse. L'utilisation du modèle murin semble adaptée à ces types d'investigation et à l'emploi de techniques en cinétique au cours du développement embryonnaire, malgré l'existence de différences entre le développement murin et humain. Afin d'utiliser un modèle humain, l'utilisation de cellules iPS pour générer des *gastruloids* (modèle organoïde en 3D) qui reproduisent le développement précoce du cœur, de la génération des précurseurs à la spécification des champs cardiaques (Rossi et al., 2021), serait également adaptée à ces approches combinées d'imagerie et de *single-cell* RNA-Seq.

V.1.B. Impact des facteurs de transcription IRX3 et IRX5 dans le gradient ventriculaire transmural d'expression des gènes

Les facteurs de transcription IRX3 et IRX5 ont une expression en gradient au sein de la paroi ventriculaire, du myocarde sous-endocardique (où ils sont le plus exprimés) vers le myocarde sous-épicaire (Gaborit et al., 2010) et IRX5 y a notamment déjà été impliqué dans la régulation de certains gènes associés à l'activité électrique (Al Sayed et al., 2021a; Costantini et al., 2005). D'autres gènes présentent aussi une expression en gradient au sein de la paroi ventriculaire, tel que des gènes codant des canaux ioniques, garantissant ainsi une repolarisation correcte des ventricules (Gaborit et al., 2007; Kim et al., 2022; McKinnon and Rosati, 2016). Ces études, principalement focalisées sur les canaux ioniques, ne présentent cependant pas l'intégralité des variations transcriptomiques s'établissant au sein de la paroi ventriculaire, qui reste donc à étudier de manière non biaisée. L'étude du gradient d'expression transmural est importante car, son rôle étant essentiel dans la repolarisation cardiaque, des altérations de ce gradient vont participer à la génération d'un substrat arythmogène responsable de troubles électriques (Akar and Rosenbaum, 2003; Boukens et al., 2016). D'autres pathologies semblent aussi être en partie associées à des défauts du gradient d'expression ventriculaire transmural, comme l'insuffisance cardiaque par exemple (Lou et al., 2011). De plus, la présence des cellules M (ou cardiomyocytes du mid-myocarde), présentant un potentiel d'action plus long que les cardiomyocytes sous-endocardiques et sous-épicaire, est toujours controversée (Boukens et al., 2016). Ainsi, il est encore nécessaire d'approfondir et d'élargir les études sur les gradients d'expression au sein de la paroi ventriculaire. Dans ce contexte, par leur expression en gradient, identifier comment IRX3 et IRX5 participent plus largement à la mise en place et au maintien du gradient d'expression ventriculaire transmural, serait particulièrement intéressant.

D'autres facteurs de transcription ont aussi une expression en gradient au sein de la paroi, et pourraient agir en coopération ou en antagonisme avec IRX3 et IRX5 dans cette mise en place du gradient transmural. D'une part, nous avons montré qu'IRX3 et IRX5 pouvaient agir en coopération fonctionnelle avec les facteurs de transcription GATA4, NKX2-5 et TBX5, notamment sous la forme de complexes multiprotéiques. Contrairement à Gata4 et Nkx2-5 (exprimés de manière homogène dans la paroi ventriculaire), Tbx5 présente une expression

en gradient similaire à IRX3 et IRX5 (Boukens et al., 2009), suggérant qu'IRX3, IRX5 et TBX5 pourraient orienter l'action du complexe dans les cellules où ils sont co-exprimés, pour permettre la mise en place de l'expression en gradient de gènes cibles. D'autre part, le facteur de transcription Hey2 a une expression en gradient inverse par rapport à IRX3 et IRX5 dans le ventricule (Veerman et al., 2017) et nous avons déjà montré l'augmentation de son expression dans des iPS-CMs porteurs de mutations perte-de-fonction Hamamy (Al Sayed et al., 2021a), suggérant qu'il est normalement inhibé par IRX3 et IRX5, pour permettre l'établissement du gradient d'expression ventriculaire.

Pour pouvoir étudier en détails ces régulations potentielles, des modèles humains adéquats sont nécessaires. En effet, bien que nous ayons construit un réseau global de facteurs de transcription au cours de la différenciation cardiaque, la fonctionnalité des complexes de facteurs de transcription étant associée à leur co-expression dans les mêmes cellules, l'utilisation d'analyses transcriptomiques en *bulk* ne permet pas d'obtenir ce niveau d'information. Des analyses en *single-cell* sont ainsi à considérer pour pouvoir étudier la mise en place de réseaux de régulation cellule-spécifiques et pour ainsi mieux comprendre l'établissement du gradient ventriculaire transmural. De plus, les analyses en *single-cell* sont à combiner avec des modèles reflétant plus fidèlement la complexité de la paroi ventriculaire, et par conséquent les modèles de différenciation d'iPS en 2D sont peu pertinents pour ces questions ([voir partie I.4.C.b Cardiomyocytes dérivés de cellules souches humaines](#)). A l'inverse, à l'image des stratégies de création d'atlas cellulaires du cœur (Litviňuková et al., 2020; Tucker et al., 2020), l'utilisation d'échantillons de paroi ventriculaire de cœurs humains, combinée à des analyses de *single-cell* RNA-Seq, permettrait d'étudier la complexité spatiale des réseaux de régulation des gènes et des types cellulaires, associés à IRX3 et IRX5 au niveau transmural. Afin d'étudier plus spécifiquement la mise en place temporelle du gradient d'expression transmural, des modèles dérivés d'iPS en 3D sont aussi à considérer, comme les organoïdes cardiaques (ou cardioïdes) par exemple (Cho et al., 2022; Hofbauer et al., 2021). En effet, ces modèles 3D miment le développement des chambres cardiaques (notamment les ventricules) et aboutissent à la formation de structure 3D composées de cardiomyocytes sous-endocardiques et sous-épicaudiques, ainsi que de cellules endocardiques et épicaudiques. Ils devraient donc permettre l'étude du gradient d'expression transmural, notamment par l'association avec des techniques transcriptomiques et épigénétiques en *single-cell*. Par ailleurs, augmentant encore d'avantage la puissance de ces technologies 3D, elles peuvent

être associées à des éditions génétiques (notamment par CRISPR-Cas9), rendant ainsi possible l'étude spécifique de gènes d'intérêt, comme *IRX3* et *IRX5*, dans ces contextes.

De manière complémentaire à l'étude des rôles d'*IRX3* et *IRX5* dans l'établissement et le maintien du gradient transmural d'expression de gènes, la question de leur propre régulation est aussi à développer. En effet, les mécanismes de régulation génique étant dépendants de la co-expression de facteurs de transcription, l'identification des mécanismes déclencheurs de l'expression d'*IRX3* et *IRX5* permettrait de mieux comprendre le développement et le maintien du gradient ventriculaire transmural, mais aussi des pathologies associées. Notre étude a permis l'identification d'un réseau global de facteurs de transcription dans un contexte de différenciation cardiaque et nous avons choisi d'étudier les cibles d'*IRX3* et *IRX5* apparaissant plus tardivement dans la différenciation ([IV.1.A - Article 1](#)). Par une stratégie inverse, nos données peuvent également permettre l'identification des régulateurs d'*IRX3* et *IRX5* dans la différenciation cardiaque. Des régions ADN intergéniques (non codantes) entre *Irx3* et *Irx5*, très conservées entre les espèces, ont déjà été identifiées comme étant des enhancers potentiels de ces gènes (Calle-Mustienes et al., 2005; Chua et al., 2022; Tena et al., 2011). De manière concordante, une récente étude de GWAS (*Genome Wide Association Study*) comparant les génotypes de 2 820 patients atteint du syndrome de Brugada à ceux de 10 001 individus contrôles, a identifié deux polymorphismes de l'ADN (SNPs) surreprésentés chez les patients Brugada et localisés dans la région intergénique entre *IRX3* et *IRX5* (Barc et al., 2022). Les patients avec un syndrome de Brugada, dont l'étiologie n'est pas connue pour la majorité des cas, présentent une élévation du segment ST sur l'ECG (caractéristique de la pathologie), associée à un risque accru de mort subite. La découverte de ces polymorphismes dans des régions identifiées par ailleurs comme étant de potentiels enhancers d'*IRX3* et *IRX5*, suggère un rôle de ces facteurs de transcription dans le développement du syndrome de Brugada et souligne l'importance d'étudier leur propre régulation, en plus de leurs modes d'action et cibles. Cependant, pourquoi ces SNPs augmentent le risque pathologique, reste encore à être déterminé. Ces polymorphismes ne constituent pas des mutations rares et causales à eux seuls mais à l'inverse sont fréquemment retrouvés dans la population générale (c'est l'association de plusieurs polymorphismes fréquents qui augmente le risque de déclenchement de la pathologie). Ainsi, pour étudier leur impact sur la fonction cardiomyocytaire, on peut faire l'hypothèse que la modification

spécifique et individuelle de ces SNPs ne sera pas suffisante pour observer un effet sur le phénotype cellulaire. Par contre, de manière intéressante, ils sont localisés dans des régions non-codantes de quelques centaines de paires de bases dont les caractéristiques épigénétiques leur confèrent un possible caractère de région régulatrice ou *enhancer* (région ouverte, ATACseq-positive, et transcriptionnellement active, H3K27Ac-positive). Il paraît donc plus pertinent d'étudier l'impact de la délétion de l'ensemble de la région concernée, afin d'identifier son potentiel rôle d'*enhancer* participant à la régulation de gènes cardiaques.

V.1.C. Implication d'IRX5 dans l'activité électrique cardiaque

Le facteur de transcription IRX5 a été montré initialement chez la souris comme central dans la régulation de la repolarisation ventriculaire par le maintien du gradient transmural d'expression du gène *Kcnd2* (Costantini et al., 2005). En effet, dans le cadre d'une repolarisation ventriculaire normale, le potentiel d'action des cardiomyocytes sous-épicaux est plus court que celui des cardiomyocytes sous-endocardiaux. Cette variabilité physiologique des potentiels d'action ventriculaires transmuraux permet d'empêcher ainsi la propagation du potentiel d'action dans le sens inverse (du myocarde sous-épicaux vers le sous-endocardiaux) qui engendrerait des phénomènes arythmiques. La différence de durée entre ces deux types de potentiel d'action est due à la présence du courant potassique I_{to} dans les cardiomyocytes sous-épicaux et à son absence dans les cardiomyocytes sous-endocardiaux (voir partie I.1 Physiologie cardiaque). Chez la souris, le courant I_{to} est généré par les sous-unités *Kcnd2* (pore) et *Kcnp2* (régulatrice). En réprimant le gène *Kcnd2* dans les cardiomyocytes sous-endocardiaux, *Irx5* permet d'assurer le contrôle de l'hétérogénéité physiologique transmurale du potentiel d'action ventriculaire. Chez l'humain, le courant I_{to} est généré par les sous-unités *KCND3* (pore) et *KCNIP2* (régulatrice) et le rôle d'IRX5 dans sa régulation n'avait jusqu'à présent pas été décrit. Nous avons mis en évidence une augmentation d'expression de *KCNIP2* dans le modèle d'iPS-CMs IRX5^{KO} (voir IV.2.C Résultats complémentaires), suggérant qu'IRX5 inhibe son expression dans les cellules contrôles. IRX5 semble donc avoir un rôle dans la régulation de la repolarisation ventriculaire chez l'humain, en inhibant l'expression de *KCNIP2* à l'image de celle de *Kcnd2* chez la souris (Figure 31). Dans le but de valider cette implication d'IRX5 dans la paroi ventriculaire humaine, le modèle de cardioïdes issus d'iPS (Cho et al., 2022; Hofbauer et al., 2021), qui semble adapté à l'étude du gradient d'expression ventriculaire transmural, constituerait un outil adéquat.

L'impact du KO d'IRX5 pourrait ainsi être étudié en tenant compte de la complexité transmurale de la paroi ventriculaire, ce qui n'est pas réalisable avec les modèles iPS en 2D.

Une étude précédemment réalisée dans l'équipe (Al Sayed et al., 2021a, [voir Annexe 1](#)) a permis de mettre en évidence l'implication d'IRX5 dans la régulation de la phase de dépolarisation rapide du potentiel d'action ventriculaire par activation de l'expression des gènes *SCN5A* et *GJA5* (Cx40).

Dans le faisceau de His et les branches droite et gauche du faisceau, la régulation de l'expression de la Cx40 est importante pour assurer une transmission rapide de l'influx électrique dans les ventricules. Cette régulation d'expression est en partie réalisée par *Irx3* chez la souris (Zhang et al., 2011). Des actions coopératives et redondantes entre *Irx3* et *Irx5* ayant déjà été suggérées, et *Irx5* ayant également une expression dans ces compartiments cardiaques, ces données suggèrent l'existence d'une association potentielle entre IRX3 et IRX5 pour la régulation de l'expression de la Cx40 dans le système de conduction ventriculaire.

L'impact d'IRX5 dans la régulation de *SCN5A*, identifié dans l'étude précédente de l'équipe ([voir Annexe 1](#)), a été confirmé et approfondi avec les résultats des deux projets présentés précédemment. En effet, les résultats obtenus dans le projet 1 ont permis d'établir que cette régulation de *SCN5A* pourrait être médiée par une association d'IRX5 avec d'autres facteurs de transcription partenaires : IRX3, GATA4, TBX5 et NKX2-5. De plus, ce résultat est en cohérence avec le rôle potentiel d'IRX5 identifié dans le projet 2 : il orienterait l'action régulatrice du complexe GATA4/TBX5/NKX2-5 dans le cadre de la régulation de l'activité électrique cardiomyocytaire.

L'ensemble de ces résultats suggère donc des rôles d'IRX5 dans la régulation de l'activité électrique, à la fois dans le système de conduction ventriculaire (faisceau de His et branches du faisceau), ainsi que dans le myocarde ventriculaire.

Nos résultats suggèrent également que les mécanismes de régulation de l'activité électrique par IRX5 soient soumis à des redondances fonctionnelles et des compensations par d'autres facteurs de transcription. En effet les analyses transcriptomiques réalisées (1) sur des cardiomyocytes (J30) porteurs de mutations Hamamy dans le gène *IRX5* ([voir Annexe 1](#)) et (2) sur des cardiomyocytes (J30) délétés pour IRX5 ([voir IV.2.C Résultats complémentaires](#)), n'ont pas identifié les mêmes gènes différentiels, ni les mêmes catégories d'annotation fonctionnelle. Le niveau d'expression d'un nombre de gènes plus élevé a été impacté par la

présence des mutations Hamamy (n=745) que par le KO complet d'IRX5 (n=236). Dans le cas des cardiomyocytes IRX5^{Mut}, l'analyse d'annotation des gènes différentiels a de plus permis l'identification de catégories associées à l'activité électrique, contrairement aux cardiomyocytes IRX5^{KO}. Nous pouvons faire l'hypothèse que la présence de la protéine IRX5 mutée pourrait avoir des impacts plus larges que son absence totale (IRX5^{KO}), par des actions dominantes négatives sur l'activité régulatrice de certains de ses partenaires protéiques, tels que IRX3 ou GATA4. Les mutations pourraient ainsi, en plus d'altérer la fonction d'IRX5 seul, perturber l'action des complexes protéiques auxquels IRX5 participe et donc altérer la fonction d'autres facteurs de transcription. Une autre hypothèse serait que la différence de nombre de gènes différentiels serait due à la mise en place de mécanismes de compensation par d'autres facteurs de transcription partenaires dans les cardiomyocytes IRX5^{KO} mais pas dans les cardiomyocytes IRX5^{Mut}.

L'augmentation de l'expression de *KCNIP2* dans les cardiomyocytes IRX5^{KO} à J30 indique que des potentiels mécanismes de compensation n'ont pas été suffisant pour pallier l'absence d'IRX5, suggérant qu'IRX5 est un acteur indispensable de la régulation de ce gène. A l'inverse, la diminution de l'expression de *SCN5A* dans les cardiomyocytes IRX5^{Mut} (voir Annexe 1), qui n'a pas été observée dans les cardiomyocytes IRX5^{KO}, suggèrent qu'IRX5 n'est pas le seul acteur impliqué dans la régulation de *SCN5A*. Ce résultat est cohérent avec les conclusions de l'Article 1 et indique que d'autres partenaires protéiques ont pu compenser l'absence d'IRX5 dans les cardiomyocytes IRX5^{KO} pour maintenir l'expression de *SCN5A*.

L'absence de catégories d'annotation fonctionnelle associées à l'activité électrique dans les résultats d'analyse transcriptomique des cardiomyocytes IRX5^{KO} suggèrent donc que, excepté pour la régulation de l'expression de *KCNIP2*, IRX5 participe à la régulation de l'expression génique associé à l'activité électrique en coopération et redondance fonctionnelle avec d'autres facteurs de transcription. Ainsi, pour la poursuite de l'investigation du rôle d'IRX5 dans la régulation de l'activité électrique cardiaque, il sera nécessaire de tenir compte de ces mécanismes de compensation. Pour cela, la réalisation de KO ou d'inhibition d'expression d'autres facteurs de transcription pourra être entreprise dans les cellules IRX5^{KO}, pour empêcher la mise en place des mécanismes de compensation.

V.1.D. Implication des facteurs de transcription Iroquois dans les mécanismes de fibrose cardiaque

L'étude transcriptomique de la différenciation cardiaque d'iPS IRX5^{KO} nous a permis de mettre en évidence le fait que la délétion d'IRX5 impacte des gènes associés à la matrice extracellulaire, qui a une place centrale dans les mécanismes de fibrose cardiaque, s'établissant dans des contextes de remodelage cardiaque ([voir partie I.1 Physiologie cardiaque](#)).

Le facteur de transcription IRX1 a déjà été associé à des mécanismes de remodelage cardiaque et notamment de fibrose. En effet, Zheng *et al.*, ont montré dans un modèle de constriction aortique chez le rat, qui va déclencher une insuffisance cardiaque et développer une hypertrophie compensatrice et de la fibrose, que le gène *IRX1* était hyperméthylé et avec une expression diminuée par comparaison aux rats contrôles (Zeng et al., 2019). L'hypométhylation induite pharmacologiquement par la 5-Aza-2'-deoxycytidine restaure l'expression d'IRX1 et induit également une diminution de l'insuffisance cardiaque et de la fibrose cardiaque. Ils suggèrent également qu'IRX1 pourrait donc servir de cible thérapeutique contre l'insuffisance cardiaque, notamment pour son effet anti-apoptotique sur les cardiomyocytes, ainsi que pour son effet anti-fibrotique.

De manière intéressante, des expériences réalisées sur la souris *Irx5-KO* et publiées dans une thèse (Kim, 2012, thèse) ont mis en évidence un lien entre *Irx5* et les mécanismes d'hypertrophie et de fibrose cardiaque. A l'état basal, les souris *Irx5-KO* ne semblent pas développer de fibrose ni d'hypertrophie, malgré une fonction contractile réduite ([voir partie I.4.B Les facteurs de transcription Iroquois homeobox](#)). Cependant après constriction aortique (modèle d'insuffisance cardiaque), bien que les souris contrôles développent une hypertrophie compensatrice et de la fibrose, les souris *Irx5-KO* ne développent pas d'hypertrophie et ont une fibrose exacerbée. Ces résultats suggèrent qu'*Irx5* est impliqué dans la régulation de l'hypertrophie compensatrice (effet pro-hypertrophique) et le développement de fibrose (effet anti-fibrotique), suite à un stress mécanique ([Figure 31](#)).

D'après nos résultats ([voir IV.2.C Résultats complémentaires](#)), la délétion d'IRX5 provoque également des variations de la dynamique du facteur de transcription KLF15, qui a une fixation à l'ADN plus importante dans les échantillons contrôles comparés à ceux IRX5^{KO}

(données d'ATAC-seq, Projet 2). La souris Klf15-KO ne présente pas de défauts cardiaques majeurs à l'état basal, mais développe une insuffisance cardiaque plus marquée que les souris contrôles, associée à l'activation de voies transcriptionnelles pro-hypertrophiques, suite à une constriction aortique ou une stimulation pharmacologique à l'angiotensine II (Fisch et al., 2007; Prosdocimo et al., 2015). De plus, d'autres expérimentations réalisées chez le rat ont permis de montrer un rôle de Klf15 dans la régulation de la fibrose cardiaque (Yu et al., 2015). Ces résultats suggèrent ainsi que KLF15 est impliqué, comme IRX5, dans la régulation de l'hypertrophie et de la fibrose cardiaque suite à un stress pathologique. IRX5 et KLF15 pourraient donc avoir des mécanismes conjoints dans ces régulations, qu'il serait intéressant d'approfondir.

Pour étudier les liens entre IRX5 et la fibrose cardiaque, nous avons mis en place le modèle de fibroblastes cardiaques dérivés d'iPS (iPS-CFs). Comme nous n'avons pas observé d'expression protéique d'IRX5 dans les iPS-CFs, suggérant que son rôle potentiel dans la fibrose passerait par son action dans les cardiomyocytes, des expériences de co-cultures entre iPS-CMs et iPS-CFs seront à réaliser ([voir IV.2.D Perspectives pour la poursuite du projet 2](#)). Du fait de l'apparition exacerbée de fibrose uniquement après stimulation chez les souris *Irx5*-KO, les co-cultures entre iPS-CMs *IRX5*^{KO} et iPS-CFs contrôles devront probablement être réalisées sous stimulation (pharmacologique par exemple), afin de pouvoir étudier la présence de facteurs profibrotiques excrétés dans les surnageants de co-cultures. Les modèles de cardioïdes en 3D, intégrant différents types cellulaires (notamment des cardiomyocytes et des fibroblastes), constitueraient également un bon modèle d'étude de la fibrose (Hofbauer et al., 2021), et permettraient l'étude de l'impact d'IRX5 sur les mécanismes de fibrose cardiaque.

V.1.E. Importance d'étudier les mécanismes épigénétiques dans le contexte du développement cardiaque

L'étude transcriptomique que nous avons réalisée tout au long de la différenciation cardiaque des cellules iPS nous a permis d'identifier des transitions d'expression de gènes qui sont associées à des spécifications vers des lignages cellulaires cardiaques. Pour aller plus loin, la dissection complète des mécanismes régulant ces transitions géniques, qui est nécessaire à la compréhension du développement cardiaque sain et pathologique, implique l'étude des

processus épigénétiques. En effet, les changements dynamiques de modification d'histones et de structuration de la chromatine vont avoir un impact sur l'établissement des lignages cellulaires et sont le substrat de pathologies cardiaques (Moore-Morris et al., 2018). De nombreuses mutations de protéines modificatrices d'histones (agissant en tant que *writer* ou *reader*) ont d'ailleurs été impliquées dans des cardiopathies congénitales (Zaidi et al., 2013; Zhang and Liu, 2015). De plus, de nombreux impacts environnementaux (facteurs de stress paternels et maternels) peuvent avoir lieu durant le développement embryonnaire, impliquant des modifications épigénétiques, et peuvent perturber le développement cardiaque (Lamberto et al., 2021), renforçant l'intérêt d'étudier les mécanismes épigénétiques dans le contexte du développement cardiaque.

Il est également important d'étudier les mécanismes épigénétiques dans le cadre de l'étude des facteurs de transcription, car ces derniers peuvent notamment avoir un impact sur l'expression des gènes par le recrutement de partenaires modificateurs d'histones. Par exemple chez la souris, un mécanisme supposé d'inhibition de l'expression du gène *Kcnd2* par *Irx5* passe par le recrutement du cofacteur mBop qui recrute lui-même le modificateur d'histones HDAC, dont la fonction est de retirer les marques d'acétylation pour induire une compaction de la chromatine et une répression de l'expression génique (Costantini et al., 2005).

L'étude du paysage épigénétique est également importante dans l'annotation du génome, toujours dans le but de mieux comprendre les mécanismes sous-jacents à l'action des facteurs de transcription et l'expression de gènes dans le développement cardiaque. En effet, à l'aide du code histone ([voir partie I.3.E.b Marques d'histones](#)) il est possible de prédire l'activité transcriptionnelle, ainsi que la présence d'éléments *cis* régulateurs (comme des *enhancers* ou *super-enhancers*), en analysant la composition en marques épigénétiques de régions génomiques (VanOudenhove et al., 2020).

L'analyse épigénétique que nous avons débutée par l'étude de la marque H3K27Ac, nous permettra de mieux comprendre les mécanismes d'action d'IRX5 et son action directe potentielle sur les marques d'histones, ainsi que sa fixation sur des éléments *cis* régulateurs, car nous avons montré qu'il semblait se fixer préférentiellement à distance des gènes sur l'ADN ([voir IV.2.C Résultats complémentaires](#)).

Enfin, les gènes qui codent des protéines ne représentant qu'environ 2% du génome, l'annotation et l'étude des mécanismes de régulation dans les régions non-codantes sont indispensables, notamment dans la compréhension des régulations géniques impliquées dans le développement cardiaque (Chahal et al., 2019). De plus, de nombreux variants génétiques associés à des cardiopathies congénitales ont été retrouvés dans des régions non-codantes du génome, renforçant l'intérêt de l'étude de ces régions (Richter et al., 2020).

Afin d'étudier le contexte épigénétique impliqué dans les mécanismes d'action d'IRX5 dans le développement cardiaque, nous avons choisi d'utiliser le modèle de différenciation cardiaque de cellules iPS, mais ce modèle est-il vraiment adapté à l'étude des mécanismes épigénétiques ?

Nous avons tout d'abord montré que ce modèle de différenciation cardiaque de cellules iPS était adapté à l'étude du développement cardiaque humain (IV.1.A - Article 1). Il est également adapté à la réalisation de techniques permettant l'étude des mécanismes épigénétiques (ATAC-Seq, ChIP-Seq, ..., voir partie I.4.C.c Techniques d'analyse). Enfin, ce modèle permet aussi d'étudier les impacts potentiels de variants génétiques ou d'éditeurs génomiques d'intérêt sur ces mécanismes, par la reprogrammation de cellules de patients ou l'utilisation de techniques d'édition du génome. Tous ces éléments sont en faveur de l'intérêt du modèle iPS dans la caractérisation des mécanismes épigénétiques, et il a d'ailleurs déjà été utilisé dans cet objectif dans différentes études (Lamberto et al., 2021; Pozo et al., 2022; Wu et al., 2015).

La principale limite potentielle à l'utilisation du modèle iPS dans les études épigénétiques est la persistance d'une mémoire épigénétique de la cellule somatique d'origine. En effet, une reprogrammation incomplète induit une persistance des marques épigénétiques de la cellule spécialisée de départ (Cahan and Daley, 2013). Ce sont principalement des méthylations de l'ADN qui persistent et qui vont être retirées par un processus passif et long (Kim et al., 2010). Bien que cette mémoire épigénétique persistante puisse induire des variations dans les études de différenciation en aval, elle n'a pas obligatoirement un impact sur l'expression des gènes, notamment par manque d'expression de facteurs de transcription spécifiques des cellules spécialisés de départ, dans les cellules pluripotentes (Efrat, 2021). Enfin, plusieurs études ont montré que l'expression du facteur de transcription Nanog était un marqueur d'une reprogrammation complète, sans conservation

d'une mémoire épigénétique (Mattout et al., 2011; Okita et al., 2007). Pour les différents projets présentés précédemment, nous avons testé toutes les lignées iPS utilisées et confirmé la bonne expression de Nanog, validant leur reprogrammation complète avant leur utilisation pour des différenciations cardiaques (au cours desquelles son expression est complètement annulée). Cette validation a par exemple été effectuée dans le cas de la génération des nouvelles lignées iPS IRX5^{KO} et IRX5-HA (IV.2.B - Article 2).

V.1.F. Intérêts et perspectives de l'étude des réseaux de facteurs de transcription dans le développement cardiaque

Les facteurs de transcription ont un rôle central dans le développement cardiaque, en régulant finement l'établissement de programmes d'expression génique spécifiques (Luna-Zurita et al., 2016). De plus, leur étude au cours du développement cardiaque est également importante pour la compréhension des pathologies apparaissant à l'âge adulte (comme l'insuffisance cardiaque par exemple), car ces facteurs de transcription peuvent y être également exprimés et impliqués dans des altérations d'expression génique, ou encore dans le cadre du développement de nouvelles stratégies thérapeutiques (Epstein and Parmacek, 2005).

Différentes études sur les facteurs de transcription cardiaques se sont focalisées sur l'analyse d'un ou de quelques facteurs de transcription en parallèle. Bien que ces analyses soient importantes pour l'identification de fonctions spécifiques des facteurs de transcription, pour obtenir une vision plus proche de la complexité biologique, il est nécessaire d'étudier ces derniers dans un contexte systémique, intégrant l'ensemble des régulations des facteurs de transcription (qui vont également se réguler entre-eux). En effet, dans le processus de développement cardiaque, la progression des cellules dans les lignages cardiaques nécessite l'expression spécifique et chronologique de gènes, et est régulée par l'action combinée et coordonnée de facteurs de transcription (Kathiriya et al., 2015).

Il est donc nécessaire d'étudier les interactions entre les facteurs de transcription cardiaques sous la forme de réseaux intégrés. Différentes études récentes ont utilisé des approches combinées de transcriptomique et de ChIP-Seq pour identifier des interactions entre les facteurs de transcription (Churko et al., 2018; Kathiriya et al., 2021). Dans l'étude de Gonzalez-Teran *et al.*, des données de transcriptomique et de ChIP-Seq ont été combinées à des

données génétiques, associées à des cardiopathies congénitales, et des approches de spectrométrie de masse (Gonzalez-Teran et al., 2022). Cependant, ces études ont également été focalisées sur les réseaux associés à quelques facteurs de transcription uniquement. Dans le cadre du développement cardiaque, qui est un processus dynamique, la notion de temps est nécessairement à intégrer dans l'analyse des réseaux de facteurs de transcription. En tenant compte de ces différents éléments, notre première étude a permis l'identification d'un réseau global de facteurs de transcription au cours de la différenciation cardiaque. Plus de 11 000 relations d'activation et d'inhibition respectivement, ont été identifiées entre les facteurs de transcription, en tenant compte du temps, permettant d'obtenir une vision intégrée et temporelle des régulations entre les facteurs de transcription au cours de la différenciation cardiaque. Ce réseau global de facteurs de transcription pourrait ensuite servir à l'étude des mécanismes de régulation impliqués dans des pathologies cardiaques, et de manière intéressante, 69% des facteurs de transcription du réseau ont déjà été associés à des phénotypes cardiaques physiopathologiques (IV.1.A - Article 1).

Pour aller plus loin, la validation biologique complète du réseau global de facteurs de transcription, et notamment l'identification et l'intégration dans le réseau des mécanismes de coopération et de compensation s'établissant entre eux, permettrait de rendre le réseau dynamique et prédictif. En effet, il permettrait de prédire l'effet de l'altération de facteurs de transcription (par des KO *in silico*) sur le réseau global associé au développement cardiaque, à l'image du travail réalisé par Argelaguet *et al.*, dans la différenciation des progéniteurs neuro-mésodermiques (Argelaguet et al., 2022). Enfin, avec la grande quantité de données publiques déjà disponible, le réseau global actuellement basé sur des données transcriptomiques, pourrait être enrichi de nouveaux niveaux de complexité (génomique, épigénétique, protéomique, métabolomique, ...). L'intégration de l'ensemble de ces données formerait un réseau *in silico* complexe, qui pourrait permettre l'identification de mécanismes pathologiques et de cibles thérapeutiques potentielles, à partir d'informations spécifiques de patients (variants rares ou fréquents, données de transcriptomique, ...), à l'image du projet *Connectivity Map* (basé sur l'expression de gènes, Lamb et al., 2006), mais étendu à l'ensemble des couches d'informations omiques.

V.1.G. Conclusion

Par l'étude transcriptomique réalisée en cinétique au cours de la différenciation cardiaque de cellules iPS humaines, puis par le focus sur les facteurs de transcription, nous avons mis en évidence un réseau global de régulation. Cette étude nous a permis de positionner les facteurs de transcription IRX3 et IRX5 dans le contexte intégré du réseau des facteurs de transcription impliqués dans la différenciation cardiaque. De nouveaux mécanismes de régulation ont ainsi pu être mis en évidence entre IRX3, IRX5 et les facteurs de transcription cardiaque majeurs, GATA4, TBX5 et NKX2-5. L'étude spécifique d'IRX5 a permis de renforcer le lien fonctionnel d'interaction identifié, et d'apporter des informations complémentaires sur les fonctions biologiques qu'il régule dans la différenciation cardiomyocytaire, par différentes approches, notamment multi-omiques. Ainsi, l'ensemble de ces résultats s'insère dans l'objectif plus large, d'une meilleure compréhension du développement cardiaque et des régulations moléculaires sous-jacentes. La poursuite des apports de connaissances robustes dans ce sens, sera indispensable à la compréhension des dérégulations pathologiques et l'identification des acteurs impliqués, dans le but de développer et d'apporter de nouvelles solutions thérapeutiques.

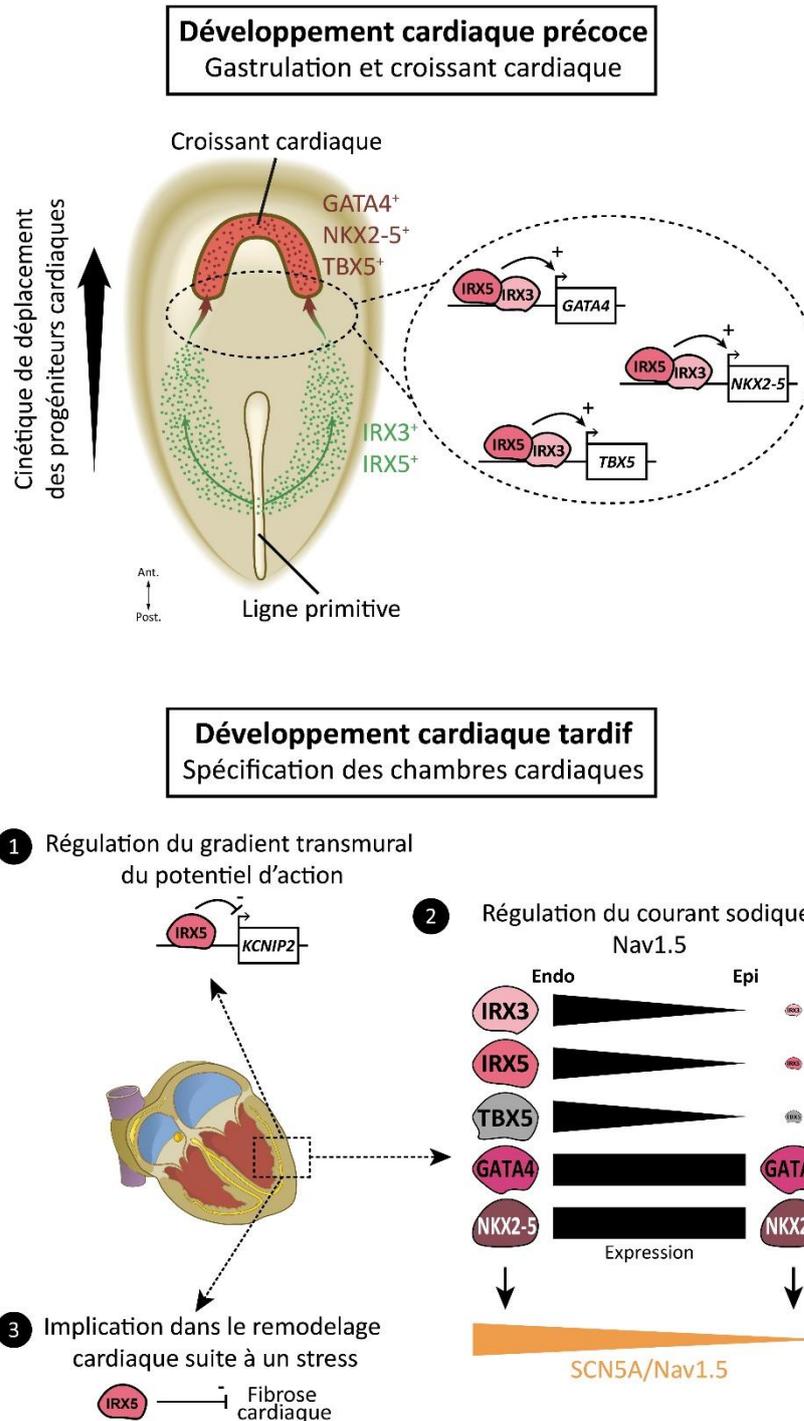


Figure 31 : Proposition de modèle d'implication d'IRX5 dans le développement cardiaque

Ce modèle hypothétique est proposé suite à l'intégration des résultats obtenus au cours de cette thèse et nécessitera approfondissement et validation. Panel du haut : IRX5 serait exprimé dès la gastrulation dans les progéniteurs mésodermiques cardiaques. Ces progéniteurs vont migrer vers la région antérieure de l'embryon pour former le croissant cardiaque (ou premier champ cardiaque), où IRX5 en collaboration avec IRX3, va participer à l'activation de l'expression de GATA4, NKX2-5 et TBX5. Panel du bas : plus tardivement dans le développement cardiaque, une fois les chambres cardiaques formées, IRX5 participerait à l'établissement d'un gradient d'expression dans la paroi ventriculaire. D'une part, il inhiberait l'expression de KCNIP2 dans le myocarde sous-endocardique pour réguler le gradient ventriculaire transmurale du potentiel d'action. D'autre part, il s'associerait avec IRX3, GATA4, NKX2-5 et TBX5 pour réguler l'expression en gradient de SCN5A dans la paroi ventriculaire. Enfin, dans le cadre de la survenue d'un stress pathologique post-développemental (ex : insuffisance cardiaque), IRX5 serait également impliqué dans la régulation du remodelage cardiaque en inhibant la fibrose.

VI

Annexes

Annexe 1

Human model of IRX5 mutations reveals key role for this transcription factor in ventricular conduction

Zeina R Al Sayed, **Robin Canac**, Bastien Cimarosti, Carine Bonnard, Jean-Baptiste Gourraud, Hanan Hamamy, Hulya Kayserili, Aurore Girardeau, Mariam Jouni, Nicolas Jacob, Anne Gaignerie, Caroline Chariou, Laurent David, Virginie Forest, Céline Marionneau, Flavien Charpentier, Gildas Loussouarn, Guillaume Lamirault, Bruno Reversade, Kazem Zibara, Patricia Lemarchand, Nathalie Gaborit

DOI : 10.1093/cvr/cvaa259 ; PMID : 32898233

Publié dans *Cardiovascular Research*, 2021

Human model of *IRX5* mutations reveals key role for this transcription factor in ventricular conduction

Zeina R. Al Sayed¹, Robin Canac¹, Bastien Cimarosti¹, Carine Bonnard², Jean-Baptiste Gourraud^{1,3}, Hanan Hamamy⁴, Hulya Kayserili⁵, Aurore Girardeau¹, Mariam Jouni¹, Nicolas Jacob¹, Anne Gaignerie⁶, Caroline Chariou⁶, Laurent David^{6,7,8}, Virginie Forest¹, Céline Marionneau¹, Flavien Charpentier^{1,3}, Gildas Loussouarn¹, Guillaume Lamirault^{1,3}, Bruno Reversade^{1,2,5,9,10,11}, Kazem Zibara¹², Patricia Lemarchand^{1,3*}, and Nathalie Gaborit^{1*}

¹Université de Nantes, CNRS, INSERM, l'institut du thorax, 8 quai Moncoussu, F-44000 Nantes, France; ²Institute of Medical Biology, A*STAR, 8A Biomedical Grove, Singapore 138648, Singapore; ³Université de Nantes, CHU Nantes, CNRS, INSERM, l'institut du thorax, 8 quai Moncoussu, F-44000 Nantes, France; ⁴Department of Genetic Medicine and Development, Geneva University, 1 rue Michel-Servet, Geneva 1211, Switzerland; ⁵Medical Genetics Department, Koç University School of Medicine(KUSOM), Rumelifeneri Yolu 34450, Istanbul, Turkey; ⁶Université de Nantes, CHU Nantes, Inserm, CNRS, SFR Santé, Inserm UMS 016, CNRS UMS 3556, 8 Quai Moncoussu, F-44000 Nantes, France; ⁷Université de Nantes, INSERM, CRTI, 30 Bd Jean Monnet, F-44093 Nantes, France; ⁸ITUN, CHU Nantes, 30 Bd Jean Monnet, F-44093 Nantes, France; ⁹Department of Paediatrics, National University of Singapore, 1E Kent Ridge Road, Singapore 119228, Singapore; ¹⁰Institute of Molecular and Cellular Biology, A*STAR, 61 Biopolis Drive, Singapore 138673, Singapore; ¹¹Reproductive Biology Laboratory, Amsterdam UMC, Meibergdreef 9 1105, Amsterdam-Zuidoost, Netherlands; and ¹²ER045, Laboratory of stem cells, DSST, Biology department, Faculty of Sciences, Lebanese University, Rafic Hariri Campus - Hadath, Beirut 1700, Lebanon;

Received 7 October 2019; revised 15 January 2020; editorial decision 24 August 2020; accepted 28 August 2020; online publish-ahead-of-print 8 September 2020

Time for primary review: 22 days

Aims

Several inherited arrhythmic diseases have been linked to single gene mutations in cardiac ion channels and interacting proteins. However, the mechanisms underlying most arrhythmias, are thought to involve altered regulation of the expression of multiple effectors. In this study, we aimed to examine the role of a transcription factor (TF) belonging to the Iroquois homeobox family, *IRX5*, in cardiac electrical function.

Methods and results

Using human cardiac tissues, transcriptomic correlative analyses between *IRX5* and genes involved in cardiac electrical activity showed that in human ventricular compartment, *IRX5* expression strongly correlated to the expression of major actors of cardiac conduction, including the sodium channel, Nav1.5, and Connexin 40 (Cx40). We then generated human-induced pluripotent stem cells (hiPSCs) derived from two Hamamy syndrome-affected patients carrying distinct homozygous loss-of-function mutations in *IRX5* gene. Cardiomyocytes derived from these hiPSCs showed impaired cardiac gene expression programme, including misregulation in the control of Nav1.5 and Cx40 expression. In accordance with the prolonged QRS interval observed in Hamamy syndrome patients, a slower ventricular action potential depolarization due to sodium current reduction was observed on electrophysiological analyses performed on patient-derived cardiomyocytes, confirming the functional role of *IRX5* in electrical conduction. Finally, a cardiac TF complex was newly identified, composed by *IRX5* and *GATA4*, in which *IRX5* potentiated *GATA4*-induction of *SCN5A* expression.

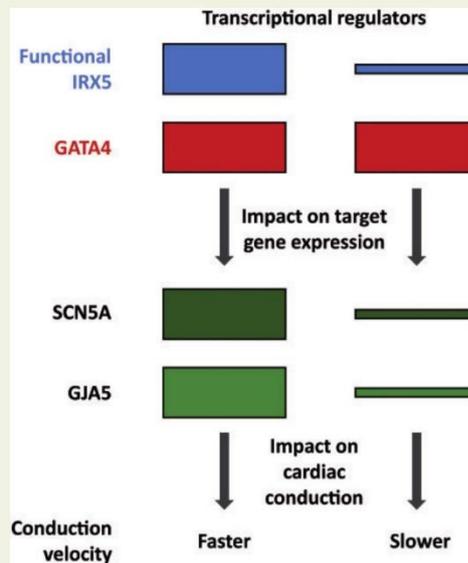
Conclusion

Altogether, this work unveils a key role for *IRX5* in the regulation of human ventricular depolarization and cardiac electrical conduction, providing therefore new insights into our understanding of cardiac diseases.

*Corresponding authors. E-mail: nathalie.gaborit@univ-nantes.fr (N.G.); patricia.lemarchand@univ-nantes.fr (P.L.)

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author(s) 2020. For permissions, please email: journals.permissions@oup.com.

Graphical Abstract



Keywords

IRX5 • Transcription factors • Conduction • Arrhythmia • Hamamy syndrome • Human-induced pluripotent stem cells

1. Introduction

Inherited cardiac arrhythmias account for ~20% of sudden cardiac deaths. Only a small portion of these arrhythmias are monogenic familial diseases, having been successfully linked to rare mutations in cardiac ion channels and related proteins.¹ However, pathogeny of inherited cardiac arrhythmias is increasingly thought to be based on complex mechanisms involving polygenic inheritance and/or altered regulation of the expression of multiple effectors.² In this context, investigating how dysfunction of transcriptional regulators participates in the mechanism of these diseases is of major importance.

The regulation of cardiac electrical transcriptional programme, including ion channels and gap junctions, is a major determinant of proper initiation and propagation of action potential (AP) through the cardiac muscle. In animal models, several transcription factors (TFs) are implicated in cardiac electrical activity and arrhythmias,³ including members of the Iroquois (IRX) family of TFs.⁴

Autosomal recessive mutations in *IRX5*, an IRX homeobox TF, is linked to Hamamy syndrome (HMMS, OMIM611174), an inherited congenital disorder. The patients' cardiac clinical phenotype combines bradycardia and QRS complex prolongation, reflecting a delayed ventricular electrical conduction.^{5,6} Surprisingly, in mice, deletion of *Irx5* leads to a different phenotype, a ventricular repolarization defect, due to the role of *Irx5* in controlling potassium channel gene expression.⁷ Inversely, another murine member of the *Irx* family, *Irx3*, is essential for fine regulation of intercellular coupling within the ventricular conduction system, through the regulation of connexins expression.^{8,9} This suggests that

distinct regulatory pathways govern the differences between human and animal cardiac electrical physiology.

In this study, our goal was to decipher whether IRX5 controls human cardiac electrical activity. First, correlative transcriptomic analyses between IRX5 and cardiac electrophysiology genes in human cardiac tissues, suggested that IRX5 may regulate several major actors of cardiac conduction. Then, human-induced pluripotent stem cells (hiPSCs) derived from HMMS patients carrying *IRX5* mutations were used to elucidate the role of IRX5 in regulating expression of these actors of cardiac conduction and to uncover a cooperative activity between TFs. This work reveals how loss-of-function (LOF) mutation in one key TF affects the expression of multiple effectors of ventricular electrical conduction and improves our understanding of cardiac disease, paving the way therefore for more effective therapies.

2. Methods

Ethical statement

The study was conducted according to the principles set forth under the Declaration of Helsinki (1989) and European guidelines for clinical and genetic research. Institutional review board approvals of the study were obtained before the initiation of patient enrolment. Regarding the patient-derived biological samples, signed informed consent allowing the experiments to be conducted have been received from all individuals. Any related health information was collected in compliance with applicable law/regulation and with any applicable policy of the ethics committee

with jurisdiction over the biological sample collection. All biological samples and their related health information have been provided in coded form such that subjects cannot be identified directly. The provisions of French law, article L1110-4 of the Code de la santé publique, related to the privacy and confidentiality of information regarding patients, have been observed. The transfer of the HMMS-affected patient's biological samples has complied with all applicable laws as well as legal, regulatory, and ethical requirements in transferring and obtaining such Material and the corresponding informed consents (Authorization IE 2009-505(1)).

2.1 Transcriptional analysis of human and murine tissues

RNA sequencing data in human heart were collected from the GTEx portal website (<http://gtexportal.org>). Briefly, transcripts per million were downloaded for 297 atrial and 303 ventricular available tissue samples.

Regarding murine samples, all animal experiments obtained local approval and all animal procedures follow the guidelines from Directive 2010/63/EU of the European Parliament on the protection of animals used for scientific purposes. Mice were euthanized in their home cage using CO₂ inhalation. Hearts were excised from six 10-week-old C57BL/6 male mice and atria and ventricles were carefully dissected as previously described.¹⁰

Two-way hierarchical agglomerative clustering was applied to gene-expression matrices (see Section 2.5.2 below), by centring gene expression values on the values-median and using average linkage-clustering with uncentred correlation (Cluster 3.0). GTEx left ventricular tissues RNA sequencing data of cardiac sodium channels and connexins were ranked according to ascending expression of *IRX5*. Genes encoding sodium channel subunits and connexins that presented a median GTEx expression level in left ventricular samples >1 transcripts per million were selected. Expression correlations were tested using Pearson correlation package on R software.

2.2 hiPSC generation

Primary dermal fibroblasts were derived from skin biopsies obtained from donors of two different consanguineous families, previously described with HMMS.^{5,6} The first diseased donor is a 4-year-old boy, homozygous carrier of the c.448G>C *IRX5* mutation (p.Ala150Pro annotated as A150P) in whom a QRS prolongation on the ECG was diagnosed. The second donor is a 17-year-old boy belonging to another family harbouring the homozygous c.498C>A *IRX5* mutation (p.Asn166Lys annotated as N166K) who presents a prolonged QRS duration on the ECG, in addition to a tiny patent ductus arteriosus. In parallel, fibroblasts from three individuals lacking the *IRX5* mutations were used as controls: hiPSCs were generated from an intra-familial control of the patient carrying the A150P-*IRX5* mutation, his unaffected sister. hiPSCs were also generated from an extra-familial control of the patient carrying the N166K-*IRX5* mutation, who has been chosen because of its match in terms of age, sex, and ethnicity to the patient. A third control hiPSC line, which was not related to both HMMS patients, and that had previously been described, has been also used.^{11–13} hiPSC lines were generated using Sendai virus method. For each line, two to three hiPSC clones were selected and expanded on stem cell-qualified Matrigel-coated plates (0.1 mg/mL; BD Bioscience) with StemMACS™ iPS Brew XF medium (Miltenyi Biotec).

2.3 hiPSC characterization

For all clones, genomic DNA was extracted and mutations were verified by sequencing using primers specific for the genomic regions surrounding the A150P and N166K mutations (Supplementary material online, Table S1). Pluripotency marker gene expression was assessed using flow cytometry, qRT-PCR, and immunofluorescence as previously described.¹¹ Single nucleotide polymorphism (SNP) analysis of hiPSCs compared to their parental skin fibroblast cells was used to confirm genome integrity after reprogramming. DNA was extracted from somatic and hiPSC samples using the QIAGEN QiaAmp kit, according to the manufacturer's recommendations. The gDNA was quantified using a nanodrop. About 200 ng of gDNA was outsourced to Integragen Company (Evry, France) for karyotype analysis using HumanCore-24-v1 SNP arrays. This array contains over 300 000 probes distributed throughout the genome with a median coverage of one probe every 5700 bases. All genomic positions were based on Human Genome Build 37 (hg19). Analysis was performed with GenomeStudio software. Chromosome abnormalities were determined by visual inspection of logR ratios and B-allele frequencies (BAF) values and comparing parental cells and hiPSC-derived samples. LogR ratio, the ratio between observed and expected probe intensity, is informative of copy number variation (i.e. deletions/duplications), whereas BAF is informative of heterozygosity. SNP data were used to compute copy number variation. In particular, this type of chips allows detecting loss of heterozygosity (LOH), an important concern for hiPSC, which is not detectable with classical CGH arrays.

2.4 Differentiation of hiPSCs into cardiomyocytes

At 80% of cell confluency, hiPSC lines were passaged using Gentle Cell Dissociation Reagent (STEMCELL Technologies) and Y-27632 Rho-kinase inhibitor (STEMCELL Technologies). The matrix sandwich method was used to generate cardiomyocytes (CMs), as previously described.¹² From Day 5 of differentiation, cell medium was changed every other day with RPMI-1640 medium supplemented with B27 (with insulin; Thermo Fisher Scientific). After 20–28 days following the beginning of differentiation, spontaneously beating cell clusters were mechanically isolated using a needle and then used for electrophysiological, biochemical, and molecular analyses. All experiments have been performed on at least three independent differentiations for each clone, performed on different hiPSC cell passages. Data obtained from the various clones of the three control lines have been pooled and averaged and correspond to the *n* values provided in the figure legends. For each Hamamy patient, data obtained from the different clones have been pooled and averaged.

2.5 RNA expression

2.5.1 RNA extraction

Samples of total RNA were obtained using the NucleoSpin RNA kit (MACHEREY-NAGEL) following the manufacturer's specifications.

2.5.2 TaqMan low-density array

TaqMan low-density array (TLDA) studies were conducted using six ventricular and six atrial control murine tissues as well as beating clusters of hiPSC-CMs obtained from 12 control and 9 *IRX5*-mutant CM differentiations. About 1 µg of RNA was reverse transcribed into cDNA using SuperScript IV Vilo Master Mix (Thermo Fisher Scientific). For murine tissues, three specifically designed 384-well TLDA were used: one containing 4 × 96 primers and probes for cardiac ion channels¹⁰ and two

others, each containing a set of 4 × 96 primers and probes for cardiac TFs (Supplementary material online, Table S2A). For hiPS-CMs, TLDA probe selection covered gene families implicated in cardiac ion channel expression and regulation, and CM structure (Supplementary material online, Table S2B). Genes with average Ct > 32 in all compared groups were considered undetectable and excluded from the analysis (SCN10A and ABCC8). Average Ct of remaining genes for each sample was used for data normalization.^{14–16}

2.5.3 3'-sequencing RNA profiling

3'-Sequencing RNA profiling (3'-SRP) protocol was performed according to Kilens *et al.*¹⁷ Briefly, the libraries were prepared from 10 ng of total RNA. About 40 RNA samples were extracted from 8 IRX5-mutated and 13 control hiPSC samples (a duplicate for each clone obtained at different cell passages) and 19 samples of corresponding hiPS-CMs. The mRNA poly(A) tail was tagged with universal adapters, well-specific barcodes, and unique molecular identifiers during template-switching reverse transcriptase. Barcoded cDNAs from multiple samples were then pooled, amplified, and tagged using a transposon-fragmentation approach which enriches for 3' ends of cDNA. A library of 350–800 bp was run on an Illumina HiSeq 2500 using a HiSeq Rapid SBS Kit v2 (50 cycles; FC-402-4022) and a HiSeq Rapid PE Cluster Kit v2 (PE-402-4002). Read pairs used for analysis matched the following criteria: all 16 bases of the first read had quality scores of at least 10 and the first six bases correspond exactly to a designed well-specific barcode. The second reads were aligned to RefSeq human mRNA sequences (hg19) using bwa version 0.7.4.4. Reads mapping to several positions into the genome were filtered out from the analysis. Digital gene expression profiles were generated by counting the number of unique molecular identifiers associated with each RefSeq genes, for each sample. R package DESeq2 was used to normalize gene expression. Differentially expressed genes were finally clustered using cluster 3.0 software. Enriched biological pathways were assigned for differentially expressed genes in IRX5-mutated vs. control hiPS-CMs. For that, Reactome pathways were attributed to the most variant genes with *P*-value < 0.05.^{18,19} However, genes encoding for ion channels exhibited low expression levels and therefore they were not taken into consideration for differentially expressed genes analysis. To overcome this issue, *P*-values for these genes were calculated based on TLDA analysis. Babelomics was then used to calculate enrichment of Reactome pathways for genes differentially expressed in both TLDA and 3'-SRP analyses. Finally, enrichment scores (ES) were calculated.^{20,21}

2.6 Protein expression and interaction

2.6.1 Protein extraction

IRX5, Nav1.5, Connexin 40 (Cx40), and GATA4 expression were evaluated in human left and right ventricular tissues obtained from three non-diseased individuals as previously described.²² Corresponding subendocardial (also containing Purkinje cells) and subepicardial tissue slices were obtained at the cardiac base and snap-frozen in liquid nitrogen. Protein expressions and interactions were also investigated in beating hiPS-CMs at Day 28 of differentiation. For protein extraction, cells were homogenized on ice using lysis buffer containing in mM: 1% Triton X-100; 100 NaCl; 50 Tris-HCl; 1 EGTA; 1 Na₂VO₄; 50 NaF; 1 phenylmethylsulfonyl fluoride; and protease inhibitors cocktail (P8340, Sigma-Aldrich). Extracted samples were sonicated and centrifuged at 15 000 g for 15 min at 4°C. Protein quantification was carried out using PierceTM BCA Protein Assay Kit (Thermo Fisher). Immunoblot was finally done

using the corresponding primary antibodies (Supplementary material online, Table S3).

2.6.2 Subcellular protein fractionation

To unveil the impact of IRX5 mutations on its cellular localization, nuclear, and cytosolic proteins were separated. For that, cytoplasmic membranes were disrupted on ice in a hypotonic solution including in mM: 500 HEPES, 500 MgCl₂, 500 KCl, 0.1 DTT, and protease inhibitor cocktail. Cells nuclei were pelleted by 400 × g centrifugation and the supernatant contained the cytosolic fraction. In order to extract nuclear proteins, a protein extraction buffer was added to pellets followed by sonication.

2.6.3 Co-immunoprecipitation

Co-immunoprecipitation was performed to investigate protein-to-protein interactions between IRX5 and GATA4 in hiPS-CMs. A total of 12.5 μL magnetic Dynabeads[®] Protein G (Thermo Fisher Scientific) were added to 3 μg anti-IRX5 antibody (VH0010265M1, Sigma-Aldrich) or anti-mouse IgG antibody as control (02-6502, Thermo Fisher Scientific) and conjugated by incubation for 40 min at room temperature with rotation. The complex beads-antibody was cross-linked using 27 mg DMP (dimethyl pimelimidate, Thermo Fisher Scientific), and then incubated for 2 h with 800 μg of total lysate at 4°C. The supernatant was then discarded and the beads were washed with the lysis buffer. Proteins were eluted in 30 μL of a mixture of NuPAGE[®] Sample Reducing Agent (10×) and LDS Sample Buffer (4×) (Invitrogen) and heated at 60°C. The samples were then loaded onto a 4–15% precast polyacrylamide gel (Bio-Rad).

2.7 Immunofluorescence

hiPSCs or hiPS-CMs at Day 20 of differentiation were dissociated using collagenase II (Gibco) and seeded onto 8-wells iBidi plates (Biovalley) coated with Matrigel (Corning). Twelve days later, cultured cells were immunostained using a staining buffer set according to the manufacturer's instructions (00-5523-00, Thermo Fisher Scientific). HiPS-CMs were labelled by cardiac differentiation markers or TF. Primary antibodies are listed in Supplementary material online, Table S3. DNA was highlighted using DAPI and pictures were taken using confocal microscopy (Nikon; MicroPiCell core facility).

2.8 Chromatin immunoprecipitation

2.8.1 IRX5 binding sites

Promoter DNA sequences were found using Ensembl. Consensus sequences known to mediate IRX TF interaction with DNA^{23–25} were screened on gene promoters. Each site was verified for its conservation among species using rVista 2.0. Primers surrounding the identified conserved potential binding sites (BS) were designed using Primer 3. Corresponding primer sequences are detailed in Supplementary material online, Table S4.

2.8.2 Chromatin immunoprecipitation

At Day 28 of differentiation, beating CM clusters were isolated and crosslinked. After chromatin sonication, IRX5 antibody (VH0010265M1, Sigma-Aldrich) was added followed by the addition of protein G-coated dynabeads. Following overnight incubation, the chromatin was then eluted and reverse-crosslinked. SYBER green qRT-PCR method was finally performed using the designed primers

(Supplementary material online, Table S4) to test for enrichment of the targeted sequences. Results are shown as percentage of input.

To confirm the specificity of IRX5 antibody, HEK293 cells were transfected using plasmids encoding human sequences of WT IRX5, N166K IRX5, and WT IRX3, as well as an empty pcDNA3 plasmid. Testing the potential non-specificity of IRX5 antibody towards IRX3 was performed as IRX3 and IRX5 have been shown to share a high level of transcript and protein homology,²⁶ strong similarities in their expression pattern²⁷ and to have coordinated functional roles.²⁸ (i) Confirming the IRX3 protein expression in all cell samples transfected by IRX3, a band at the expected size for IRX3 (<100 kDa, as IRX3 was Myc-tagged) was indeed obtained (Supplementary material online, Figure S1). (ii) When HEK293 cells were transfected by an empty plasmid or IRX3 plasmid, as expected, on corresponding western blots with IRX5 antibody, no band was detected, demonstrating that IRX5 antibody did not bind other protein than IRX5, including IRX3. (iii) Cells were transfected by IRX5 (left panel) or IRX5-N166K (right panel) plasmids and a unique band of the expected size for IRX5 (~100 kDa, as IRX5 was GFP-tagged) was observed, further demonstrating that the antibody specifically bound to the WT or mutated forms of IRX5 protein (Supplementary material online, Figure S1).

2.9 Cellular electrophysiology

2.9.1 Cell dissociation

hiPS-CMS were enzymatically dissociated into single cells using collagenase II (Gibco) at 37°C for 20 min. Cells were then incubated at RT for 30 min in Kraft–Bruhe solution containing (in mM): 85 KCl; 30 K₂HPO₄; 5 MgSO₄; 20 glucose; 20 taurine; 5 Na pyruvate; 5 creatine; 2 K₂-ATP; 1 EGTA; pH 7.2 adjusted with KOH. Single cells were then plated at low density on Matrigel-coated Petri-dishes (Nunc) and recorded 11–14 days after dissociation.

2.9.2 Data recording and analysis

Stimulation and data recording were performed at 37°C using an Axopatch 200B amplifier controlled by Axon pClamp 10.6 software through an A/D converter (Digidata 1440A) (all Molecular Devices). Data were analyzed using Clampfit 10.6 software. For all cell lines, data were collected from at least three independent differentiations. Experimental details regarding solutions composition are described in Supplementary material online, Table S5.

2.9.3 Current clamp

Action potential recordings were obtained using the amphotericin-B-perforated patch-clamp technique. Patch pipettes were pulled from borosilicate glass capillaries (Sutter instruments) to obtain a tip resistance of 2–3 MΩ. APs were recorded in cells that were first allowed to beat spontaneously and then paced. A stimulus of an amplitude of 30–35 pA/pF during 1 ms (in order to avoid interference with the upstroke phase of the AP) was used at various frequencies, in absence of I_{K1} injection (see below). Cells were classified as either ventricular-like, atrial-like, or nodal-like, based on several parameters previously described (maximum upstroke velocity, dV/dt_{max}; AP duration, APD; resting membrane potential, RMP).¹¹ Only ventricular-like cells were selected for further analysis. I_{K1} current was not observed in hiPS-CMs (not shown), consistent with the observation of phase-4 depolarization in these cells. Without I_{K1}, the APs RMP was about -60 mV (Figure 5B), which corresponds to immature ventricular CMs. Thus electronic expression of the inward rectifier potassium current, I_{K1}, was realized using dynamic patch-clamp.^{29,30} With

2 pA/pF of *in silico* I_{K1} injection, hiPS-CM membrane resting potential was about -80 mV, closer to physiological RMP of adult ventricular CMs. Averaged parameters were obtained from seven consecutive APs. Both cell stimulation and injection of I_{K1} were realized by a custom-made software running on RT-Linux and a National Instrument A/D converter (NI PCI-6221) connected to the current command of the amplifier.

2.9.4 Voltage clamp

I_{Na} and I_{Ca,L} recordings were performed in the ruptured whole-cell configuration using voltage clamp protocol as depicted in the figures. Pipettes (Sutter instruments) with 1.7–2.5 MΩ tip resistance were waxed before use to decrease pipette capacitive currents. I_{Ca,L} and I_{Na} recordings were low-pass filtered at, respectively, 3 and 10 kHz. Holding potentials were, respectively, -80 and -100 mV. The capacitance and series were compensated using a VE-2 amplifier (Alembic Instrument, Qc, Canada). Leak current was subtracted. Current densities were calculated by dividing the current to the membrane capacitance (Cm). Steady-state activation and inactivation curves were fitted using a Boltzmann equation: $I/I_{max} = A / \{1.0 + \exp [(V_{1/2} - V)]/k\}$ with $k > 0$ for the activation curve and $k < 0$ for the inactivation curve.

2.9.5 Action potential clamp

In order to record specific currents during AP cycles, and especially during the depolarization phase, cells membrane potential was clamped according to voltage-AP-like protocols. Representative APs acquired from spontaneously beating control hiPS-CMs with (dV/dt_{max} = 100 V/s; APD₉₀ = 280 ms; RMP = -83 mV) or without (dV/dt_{max} = 12 V/s; APD₉₀ = 300 ms; RMP = -57 mV) *in silico* I_{K1} injection were selected. Activated sodium current during AP was measured by subtracting traces before and after adding 30 μM tetrodotoxin TTX in I_{Na}-specific intracellular and extracellular (with 130 mM NaCl concentration) solutions (Supplementary material online, Table S5). I_{Ca,L} was considered as the Nifedipine-sensitive (4 μM) calcium current recorded using I_{Ca,L} solutions (Supplementary material online, Table S5).

2.10 Luciferase assay

SCN5A BS2- (from -261 to +613 bp of the TSS) and BS2+ (from -261 to +1072 bp from TSS) Plasmids (150 ng), pGL2-Renilla (15 ng) with or without GATA4 expressing vector (150 ng) and with or without either of the wild-type or mutated forms of IRX5 expressing vectors (150 ng) were transfected into HEK293 cells using Lipofectamine 2000 (Thermo Fisher Scientific). DNA quantities were equalized in each condition using empty pcDNA3.1. Cells were harvested 48 h later and luciferase activity was measured using Dual Luciferase reporter assay system (Promega).

2.11 Microscale thermophoresis

Microscale thermophoresis (MST) was used to verify protein interaction with DNA according to Nanotemper instructions. Briefly, HEK293 cells transfected by IRX5-GFP expressing vector were used to produce fluorescently labelled IRX5 protein. Cell lysates were prepared using a buffer containing 20 mM Tris-HCl, 130 mM NaCl, 1% NP40, and protease inhibitors cocktail. Titration series were performed to obtain varying concentrations of SCN5A plasmids. Measurements were accomplished in standard treated capillaries on Monolith NT.115 instrument using 80% of laser power.

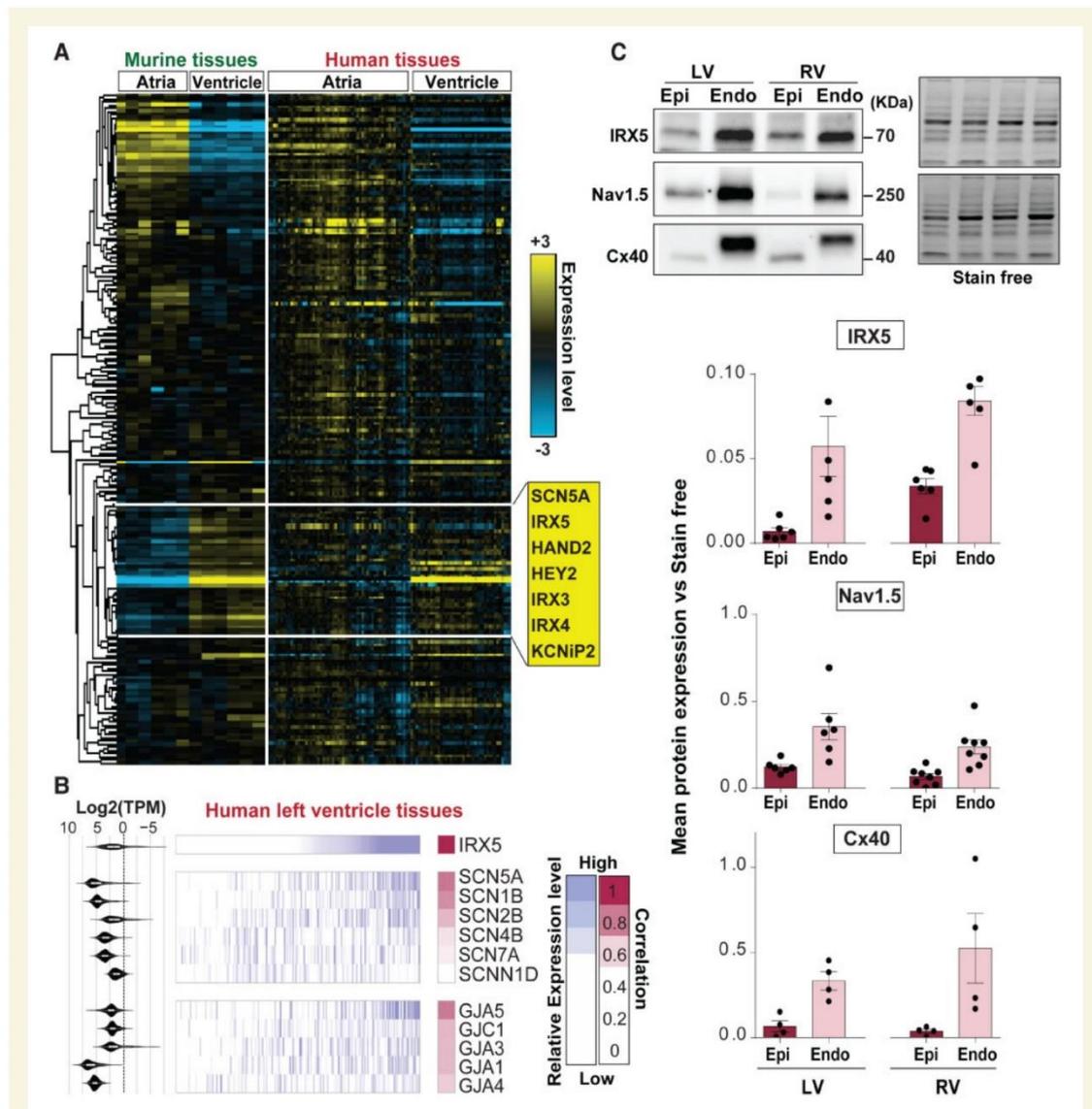


Figure 1 IRX5 TF and conduction genes expression in human cardiac tissues. (A) Heatmap displaying expression level of 212 cardiac genes (ion channels and TFs) in murine atrial and ventricular tissues measured by TLDA ($-\Delta\Delta\text{Ct}$), and in human atrial and left ventricular samples (GTEx). Murine genes were clustered using hierarchical ascending method with an uncentred correlation metric and average linkage. Heatmap of the same genes was obtained in human tissues based on the murine cluster, in a supervised hierarchical manner. Yellow and blue indicate high and low expression levels, respectively. Representative genes of the ventricular-specific gene cluster are listed in the yellow box. (B) Comparative analysis of the expression pattern of genes encoding sodium channels and gap junction proteins in human left ventricular tissues, to the IRX5 gene expression pattern (GTEx). Samples were ordered according to their IRX5 expression levels (white/purple gradient from the lowest to the highest). Keeping the same sample organization, a white/purple colour gradient was then similarly applied to selected genes involved in ventricular conduction of electrical influx (the median GTEx expression level in left ventricular samples is indicated on the left for each gene). Pearson correlation values between each sodium channel or gap junction gene and IRX5 expression profiles are shown using white/pink colour gradients. (C) Representative immunoblots of IRX5, Na_v1.5, and Cx40 levels in human protein lysates of ventricular subendocardial (Endo) and subepicardial (Epi) samples obtained from three human left (LV) and right ventricles (RV). Corresponding averaged IRX5, Na_v1.5, and Cx40 protein expression levels obtained from tissue samples of three individuals, technically replicated one to three times and normalized to Stain free.

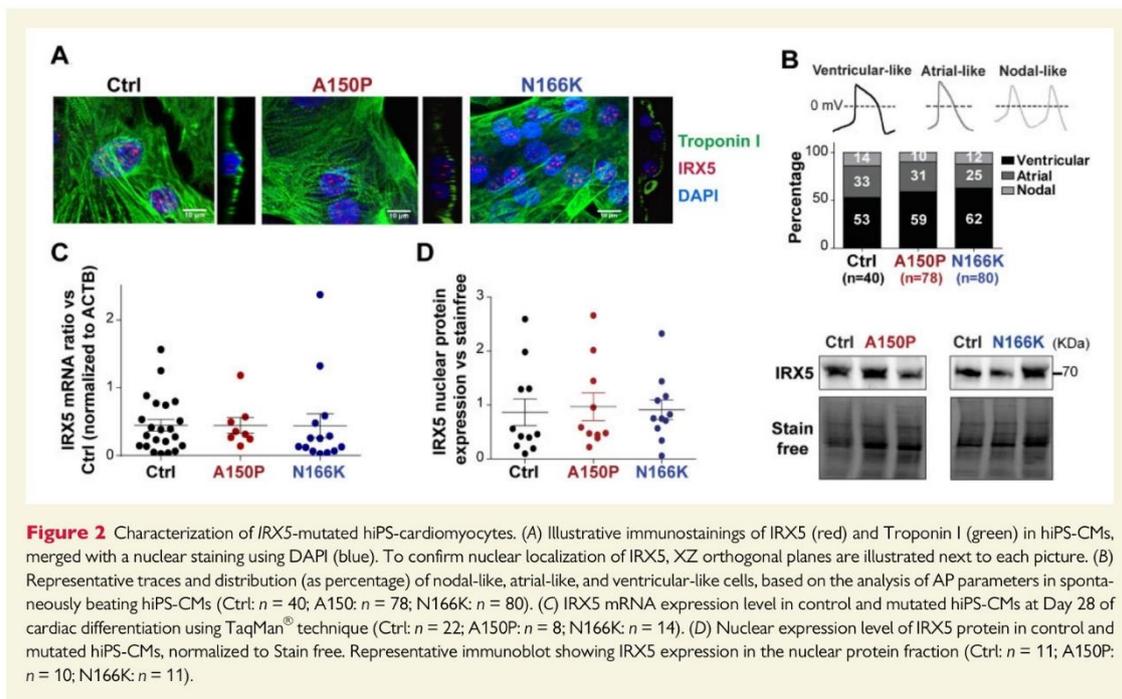


Figure 2 Characterization of *IRX5*-mutated hiPS-cardiomyocytes. (A) Illustrative immunostainings of *IRX5* (red) and Troponin I (green) in hiPS-CMs, merged with a nuclear staining using DAPI (blue). To confirm nuclear localization of *IRX5*, XZ orthogonal planes are illustrated next to each picture. (B) Representative traces and distribution (as percentage) of nodal-like, atrial-like, and ventricular-like cells, based on the analysis of AP parameters in spontaneously beating hiPS-CMs (Ctrl: $n = 40$; A150: $n = 78$; N166K: $n = 80$). (C) *IRX5* mRNA expression level in control and mutated hiPS-CMs at Day 28 of cardiac differentiation using TaqMan[®] technique (Ctrl: $n = 22$; A150P: $n = 8$; N166K: $n = 14$). (D) Nuclear expression level of *IRX5* protein in control and mutated hiPS-CMs, normalized to Stain free. Representative immunoblot showing *IRX5* expression in the nuclear protein fraction (Ctrl: $n = 11$; A150P: $n = 10$; N166K: $n = 11$).

2.12 Statistical analysis

Results are expressed as mean \pm S.E.M. Comparisons were made using Student's *t*-test, or two-way ANOVA with Bonferroni *post hoc* test for repeated measures. Values of $P < 0.05$ were considered statistically significant. Statistical analyses were performed with GraphPad Prism software.

3. Results

3.1 *IRX5* TF expression correlates with conduction genes expression in human cardiac tissues

Transcriptomic analyses were performed to obtain a systemic view of *IRX5*-driven gene regulations and to direct further functional analyses. In order to identify evolutionary-conserved *IRX5*-coregulated genes in contractile tissues, expression profiles of cardiac genes including ion channels and TFs were compared between murine and human tissues. Overall, the global gene expression contrast profile between atrial and ventricular murine tissues was conserved in human samples (Figure 1A). The cluster of genes that are more expressed in ventricular tissues, included *IRX5* and several genes involved in cardiac conduction, such as *SCN5A*, which encodes the main cardiac sodium channel $Na_v1.5$ (Figure 1A). Then, the correlations between *IRX5* gene expression level and those of sodium channels and connexins (Cx), the two main families of genes implicated in cardiac conduction in humans, were investigated in the ventricles. Interestingly, *SCN5A* and *GJA5*, encoding one of the main cardiac Cx, Cx40, showed the strongest correlation coefficient with *IRX5*, respectively, 0.62 and 0.60 (Figure 1B). In accordance with

these results and confirming previous reports,^{22,31–34} both $Na_v1.5$ and Cx40 shared a gradient of protein expression similar to that of *IRX5* throughout the human ventricular wall with a higher level in subendocardium, which also includes Purkinje fibres, as compared to subepicardium (Figure 1C).

Taken together, these analyses shed light on the importance of studying the role of *IRX5* in the regulation of cardiac ventricular genes, and more specifically in the regulation of players of cardiac electrical conduction in human ventricles. The analysis of the role of *IRX5* in humans was then pursued using hiPSCs obtained from three non-diseased individuals (shown as averaged) as well as from two patients carrying different LOF mutations in *IRX5* gene associated to a prolonged QRS interval in comparison with normal limits in children on the ECG,³⁵ together with a right bundle branch block^{6,36} (Supplementary material online, Figure S2A).

3.2 *IRX5*-mutated hiPS-CMs derived from HMMs patients are a relevant model to study *IRX5* role

Dermal fibroblasts of two Hamamy patients who are homozygous carriers of either the A150P- or the N166K-*IRX5* mutation⁶ and of three intra- and extra-familial control subjects were reprogrammed into hiPSCs. Mutations were confirmed by sequencing in hiPSCs (Supplementary material online, Figure S2B). Genome integrity was validated using SNP analysis (data not shown). Expression of pluripotency markers was verified in hiPSCs (Supplementary material online, Figure S2C–E).

We then investigated whether all hiPSC lines differentiated in a comparable manner. Principal component analysis of individual transcriptomic profiles as well as correlation analysis showed that samples mainly clustered according to their differentiation stage (Supplementary

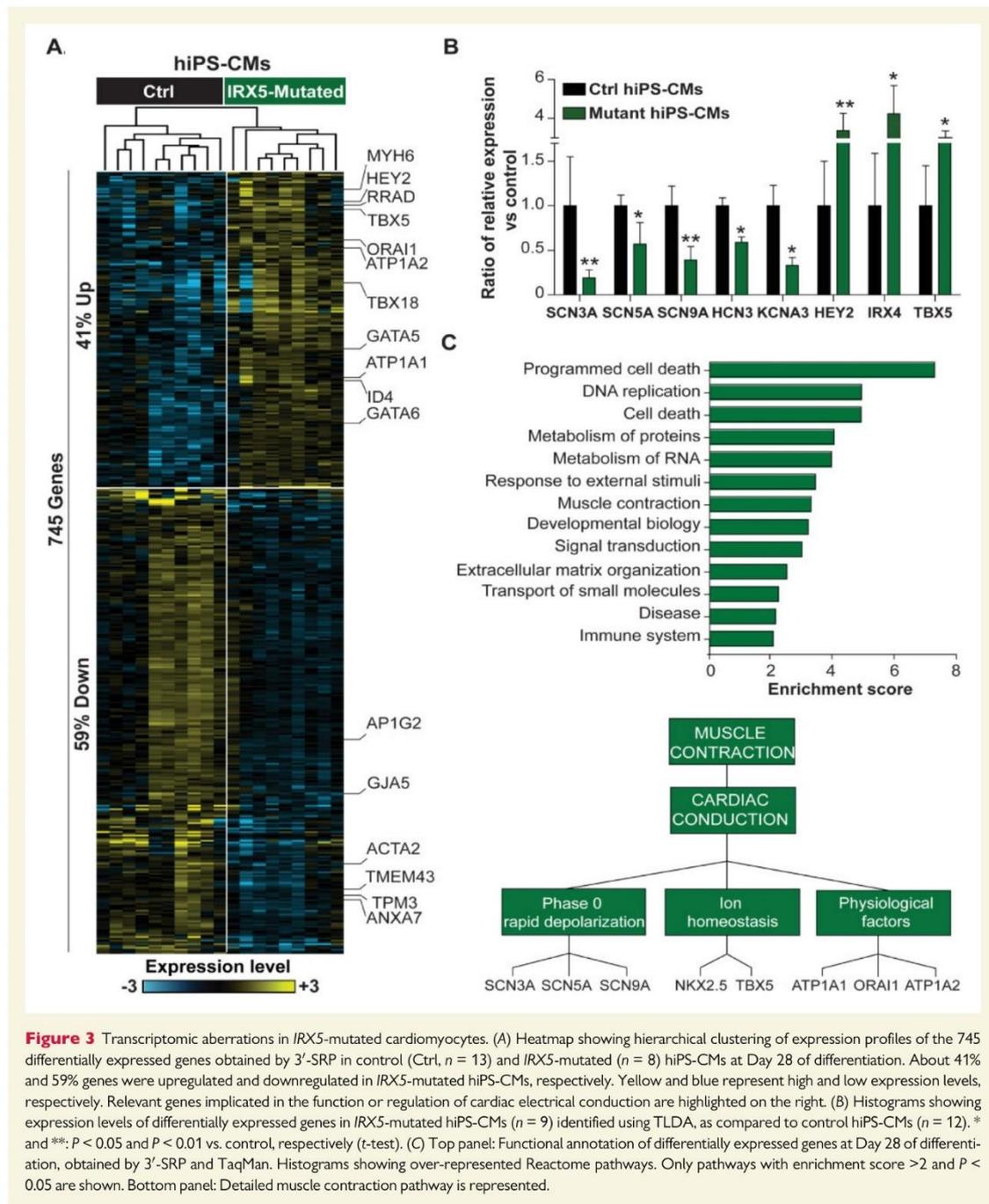
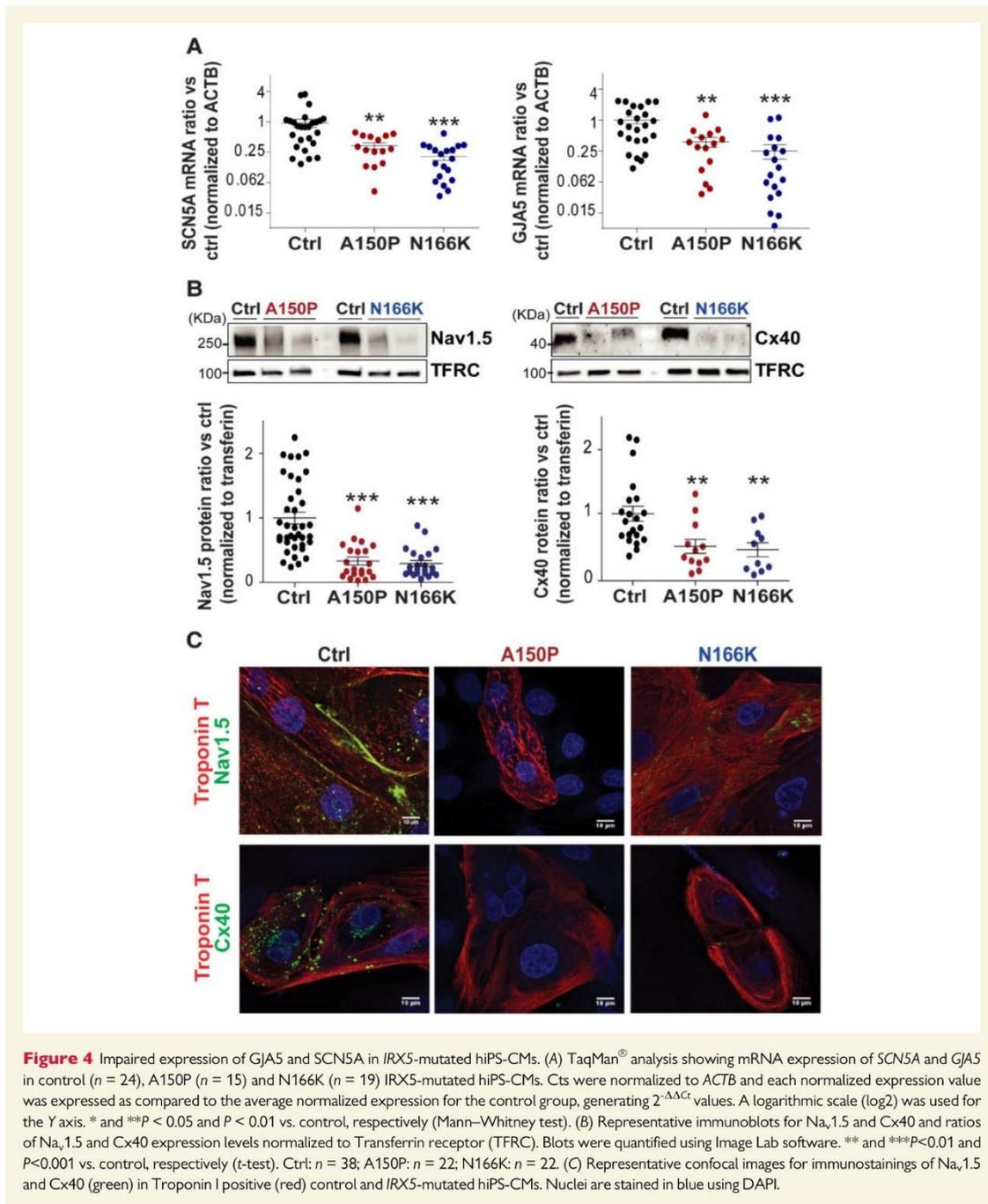


Figure 3 Transcriptomic aberrations in *IRX5*-mutated cardiomyocytes. (A) Heatmap showing hierarchical clustering of expression profiles of the 745 differentially expressed genes obtained by 3'-SRP in control (Ctrl, $n = 13$) and *IRX5*-mutated ($n = 8$) hiPS-CMs at Day 28 of differentiation. About 41% and 59% genes were upregulated and downregulated in *IRX5*-mutated hiPS-CMs, respectively. Yellow and blue represent high and low expression levels, respectively. Relevant genes implicated in the function or regulation of cardiac electrical conduction are highlighted on the right. (B) Histograms showing expression levels of differentially expressed genes in *IRX5*-mutated hiPS-CMs ($n = 9$) identified using TLDA, as compared to control hiPS-CMs ($n = 12$). * and **: $P < 0.05$ and $P < 0.01$ vs. control, respectively (t-test). (C) Top panel: Functional annotation of differentially expressed genes at Day 28 of differentiation, obtained by 3'-SRP and TaqMan. Histograms showing over-represented Reactome pathways. Only pathways with enrichment score >2 and $P < 0.05$ are shown. Bottom panel: Detailed muscle contraction pathway is represented.

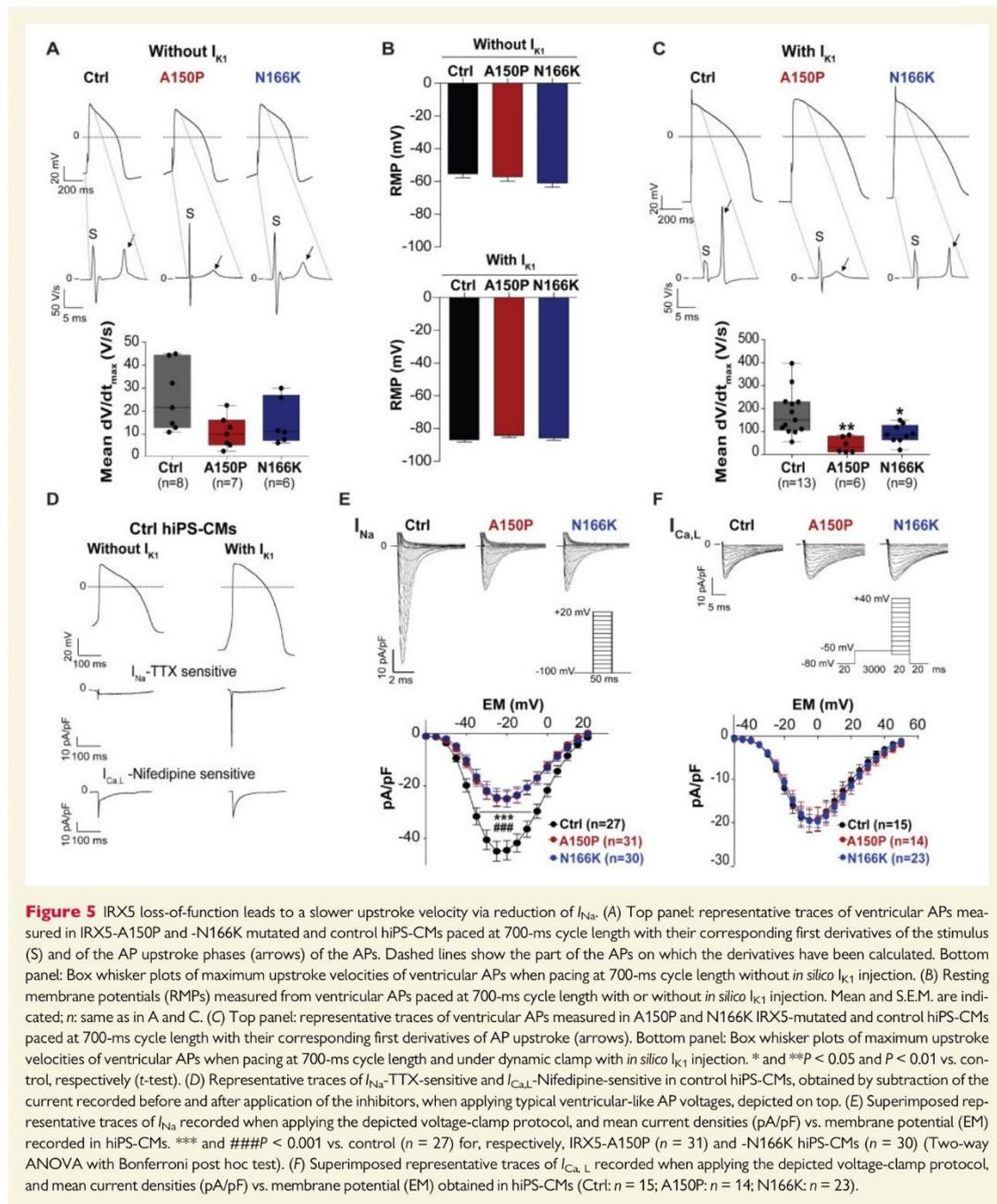
material online, Figure S3A and B). Furthermore, all control and *IRX5*-mutated hiPS-CMs presented a similar global change in gene expression compared to that of hiPSCs (Supplementary material online, Figure S3C). These data showed that all hiPSCs, irrespectively of their genetic background, differentiated similarly into cardiac cells. Immunostaining analysis

showed that striated troponin I was present in CMs derived from each hiPSC line (Figure 2A) and a known marker for ventricular electrical activity, Connexin 43, was present in all lines (Supplementary material online, Figure S4). Expression of specific markers for atrial and ventricular cells revealed the presence of both these cell types in the differentiated



hiPSCs, but not their proportions (Supplementary material online, Figure S5). AP recordings further confirmed the presence of atrial-like and ventricular-like cells, but also nodal-like cells, in proportions that were comparable in control and mutated hiPS-CMs, with the ventricular-like

type forming the majority of hiPS-CMs (Figure 2B). Altogether, these data confirmed that control and *IRX5*-mutated hiPSCs differentiated similarly into CMs, and therefore that hiPS-CMs are a relevant model to investigate the role of *IRX5* in ventricular electrical conduction.

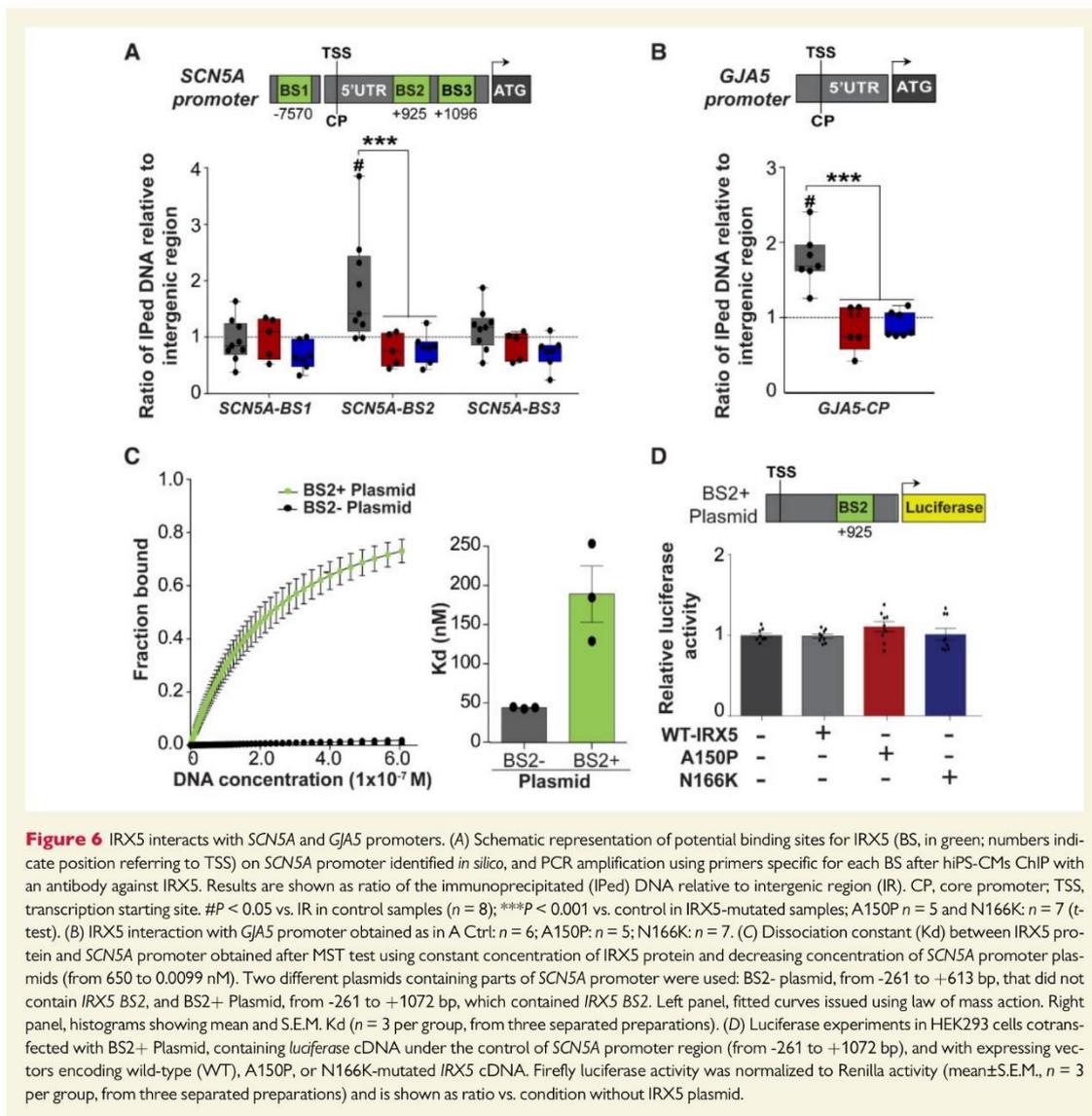


Regarding the impact of *IRX5* mutations on its own expression, transcriptional analysis and nuclear protein expression revelation showed comparable *IRX5* expression levels in control and *IRX5*-mutated hiPS-CMs (Figure 2C and D). Finally, immunostainings of hiPS-CMs constantly showed localization of *IRX5* in CM nuclei (Figure 2A). These data showed that A150P and N166K mutations neither alter *IRX5* expression nor its localization in hiPS-CMs.

3.3 Key cardiac electrical conduction gene expression depends on *IRX5*

Global transcriptomic analysis in hiPS-CMs, using 3'-SRP, unveiled 745 genes differentially expressed between control and *IRX5*-mutated cells.

These include several cardiac electrical conduction genes such as *HEY2* and *TBX5* TFs, that have been shown to present a gradient of expression within the ventricular wall opposite to the one of *IRX5*, suggesting that they may be negatively regulated by *IRX5* (Figure 3A)^{33,34,37}. Using high-throughput real-time RT-PCR (TLDA), we also analyzed a set of 96 genes, including mostly genes with low expression level that could not be properly detected using 3'-SRP (Figure 3B). Overall, for the genes detected with both techniques, expression profiles were concordant (Supplementary material online, Figure S6). Biological pathway enrichment analysis revealed that differentially expressed genes were enriched in pathways involved in cardiac structure, in the regulation of cardiac conduction, and in muscle contraction, and more specifically in the cardiac conduction and phase 0 AP-depolarization functions, mainly linked



to the activity of *SCN5A* and *GJA5* (Figure 3C and Supplementary material online, Figure S3D). More precisely, analyses of IRX5-A150P and -N166K hiPS-CMs concordantly showed major mRNA and protein expression reductions of both *SCN5A* and *GJA5* (Figure 4A and B). In addition, confocal microscopy further confirmed that Na_v1.5 and Cx40 expression was decreased in troponin T positive IRX5-A150P and -N166K mutated hiPS-CMs (Figure 4C).

Taken together, hiPS-CMs carrying IRX5 LOF mutations, exhibited altered expression of key genes involved in cardiac electrical conduction, suggesting a role for IRX5 in the direct transcriptional control of these genes.

3.4 AP depolarization and I_{Na} density are modulated by IRX5

Given the delayed ventricular conduction phenotype observed in HMMS, a slower AP depolarization in patient-derived hiPS-CMs was expected. However, the maximal depolarization velocity (dV/dt_{max}) of ventricular APs, was not statistically reduced in IRX5-mutated hiPS-CMs (Figure 5A). We hypothesized that hiPS-CMs being immature and lacking *I_{K1}*, cellular RMPs (RMPs) of hiPS-CMs were too depolarized (around -60 mV, as shown in Figure 5B), and may not be suitable for proper activation of depolarizing currents. To overcome this issue, an *in silico* *I_{K1}* injection was performed. It brought ventricular RMPs to about -80 mV (Figure 5B) and induced (i) an increase in dV/dt_{max} back to physiological values in control hiPS-CMs and (ii) a significant reduction in dV/dt_{max} in IRX5-mutated hiPS-CMs (Figure 5C). The contribution of *I_{Na}* and *I_{Ca,L}* to AP depolarization as a function of *I_{K1}* injection was then investigated. They were measured when ventricular-like AP voltages, recorded with or without *I_{K1}* injection, were applied to control hiPS-CMs. Without *I_{K1}* injection, *I_{Ca,L}* contribution to the depolarization phase was greater than that of *I_{Na}*, while with *I_{K1}* injection, *I_{Na}* became the main contributor (Figure 5D). This suggests that the dV/dt_{max} reduction observed in IRX5-mutated hiPS-CMs under *I_{K1}* injection was due to a functional reduction of *I_{Na}*, but not of *I_{Ca,L}*. Confirming this hypothesis, *I_{Na}* was reduced by half in IRX5-mutated hiPS-CMs (Figure 5E and Supplementary material online, Figure S7A) without any change in the activation and inactivation curves as well as in the recovery from inactivation kinetics (Supplementary material online, Figure S7B and C and Table S6). In contrast to *I_{Na}* reduction, no modification in *I_{Ca,L}* density was observed (Figure 5F). Together with the *SCN5A* expression alteration, these data suggest that a slower ventricular AP depolarization due to *I_{Na}* reduction may participate to the prolonged QRS observed in HMMS.

3.5 IRX5 interacts with *SCN5A* and *GJA5* promoters

Considering the altered transcription of *SCN5A* and *GJA5* induced by IRX5 LOF, we explored whether IRX5 regulates directly their promoters. An *in silico* analysis showed that *SCN5A* promoter region includes three potential TF-BS for IRX5 (Figure 6A). In accordance with previous studies on mice,⁸ human *GJA5* promoter region was free from any conserved IRX BS (Figure 6B). Immunoprecipitation of IRX5-associated chromatin showed an enrichment of *SCN5A*-BS2 in control hiPS-CMs, revealing an interaction with IRX5, but not in IRX5-mutated hiPS-CMs (Figure 6A). On the other hand, despite the absence of *in silico* identification of conserved IRX-BS on *GJA5* promoter, an interaction between IRX5 and *GJA5* core promoter was found in control hiPS-CMs, that was again absent in IRX5-mutated hiPS-CMs (Figure 6B). These data strongly

suggest that IRX5 interacts with *SCN5A* and *GJA5* promoters to regulate their expression in hiPS-CMs.

To further investigate the direct interaction between *SCN5A* promoter and IRX5, their binding affinity was evaluated. Two different DNA constructs containing *SCN5A* promoter were used: BS2+ plasmid, containing IRX5 BS2 and BS2- plasmid, lacking IRX5 BS2. Interestingly, IRX5 protein did bind to BS2+ plasmid but not to BS2- plasmid, confirming that IRX5 protein binds to *SCN5A* promoter when BS2 is present (Figure 6C).

Finally, the functional consequences of IRX5 binding on *SCN5A* promoter were further evaluated using luciferase assay with the BS2+ plasmid construct. Unexpectedly, neither IRX5 nor its mutated forms activated the luciferase reporter gene (Figure 6D), indicating that another TF may be required to allow IRX5 to regulate *SCN5A* expression. This prompted us to investigate the role of potential critical partners for IRX5, already known to be important for regulation of cardiac development and electrical function, such as GATA4. GATA4, a cardiogenic TF, has indeed been previously shown to regulate *SCN5A* expression³⁸ and an interaction between *Irx5* and another member of the Gata family of TFs, Gata3, has been unveiled during craniofacial development.⁶

3.6 IRX5 and GATA4 cooperatively regulate *SCN5A* expression in human CMs

Comparably in human cardiac ventricular samples and control hiPS-CMs, among GATA family members, GATA4 was the most strongly expressed (Figure 7A). In control hiPS-CMs, GATA4 co-immunoprecipitated with IRX5 and this interaction was not altered in IRX5-mutated hiPS-CMs (Figure 7B). As *SCN5A* promoter contains two sites for GATA4 binding (BSa and BSb),³⁸ chromatin immunoprecipitation (ChIP) experiments using IRX5 antibody showed that IRX5 interacted with GATA4 BSb region in control and IRX5-mutated hiPS-CMs (Figure 7C). Binding affinity analysis further validated the direct interaction between *SCN5A* promoter and the heterodimer IRX5-GATA4, as a binding was observed when both IRX5 and GATA4 proteins were mixed with BS2- plasmid, in opposition to when IRX5 alone was mixed with BS2- plasmid (Supplementary material online, Figure S8). Therefore, IRX5 not only binds directly to BS2 on *SCN5A* promoter but also binds the BSb site together with GATA4. Finally, to further dissect the role of IRX5 and GATA4 on *SCN5A* transcription activation, luciferase assays were performed with plasmids containing either BSb only (BS2- plasmid) or BSb and BS2 (BS2+ plasmid). Co-transfection of GATA4 expressing vector with either plasmid triggered a similar increase in luciferase activity, confirming that GATA4 activated *SCN5A* promoter through a direct interaction with BSb (Figure 7D). Interestingly, when IRX5 was added to these combinations, it potentiated by 1.6 times the effect of GATA4 on BS2- plasmid, and by 3.9 times the effect of GATA4 on BS2+ plasmid (Figure 7D). This further demonstrated that the direct binding of IRX5 to BS2 is important for proper activation of *SCN5A* expression. Conversely, no potentiation of GATA4 activation of *SCN5A* promoter was observed when A150P- or N166K-mutated IRX5 were added to the various combinations (Figure 7D). Interestingly, in contrast to IRX5 and *SCN5A*, which shared similarly graded expression patterns, GATA4 was homogeneously expressed both in atria and ventricles, and throughout the ventricular wall (Supplementary material online, Figure S9A–D). This suggests that GATA4 triggers basal expression of *SCN5A* throughout cardiac tissues, while IRX5 is responsible for the establishment of its expression gradient.

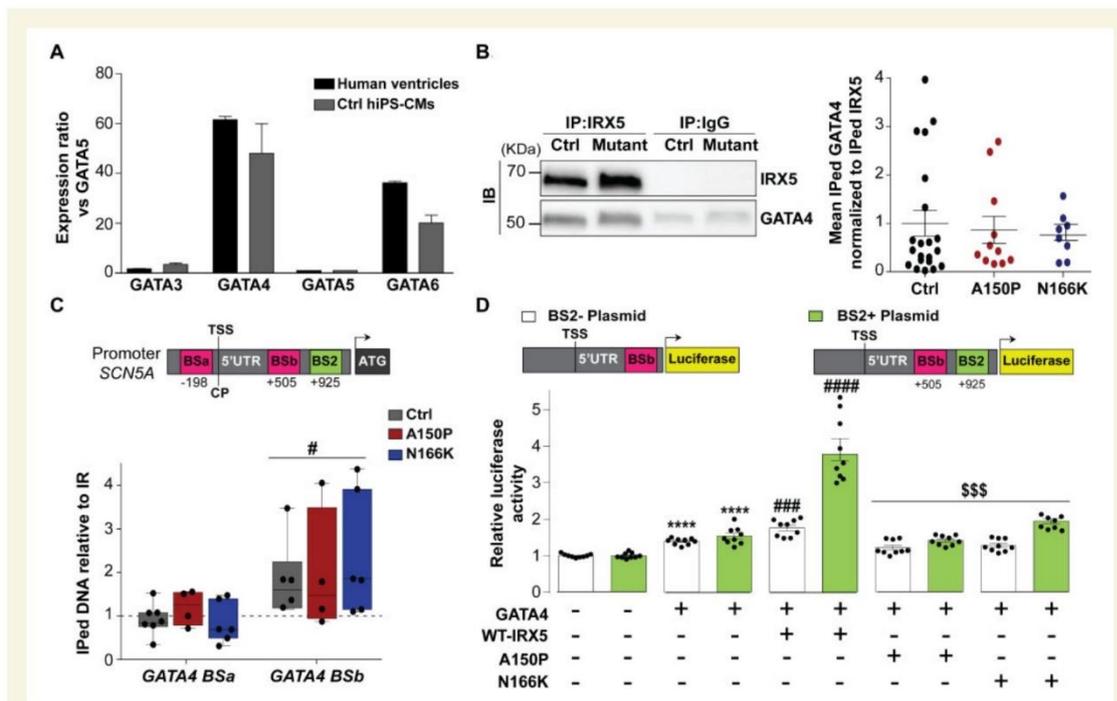


Figure 7 IRX5 and GATA4 cooperate to regulate *SCN5A* expression. (A) GATA transcripts expression ratios vs. *GATA5* (GATA TF with the lowest expression). Transcript expression ratios were calculated using GTEx data for human ventricular tissues and using TLDA analysis for control hiPS-CMs. Data are expressed as mean \pm S.E.M. (B) Representative western blots of IRX5 and GATA4 in control and IRX5-A150P mutated hiPS-CM lysates immunoprecipitated with anti-IRX5 antibody. Quantification of the amount of GATA4 co-immunoprecipitated with IRX5, in control and, IRX5-A150P and -N166K mutated hiPS-CMs. Results are shown as ratios of the co-immunoprecipitated (IPed) GATA4 normalized to the IPed IRX5 amounts (Ctrl: $n = 20$; A150P: $n = 11$; N166K: $n = 8$). (C) Schematic representation of IRX5 binding site (BS; green) and GATA4 BS (pink) on *SCN5A* promoter (numbers indicate position referring to TSS), and PCR amplification using primers specific for each BS after hiPS-CMs ChIP with an antibody against IRX5. Results are shown as ratio of the IPed DNA relative to intergenic region (IR). TSS: transcription starting site. # $P < 0.05$ vs. intergenic region (t-test). Ctrl: $n = 7$; A150P $n = 4$ and N166K: $n = 6$. (D) Luciferase experiments in HEK293 cells cotransfected with BS2- plasmid, containing *luciferase* cDNA under the control of *SCN5A* promoter region (-261 to +613 bp) with GATA4 BSb but not IRX5 BS2 (white bars), or BS2+ plasmid, containing *luciferase* cDNA under the control of *SCN5A* promoter region (-261 to +1072 bp) with GATA4 BSb and IRX5 BS2 (green bars), and with the indicated GATA4 or/and IRX5 expressing plasmids. Firefly luciferase activity was normalized to Renilla and is shown as ratio to its activity when transfected with *SCN5A* plasmids only. Results are displayed for three independent cell preparation and transfection. **** $P < 0.0001$ vs. condition with no IRX5/GATA4 plasmid; ### and ##### $P < 0.001$ and $P < 0.0001$ vs. condition with GATA4 plasmid only; \$\$\$ $P < 0.001$ vs. condition with GATA4 + WT-IRX5 plasmids (t-test).

4. Discussion

This study demonstrates that IRX5 TF is instrumental for proper control of electrical activity in human CMs. hiPS-CMs generated from patients with rare mutations in IRX5 recapitulated their clinical features of ventricular depolarization delay, through an alteration in the regulation of major sodium channel and connexin-encoding gene expression.

4.1 hiPSCs as a model to study TFs

It is commonly acknowledged that stem cell-derived CMs fail to fully recapitulate adult CM features.³⁹ Accordingly, expression of both Connexin 40 and Connexin 43 was observed in the hiPS-CMs. While the adult ventricular myocardium is known to express Connexin 43 but nearly no Connexin 40, known to be specifically expressed in adult Purkinje fibres, prenatal ventricular myocardium expresses Connexin 43 and Connexin 40 in an overlapping pattern.⁴⁰ Thus, the foetal-like

phenotype of ventricular hiPS-CMs may explain why they express not only Connexin 43 but also Connexin 40. Nevertheless, hiPS-CMs have been used successfully so far in >35 studies of cardiac arrhythmic diseases and channelopathies.³⁹ Recently, two studies showed the role of TFs in cellular dysfunction associated to cardiac morphogenetic and electrical defects using hiPS-CMs.^{41,42} The present work additionally enforces the use of hiPSCs as a relevant model for studying the role of TF in the regulation of cardiac electrical conduction. Indeed, with *in silico* I_{K1} injection, we revealed cellular phenotypic traits that are consistent with the patient's clinical phenotype, e.g. QRS complex prolongation observed on the patient's ECG.

4.2 Interspecies regulatory pathway differences

Electrophysiological properties across species are known to be different,⁴³ suggesting that the roles of cardiac TFs can differ from one species

to another. Here, in agreement with this postulate, comparably to what was previously demonstrated for *Irx3* and its regulation of fast conduction in the murine heart,⁸ an induction of *GJA5* expression by IRX5 has been found in human iPSC-derived CMs. However, *Irx5*KO mice did not exhibit conduction delay and murine *Irx5* was not shown to regulate the expression of *Cx*.²⁸

Moreover, the patients, unlike *Irx5*KO mice, do not display ventricular repolarization abnormalities, but a conduction defect. Indeed, the deletion of *Irx5* in mice leads to the abolishment of the T wave on the ECG due to the loss of the transmural gradient of expression of *Kcnd2*, at the root of one major player of ventricular repolarization in mice, I_{to} , but not in human heart, where *KCND3* constitutes the predominant alpha-subunit generating I_{to} .²² Another argument in favour of interspecies differences in gene regulation by IRX TFs is that, due to its transmural gradient of expression, *Irx5* establishes gradients of expression for its target genes. That is how, in mice, *Kcnd2* ventricular gradient is known to be established. In human, it is unlikely that IRX5 regulates *KCND3* expression, as, in opposition to IRX5, its expression is uniform across the ventricular wall.^{7,22}

Along with this rationale, in this study, using patient-derived CMs, a role for IRX5 in the regulation of the expression of the major actors of conduction of the electrical influx, sodium channel subunit, and connexin was elucidated. These data illustrate interspecies differences in the mechanism of action of IRX TFs, demonstrating the importance of using appropriate models to study human diseases.

4.3 Complex regulatory pathway underlies electrical conduction

Conduction of electrical influx throughout the ventricles is mediated not only by connexins but also by inward currents responsible for the depolarization phase of APs. Here, among these inward currents (I_{Na} and I_{Ca}), I_{Na} reduction has been identified as the trigger of the impaired AP upstroke velocity in mutated hiPS-CMs. Concordantly, *SCN5A* mRNA expression level is significantly reduced by 2.8- and 4.6-folds, in A150P and N166K hiPS-CMs as compared to control cells, respectively. At the protein level and as discussed above, Nav1.5 protein expression level is significantly reduced by 3- and 3.3-folds, in A150P and N166K hiPS-CMs as compared to control cells, respectively. And finally, at the functional level, I_{Na} current density is significantly reduced by 2- and 2.7-folds, in A150P and N166K hiPS-CMs as compared to control cells, respectively. In agreement with a direct transcriptional regulation, IRX5 binds to and activates *SCN5A* proximal promoter in control hiPS-CMs whereas mutations in IRX5 prevent this interaction. Surprisingly, at the functional level, luciferase assays showed that IRX5 alone was not able to activate *SCN5A*, suggesting that regulation of *SCN5A* is due to a combinatorial effect of several factors. In accordance, IRX5 and GATA4 have been found to interact with each other allowing IRX5 to bind to a previously described GATA4 BS on *SCN5A* promoter and to enhance GATA4 effect on *SCN5A* activation. Interestingly, while IRX5 and *SCN5A* display a comparable gradient of expression through the ventricular wall, as well as between atria and ventricles, *GATA4* has a more uniform pattern of expression. We thus hypothesized that GATA4 and IRX5 could act cooperatively where GATA4 may be important for basal expression of *SCN5A* in human CMs while IRX5 creates its gradients of expression in the heart. Genetic expression and/or function abnormalities of *SCN5A* together with conduction defects have been previously linked to other rare arrhythmic diseases, such as Brugada syndrome.^{44,45} Thus, the current study elaborating the role of IRX5 in the control of ventricular

conduction through the regulation of *SCN5A* presents a framework for future studies on its potential implication in other complex diseases such as Brugada syndrome.

4.4 Study limitations

Arrhythmic diseases often present with a wide spectrum of phenotypes, even among patients belonging to the same family. The individual genetic background is thought to play a major role in this heterogeneity. In the context of hiPSC-based disease modelling, the use of isogenic control, where the genetic variant has been corrected in the patient iPSC cell line, is increasingly employed. The present study aimed at dissecting IRX5 regulatory mechanisms of human cardiac electrical function by modelling HMMS. Therefore, we chose to perform a comparative analysis between control and two hiPSC lines obtained from two HMMS patients belonging to different families and harbouring different LOF mutations in *IRX5*. For all studied parameters, similar results have been obtained for both types of mutated hiPS-CMs as compared to control cells, strongly supporting a specific role for IRX5 in the pathological phenotype and thereby minimizing the relevance of using isogenic controls.

We suggested the presence of Cx40 and Nav1.5 protein gradients through human adult ventricular wall (Figure 1C), although the presence of Purkinje fibres in the endocardial samples likely participate to the expression differences. Indeed, higher levels of protein and or current have been observed in Purkinje fibres as compared to endocardial myocytes, in the adult heart.^{46,47}

5. Conclusion

In conclusion, the present data suggest that functional IRX5 is essential for regulation of sodium channel- and connexin-gene expression, ensuring proper conduction of the electrical influx between CMs. Furthermore, a role for this TF in the orchestration of the gene expression programme governing ventricular electrical depolarization in human was uncovered, bringing to light a new potential option for therapeutic intervention.

Supplementary material

Supplementary material is available at *Cardiovascular Research* online.

Authors' contributions

Z.R.A.S. contributed to the acquisition, interpretation of data, and writing of the report; R.C. contributed to data acquisition; B.C. contributed to data acquisition; C.B. contributed to clinical data acquisition; J.-B.G. contributed to clinical data analysis; H.H. contributed to clinical data acquisition; H.K. contributed to clinical data acquisition; A.G. contributed to data acquisition; M.J. contributed to data acquisition; N.J. contributed to data acquisition; A.G. contributed to data acquisition; C.C. contributed to data acquisition; L.D. contributed to data acquisition; V.F. contributed to data acquisition; C.M. contributed to data acquisition; F.C. contributed to writing of the report; G.L. contributed to data analysis, and writing of the report; G.L. contributed to data analysis, and writing of the report; B.R. contributed to clinical data acquisition; K.Z. contributed to the writing of the report; P.L. contributed to the study design, interpretation of the data, and writing of the report; N.G. contributed to the study design, interpretation of the data, and writing of the report. All

authors discussed the results and contributed and approved the final manuscript.

Acknowledgements

The authors would like to thank the Genomics and Bioinformatics Core Facility (GenoBiRD, Biogenouest), the Cytometry facility (CytoCell), the iPS core facility, and the Cellular and Tissular Imaging Core Facility (MicroPICell) of Nantes University, for their technical support. The authors also would like to thank Dan Roden for the gift of SCN5A plasmids, and Sophie Demolombe for helping with mouse studies.

Conflict of interest: None declared.

Funding

This work was funded by grants from The National Research Agency (HEART iPS ANR-15-CE14-0019-01), and *La Fédération Française de Cardiologie*. N.G. was laureate of fellowships from Fondation Lefoulon-Delalande and Marie Curie Actions International Incoming Fellowship FP7-PEOPLE-2012-IIF (PIIF-GA-2012-331436). Z.R.A.S. is supported by Eiffel scholarship programme of Excellence (Campus France), by Doctoral School of Science and Technology-Lebanese University and The Fondation Genavie.

Data availability

The authors declare that all data supporting the findings of this study are available within the article and its supplementary information files or from the corresponding author upon reasonable request. All read sequence data and sample annotations generated for this article are available at European Nucleotide Archive (ENA) with accession number PRJEB39654.

References

- Pazoki R, Wilde AAM, Bezzina CR. Genetic basis of ventricular arrhythmias. *Curr Cardio Risk Rep* 2010;**4**:454–460.
- Gourraud J-B, Barc J, Thollet A, Le Scouarnec S, Le Marec H, Schott JJ, Redon R, Probst V. The Brugada syndrome: a rare arrhythmia disorder with complex inheritance. *Front Cardiovasc Med* 2016;**3**:9.
- Eif VVVV, Devalla HD, Boink GJ, Christoffels VM. Transcriptional regulation of the cardiac conduction system. *Nat Rev Cardiol* 2018;**15**:617–630.
- Hu W, Xin Y, Zhang L, Hu J, Sun Y, Zhao Y. Iroquois homeodomain transcription factors in ventricular conduction system and arrhythmia. *Int J Med Sci* 2018;**15**:808–815.
- Hamamy HA, Teebi AS, Oudjhane K, Shegem N, Ajlouni K. Severe hypertelorism, midface prominence, prominent/simple ears, severe myopia, borderline intelligence, and bone fragility in two brothers: new syndrome? *Am J Med Genet A* 2007;**143A**:229–234.
- Bonnard C, Strobl AC, Shboul M, Lee H, Merriman B, Nelson SF, Ababneh OH, Uz E, Güran T, Kayserili H, Hamamy H, Reversade B. Mutations in *IRX5* impair craniofacial development and germ cell migration via *SDF1*. *Nat Genet* 2012;**44**:709–713.
- Costantini DL, Arruda EP, Agarwal P, Kim K-H, Zhu Y, Zhu W, Lebel M, Cheng CW, Park CY, Pierce SA, Guerschicoff A, Pollevick GD, Chan TY, Kabir MG, Cheng SH, Husain M, Antzelevitch C, Srivastava D, Gross GJ, Hui C, Backx PH, Bruneau BG. The homeodomain transcription factor *Irx3* establishes the mouse cardiac ventricular repolarization gradient. *Cell* 2005;**123**:347–358.
- Zhang S-S, Kim K-H, Rosen A, Smyth JW, Sakuma R, Delgado-Olguin P, Davis M, Chi NC, Puviondran V, Gaborit N, Sukonnik T, Wylie JN, Brand-Arzamendi K, Farman GP, Kim J, Rose RA, Marsden PA, Zhu Y, Zhou Y-Q, Miquerol L, Henkelman RM, Stainier DYR, Shaw RM, Hui C, Bruneau BG, Backx PH. Iroquois homeobox gene 3 establishes fast conduction in the cardiac His-Purkinje network. *Proc Natl Acad Sci USA* 2011;**108**:13576–13581.
- Kim K-H, Rosen A, Hussein SM, Puviondran V, Korogiy AS, Chiarello C, Nagy A, Hui C, Backx PH. *Irx3* is required for postnatal maturation of the mouse ventricular conduction system. *Sci Rep* 2016;**6**:19197.
- Marionneau C, Couette B, Liu J, Li H, Mangoni ME, Nargeot J, Lei M, Escande D, Demolombe S. Specific pattern of ionic channel gene expression associated with pacemaker activity in the mouse heart. *J Physiol* 2005;**562**:223–234.
- Jouni M, Si-Tayeb K, Es-Salah-Lamoureux Z, Latypova X, Champon B, Caillaud A, Rungoat A, Charpentier F, Loussouarn G, Baró I, Zibara K, Lemarchand P, Gaborit N. Toward personalized medicine: using cardiomyocytes differentiated from urine-derived pluripotent stem cells to recapitulate electrophysiological characteristics of type 2 long QT syndrome. *J Am Heart Assoc* 2015;**4**:e002159.
- Es-Salah-Lamoureux Z, Jouni M, Malak OA, Belbachir N, Al Sayed ZR, Gandon-Renard M, Lamirault G, Gauthier C, Baró I, Charpentier F, Zibara K, Lemarchand P, Beaumelle B, Gaborit N, Loussouarn G. HIV-Tat induces a decrease in I_{Kr} and I_{Ks} via reduction in phosphatidylinositol-(4,5)-bisphosphate availability. *J Mol Cell Cardiol* 2016;**99**:1–13.
- Si-Tayeb K, Noto FK, Nagaoka M, Li J, Battle MA, Duris C, North PE, Dalton S, Duncan SA. Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepato Bilim Md* 2010;**51**:297–305.
- Wang W-X, Danaher RJ, Miller CS, Berger JR, Nubia VG, Wilfred BS, Neltner JH, Norris CM, Nelson PT. Expression of miR-15/107 family microRNAs in human tissues and cultured rat brain cells. *Genomics Proteomics Bioinformatics* 2014;**12**:19–30.
- Mestdagh P, Van Vlierbergh P, De Weer A, Muth D, Westermann F, Speleman F, Vandesompele J. A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol* 2009;**10**:R64.
- Bockmeyer CL, Säuberlich K, Wittig J, Eßer M, Roeder SS, Vester U, Hoyer PF, Agustian PA, Zeuschner P, Amann K, Daniel C, Becker JU. Comparison of different normalization strategies for the analysis of glomerular microRNAs in IgA nephropathy. *Sci Rep* 2016;**6**:31992.
- Kilens S, Meistermann D, Moreno D, Chariou C, Gaignerie A, Reigner A, Lelièvre Y, Casanova M, Vallot C, Nedellec S, Flippe L, Firmin J, Song J, Charpentier E, Lammers J, Donnat A, Marec N, Deb W, Bihouée A, Caignec CL, Pecqueur C, Redon R, Barrière P, Bourdon J, Pasque V, Soumillon M, Mikkelsen TS, Rougeulle C, Fréour T, David L, The Milieu Intérieur Consortium. Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nat Commun* 2018;**9**:360.
- Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P. The Reactome pathway knowledgebase. *Nucleic Acids Res* 2014;**42**:D472–D477.
- Fabregat A, Korminger F, Viteiri G, Sidiropoulos K, Marin-Garcia P, Ping P, Wu G, Stein L, D'Eustachio P, Hermjakob H. Reactome graph database: efficient access to complex pathway data. *PLoS Comput Biol* 2018;**14**:e1005968.
- Medina PP, Nolde M, Slack FJ. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature* 2010;**467**:86–90.
- Sartor MA, Leikauf GD, Medvedovic M. LRpath: a logistic regression approach for identifying enriched biological groups in gene expression data. *Bioinforma Oxf Engl* 2009;**25**:211–217.
- Gaborit N, Varro A, Le Bouter S, Szuts V, Escande D, Nattel S, Demolombe S. Gender-related differences in ion-channel and transporter subunit expression in non-diseased human hearts. *J Mol Cell Cardiol* 2010;**49**:639–646.
- Biloni A, Craig G, Hill C, McNeill H. Iroquois transcription factors recognize a unique motif to mediate transcriptional repression in vivo. *Proc Natl Acad Sci USA* 2005;**102**:14671–14676.
- Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimeh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML, Hughes TR. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell* 2008;**133**:1266–1276.
- Noyes MB, Christensen RG, Wakabayashi A, Stormo GD, Brodsky MH, Wolfe SA. Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* 2008;**133**:1277–1289.
- Peters T, Dildrop R, Ausmeier K, Rütter U. Organization of mouse Iroquois homeobox genes in two clusters suggests a conserved regulation and function in vertebrate development. *Genome Res* 2000;**10**:1453–1462.
- Houweling AC, Dildrop R, Peters T, Mummehoff J, Moorman AF, Rütter U, Christoffels VM. Gene and cluster-specific expression of the Iroquois family members during mouse development. *Mech Dev* 2001;**107**:169–174.
- Gaborit N, Sakuma R, Wylie JN, Kim K-H, Zhang S-S, Hui C-C, Bruneau BG. Cooperative and antagonistic roles for *Irx3* and *Irx5* in cardiac morphogenesis and postnatal physiology. *Dev Camb Engl* 2012;**139**:4007–4019.
- Putten VERM, Mengarelli I, Guan K, Zegers JG, Ginneken VGAC, Verkerk AO, Wilders R. Ion channelopathies in human induced pluripotent stem cell derived cardiomyocytes: a dynamic clamp study with virtual IK1. *Front Physiol* 2015;**6**:7.
- Bett GC, Kaplan AD, Lis A, Cimato TR, Tzanakakis ES, Zhou Q, Morales MJ, Rasmuson RL. Electronic « expression » of the inward rectifier in cardiocytes derived from human-induced pluripotent stem cells. *Heart Rhythm* 2013;**10**:1903–1910.
- Gaborit N, Le Bouter S, Szuts V, Varro A, Escande D, Nattel S, Demolombe S. Regional and tissue specific transcript signatures of ion channel genes in the non-diseased human heart. *J Physiol* 2007;**582**:675–693.
- Remme CA, Verkerk AO, Hoogaars WMH, Aanhaenen WTJ, Scicluna BP, Annink C, den Hoff MJB, Wilde AAM, Veen TAB, Veldkamp MW, Bakker JMT, Christoffels VM, Bezzina CR. The cardiac sodium channel displays differential distribution in the conduction system and transmural heterogeneity in the murine ventricular myocardium. *Basic Res Cardiol* 2009;**104**:511–522.

33. Veerman CC, Podliesna S, Tadros R, Lodder EM, Mengarelli I, Jonge B, de Beekman L, Barc J, Wilders R, Wilde AAM, Boukens BJ, Coronel R, Verkerk AO, Remme CA, Bezzina CR. The Brugada syndrome susceptibility gene HEY2 modulates cardiac transmural ion channel patterning and electrical heterogeneity novelty and significance. *Circ Res* 2017;**121**:537–548.
34. Rosati B, Grau F, Mckinnon D. Regional variation in mRNA transcript abundance within the ventricular wall. *J Mol Cell Cardiol* 2006;**40**:295–302.
35. Rijnbeek PR, Witsenburg M, Schrama E, Hess J, Kors JA. New normal limits for the paediatric electrocardiogram. *Eur Heart J* 2001;**22**:702–711.
36. Surawicz B, Childers R, Deal BJ, Gettes LS, Bailey JJ, Gorgels A, Hancock EW, Josephson M, Kligfield P, Kors JA, Macfarlane P, Mason JW, Mirvis DM, Okin P, Pahlm O, Rautaharju PM, van Herpen G, Wagner GS, Wellens H, American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology, American College of Cardiology Foundation, Heart Rhythm Society, AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part III: intraventricular conduction disturbances: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. *J Am Coll Cardiol* 2009;**53**:976–981.
37. Hatcher CJ, Goldstein MM, Mah CS, Delia CS, Basson CT. Identification and localization of TBX5 transcription factor during human cardiac morphogenesis. *Dev Dyn* 2000;**219**:90–95.
38. Tarradas A, Pinsach-Abuin M, Mackintosh C, Llorà-Batlle O, Pérez-Serra A, Batlle M, Pérez-Villa F, Zimmer T, Garcia-Bassets I, Brugada R, Beltran-Alvarez P, Pagans S. Transcriptional regulation of the sodium channel gene (SCN5A) by GATA4 in human heart. *J Mol Cell Cardiol* 2017;**102**:74–82.
39. Giacomelli E, Mummery CL, Bellin M. Human heart disease: lessons from human pluripotent stem cell-derived cardiomyocytes. *Cell Mol Life Sci* 2017;**74**:3711–3739.
40. Kempen MJ, Vermeulen JL, Moorman AF, Gros D, Paul DL, Lamers WH. Developmental changes of connexin40 and connexin43 mRNA distribution patterns in the rat heart. *Cardiovasc Res* 1996;**32**:886–900.
41. Ang Y-S, Rivas RN, Ribeiro AJS, Srivas R, Rivera J, Stone NR, Pratt K, Mohamed TMA, Fu J-D, Spencer CI, Tippens ND, Li M, Narasimha A, Radzinsky E, Moon-Grady AJ, Yu H, Pruitt BL, Snyder MP, Srivastava D. Disease model of GATA4 mutation reveals transcription factor cooperativity in human cardiogenesis. *Cell* 2016;**167**:1734–1749.e22.
42. Caballero R, Utrilla RG, Amorós I, Matamoros M, Pérez-Hernández M, Tinaquero D, Alfayate S, Nieto-Marín P, Guerrero-Serna G, Liu Q-H, Ramos-Mondragón R, Ponce-Balbuena D, Herron T, Campbell KF, Filgueiras-Rama D, Peinado R, López-Sendón JL, Jalife J, Delpón E, Tamargo J. Tbx20 controls the expression of the KCNH2 gene and of hERG channels. *Proc Natl Acad Sci USA* 2017;**114**:E416–E425.
43. Muszkiewicz A, Britton OJ, Gemmell P, Passini E, Sánchez C, Zhou X, Carusi A, Quinn TA, Burrage K, Bueno-Orovio A, Rodriguez B. Variability in cardiac electrophysiology: using experimentally-calibrated populations of models to move beyond the single virtual physiological human paradigm. *Prog Biophys Mol Biol* 2016;**120**:115–127.
44. Gaborit N, Wichter T, Varro A, Szuts V, Lamirault G, Eckardt L, Paul M, Breithardt G, Schulze-Bahr E, Escande D, Nattel S, Demolombe S. Transcriptional profiling of ion channel genes in Brugada syndrome and other right ventricular arrhythmogenic diseases. *Eur Heart J* 2009;**30**:487–496.
45. Juang J-M, Tsai C-T, Lin L-Y, Liu Y-B, Yu C-C, Hwang J-J, Chen J-J, Chiu F-C, Chen W-J, Tseng C-D, Chiang F-T, Yeh H-M, Sherri Yeh S-F, Lai L-P, Lin J-L. Unique clinical characteristics and SCN5A mutations in patients with Brugada syndrome in Taiwan. *J Formos Med Assoc* 2015;**114**:620–626.
46. Lin X, Liu N, Lu J, Zhang J, Anumonwo JMB, Isom LL, Fishman GI, Delmar M. Subcellular heterogeneity of sodium current properties in adult cardiac ventricular myocytes. *Heart Rhythm* 2011;**8**:1923–1930.
47. Gros D, Jarry-Guichard T, Ten Velde I, de Maziere A, van Kempen MJ, Davoust J, Briand JP, Moorman AF, Jongsma HJ. Restricted distribution of connexin40, a gap junctional protein, in mammalian heart. *Circ Res* 1994;**74**:839–851.

Translational perspective

Inherited cardiac arrhythmias account for ~20% of sudden cardiac deaths, of which a small portion are monogenic familial diseases with mutations in cardiac ion channels. However, pathogeny of inherited cardiac arrhythmias is increasingly thought to result from complex mechanisms involving altered regulation of multiple effectors expression. Taking advantage of cardiomyocytes derived from HMMS patients, carrying loss-of-function mutations in *IRX5* transcription factor, we uncovered an important role for *IRX5* in the regulation of several major players of ventricular depolarization conduction and in arrhythmogenesis. Thus, this study supports systematic screening for genetic variants in *IRX5* in inherited cardiac arrhythmias.

Annexe 2

Generation of human induced pluripotent stem cell lines from four unrelated healthy control donors carrying European genetic background

Aurore Girardeau[#], Diane Atticus[#], **Robin Canac**, Bastien Cimarosti, Amandine Caillaud, Caroline Chariou, Floriane Simonet, Bertrand Cariou, Flavien Charpentier, Jean-Baptiste Gourraud, Vincent Probst, Nadjat Belbachir, Laurence Jesel, Patricia Lemarchand, Julien Barc, Richard Redon, Nathalie Gaborit*, Guillaume Lamirault*

: Co-premiers auteurs

* : Co-derniers auteurs

DOI : 10.1016/j.scr.2021.102647 ; PMID : 34999420

Publié dans *Stem Cell Research*, 2021



Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Lab Resource: Multiple Cell Lines



Generation of human induced pluripotent stem cell lines from four unrelated healthy control donors carrying European genetic background

Aurore Girardeau^{a,1}, Diane Atticus^{a,1}, Robin Canac^a, Bastien Cimarosti^a, Amandine Caillaud^a, Caroline Chariou^b, Floriane Simonet^a, Bertrand Cariou^a, Flavien Charpentier^a, Jean-Baptiste Gourraud^a, Vincent Probst^a, Nadjet Belbachir^{a,c}, Laurence Jesel^d, Patricia Lemarchand^a, Julien Barc^a, Richard Redon^a, Nathalie Gaborit^{a,2}, Guillaume Lamirault^{a,2,*}

^a Université de Nantes, CHU Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France^b Université de Nantes, CHU Nantes, INSERM, CNRS, SFR Santé, FED 4203, INSERM UMS 016, CNRS UMS 3556, F-44000 Nantes, France^c Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA^d CHU Strasbourg, Service de Cardiologie, Strasbourg, France

A B S T R A C T

Four human induced pluripotent stem cell (hiPSC) lines have been generated from healthy control European donors, and validated. This resource represents a useful tool for stem cell-based research, as references for developmental studies and disease modeling linked to any type of human tissue and organ, in an ethical-, sex- and age-matched context. They provide a reliable in-vitro model for single cell- and tissue-based investigations, and are also a valuable tool for genome editing-based studies.

1. Resource Table:

(continued)

Unique stem cell lines identifier	ITXi005-A ITXi006-A ITXi007-A ITXi008-A		Age: 55 Sex: Male Ethnicity: European
Alternative name(s) of stem cell lines	MS573 (ITXi005-A) WT8288 (ITXi006-A) 202CT (ITXi007-A) SD378M (ITXi008-A)		202CT (ITXi007-A) Age: 29 Sex: Female Ethnicity: European
Institution	L'institut du thorax (INSERM UMR1087, CNRS UMR6291)		SD378M (ITXi008-A) Age: 49 Sex: Female Ethnicity: European
Contact information of distributor	Guillaume Lamirault, guillaume.lamirault@univ-nantes.fr		MS573, WT8288 and SD378M: Peripheral blood mononuclear cells 202CT: Urine progenitor cells Clonal (molecular karyotyping) Method of reprogramming MS573, WT8288 and SD378M: Integration-free Sendai virus expressing human OCT4, SOX2, KLF4, and c-MYC 202CT: Episomal system
Type of cell lines	iPSC	Cell Source	
Origin	Human	Clonality	
Additional origin info required	MS573 (ITXi005-A) Age: 16 Sex: Male Ethnicity: European WT8288 (ITXi006-A)	Method of reprogramming	

(continued on next column)

(continued on next page)

* Corresponding author at: l'institut du thorax Inserm UMR 1087, CNRS UMR 6291, IRS-UN, 8 quai Moncoussu, 44007 Nantes cedex 1, France.

E-mail address: guillaume.lamirault@univ-nantes.fr (G. Lamirault).¹ Contributed equally to this work.² Contributed equally to this work.<https://doi.org/10.1016/j.scr.2021.102647>

Received 26 November 2021; Accepted 27 December 2021

Available online 28 December 2021

1873-5061/© 2022 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

<http://creativecommons.org/licenses/by-nc-nd/4.0/>.

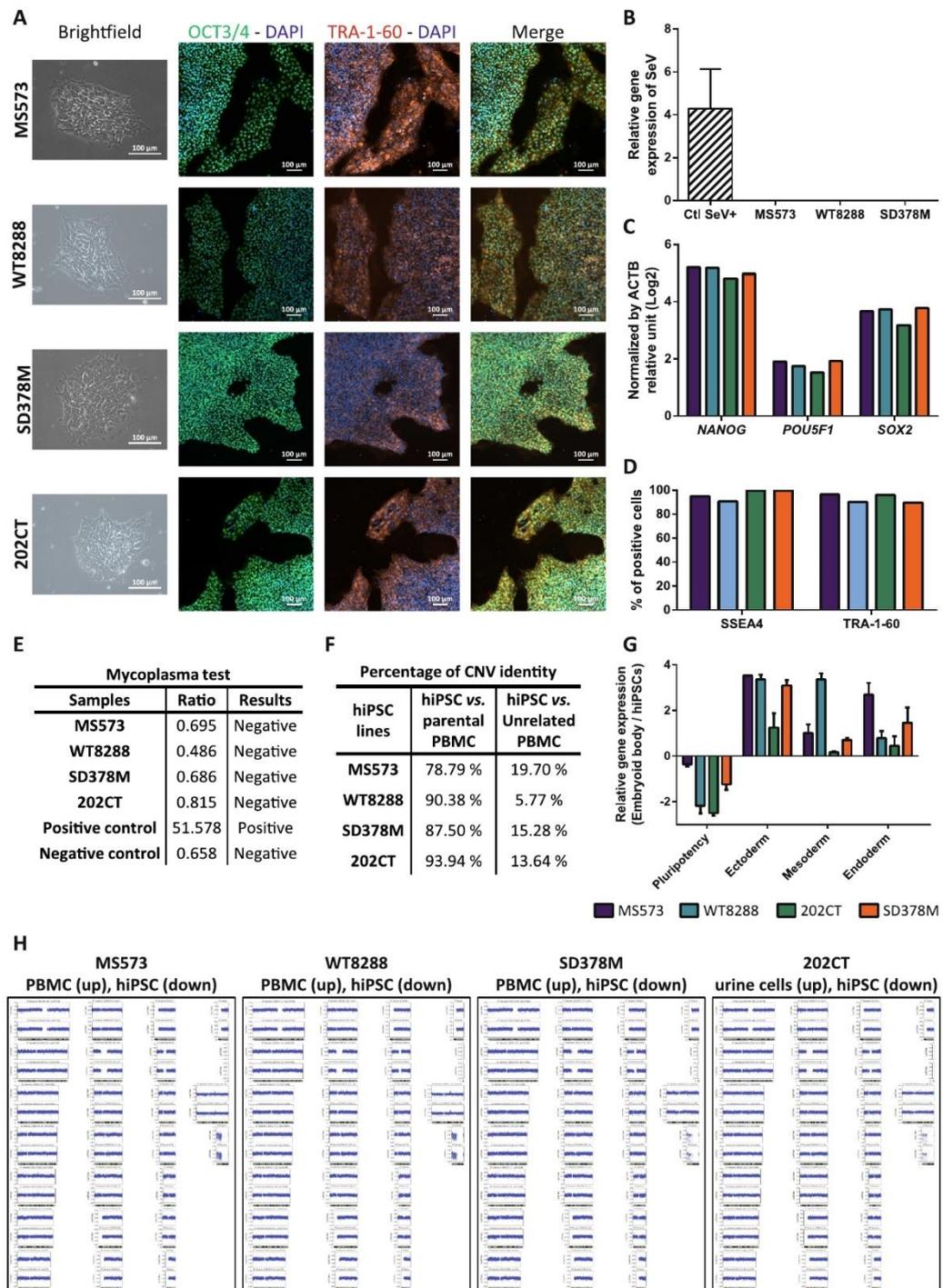


Fig. 1.

(continued)

Genetic Modification	No
Type of Genetic Modification	N/A
Evidence of the reprogramming transgene loss (including genomic copy if applicable)	RT-qPCR
Associated disease	N/A
Gene/locus	N/A
Date archived/stock date	MS573: 09-13-2016 WT8288: 04-16-2015 202CT: 03-14-2018 SD378M: 02-05-2019
Cell line repository/bank	MS573: https://hpscereg.eu/cell-line/ITXi005-A WT8288: https://hpscereg.eu/cell-line/ITXi006-A 202CT: https://hpscereg.eu/cell-line/ITXi007-A SD378M: https://hpscereg.eu/cell-line/ITXi008-A
Ethical approval	A signed informed consent has been obtained from all donors.. Authorizations have been obtained from competent ethics review board (CPP Ouest IIn°2010-33) and approval authority (MESR n°DC-2011-1399) to generate these lines.

2. Resource utility

Somatic cells from four healthy control donors, belonging to the French population, have been reprogrammed into hiPSC lines using non-integrative methods. They can be used as a reference for disease modeling, developmental investigations and genetic editing, especially in the context of ethnical-, sex- and age-matched studies.

3. Resource details

In the field of stem cell research there is a need for ethnical-, sex- and age-matched control hiPSC lines as a proper experimental reference (Volpato and Webber, 2020). Therefore, in the present study, we selected four healthy donors from the French population that varies in term of age and sex, in order to generate hiPSC lines. European ancestry of the four individuals was assessed after genotyping on Affymetrix PMRA array, pruning ($r^2 < 0.2$) and creating an identity-by-state (IBS) matrix containing (1) the four cell lines, (2) European ancestry and (3) reference populations from the 1000 Genomes Project. Two of them, MS573 and WT8288, are male subjects of age 16 and 55, respectively (Belbachir et al., 2019). The other two, 202CT and SD378M, are female subjects of age 29 and 49, respectively. Furthermore, hiPSC lines have been generated from peripheral blood mononuclear cells (PBMCs) using Sendai virus method for MS573, WT8288 and SD378M, and from urine progenitor cells using episomal system for 202CT.

All cell lines were cultured in StemMACS iPS-Brew XF medium on Matrigel®-coated plates (Fig. 1A left panels) and were mycoplasma free (Fig. 1E). Also, loss of Sendai virus expression was verified by RT-qPCR for the three lines reprogrammed by this technology (Fig. 1B). Proper expression of OCT3/4 and TRA-1-60 pluripotency proteins was assessed by immunocytochemistry (Fig. 1A right panels). Furthermore, *NANOG*, *POU5F1* and *SOX2* were highly expressed in all cell lines and more than 90% of the cells were positive for SSEA4 and TRA-1-60 expression (Fig. 1C and 1D). Differentiation potential into derivatives of the 3 embryonic germ layers was confirmed using spontaneous embryoid body formation and differentiation followed by RT-qPCR analysis. It showed upregulation of ectodermal, mesodermal and endodermal markers, while pluripotency markers were downregulated (Fig. 1G). Finally, no chromosomal aberrations were revealed with molecular

Table 1
Characterization and validation.

Classification	Test	Result	Data
Morphology	Photography Bright field	Normal	Fig. 1A left panels
Phenotype	Qualitative analysis (Immunocytochemistry)	Assess staining/ expression of pluripotency markers: OCT3/4, TRA-1-60	Fig. 1A right panels
	Quantitative analysis (Flow cytometry and RT-qPCR)	MS573: <ul style="list-style-type: none"> TRA-1-60: 96.2% SSEA-4: 94.7% Expression of <i>NANOG</i>, <i>POU5F1</i> and <i>SOX2</i> WT8288: <ul style="list-style-type: none"> TRA-1-60: 90.0% SSEA-4: 90.3% Expression of <i>NANOG</i>, <i>POU5F1</i> and <i>SOX2</i> 202CT: <ul style="list-style-type: none"> TRA-1-60: 95.6% SSEA-4: 99.8% Expression of <i>NANOG</i>, <i>POU5F1</i> and <i>SOX2</i> SD378M: <ul style="list-style-type: none"> TRA-1-60: 89.3% SSEA-4: 99.8% Expression of <i>NANOG</i>, <i>POU5F1</i> and <i>SOX2</i> 	Fig. 1C and 1D
Genotype	Molecular karyotyping (SNP array; greater than 300000 markers)	MS573: 46,XY; Normal karyotype; WT8288: 46,XY; Normal karyotype; SD378M: 46,XX; Normal karyotype; 202CT: 46,XX; Normal karyotype	Fig. 1H
Identity	CNV analysis	SNP array to compare MS573, WT8288, SD378M and 202CT with blood or urine cells. Matching genotypes confirm the origins of iPSC lines	Fig. 1F and archived with the journal
		MS573: 66 tested sites WT8288: 60 tested sites SD378M: 72 tested sites 202CT: 66 tested sites	Fig. 1F and archived with the journal
Mutation analysis (IF APPLICABLE)	Sequencing Southern Blot OR WGS	N/A N/A	N/A N/A
Microbiology and virology	Mycoplasma	Mycoplasma testing by luminescence: Negative	Fig. 1E

(continued on next page)

Table 1 (continued)

Classification	Test	Result	Data
Differentiation potential	Embryoid body formation and differentiation followed by RT-qPCR analysis	Expression of genes in embryoid bodies: Ectoderm (<i>PAX6</i>); Mesoderm (<i>HAND1</i>); Endoderm (<i>SOX17</i>);	Fig. 1G
List of recommended germ layer markers	Expression of markers at mRNA level (RT-qPCR)	<ul style="list-style-type: none"> • Ectoderm: <i>PAX6</i> • Mesoderm: <i>HAND1</i> • Endoderm: <i>SOX17</i> 	Fig. 1G
Donor screening (OPTIONAL)	HIV 1 + 2 Hepatitis B, Hepatitis C	N/A	N/A
Genotype additional info (OPTIONAL)	Blood group genotyping, HLA tissue typing	N/A	N/A

karyotyping (Fig. 1H) and parentality between each hiPSC line and its corresponding parental somatic cells was confirmed (Fig. 1F). Characterization of the four hiPSC lines is summarized in Table 1 and Table 2.

4. Materials and methods

4.1. Generation and maintenance of hiPSC lines

Collected blood cells were reprogrammed by the iPSC core facility (PF-iPSC, Nantes) using Sendai virus method (Cytotune reprogramming Kit, Life technologies) for MS573, WT8288 and SD378M. Regarding the 202CT hiPSC line, episomal integration-free reprogramming was performed on urine progenitor cells (Epi5™ Episomal iPSC Reprogramming Kit, Life Technologies). hiPSC lines were maintained at 37 °C, 5% CO₂, 21% O₂ in StemMACS™ iPS Brew XF Medium (Miltenyi Biotec) on culture plates coated with Matrigel® hESC-Qualified Matrix (0.05 mg/mL, Corning). At 75% confluency, cells were passaged using Gentle Cell Dissociation Reagent (STEMCELL™ Technologies).

4.2. Flow cytometry analysis

A total of 1x10⁵ cells in suspension were washed 3 times with PBS/BSA 0.1% and then incubated with PE-labelled antibodies for 30 min at 4 °C in the dark. Data acquisition was performed using FACSDiva™ software with LSR II instrument (Becton Dickinson, BD).

4.3. Immunofluorescence

Cells were dissociated and seeded onto 8-well IBIDI plates coated with Matrigel® (p22-MS573; p27-WT8288; p29-SD378M; p44-202CT). Cultured cells were fixed with 4% paraformaldehyde for 15 min, permeabilized with 0.5% Triton-X100 for 15 min and blocked with 3% BSA for 30 min. Immunofluorescence stainings were performed using appropriate primary antibodies (4 °C; overnight) and their corresponding Alexa conjugated antibodies (room temperature; 1 h). DAPI was used for nuclear staining. Immunostainings were examined using inverted epifluorescent microscope (Zeiss™ Axiovert 200 M).

4.4. SNP and CNV analysis

Single nucleotide polymorphism (SNP) analysis of hiPSC lines, compared to parental somatic cells, was used to confirm genome integrity. DNA was extracted from somatic and hiPSC samples (p23-MS573; p40-WT8288; p30-SD378M; p16-202CT) using the QIAGEN QiaAmp kit. 200 ng of gDNA were outsourced to Intergen Company

Table 2

Reagents details.

	Antibodies used for immunocytochemistry/flow-cytometry			
	Antibody	Dilution	Company Cat #	RRID
Pluripotency Markers Immunocytochemistry	Rat anti-OCT3/4	1:200	Thermo Fisher Scientific Cat# 14-5841-80	AB_914299
Pluripotency Markers Immunocytochemistry	Mouse anti-TRA-1-60	1:200	Thermo Fisher Scientific Cat# 14-8863-80	AB_891612
Pluripotency Markers Flow cytometry	SSEA-4-PE	5 µl/test	Thermo Fisher Scientific Cat# 12-8843-42	AB_11151520
Pluripotency Markers Flow cytometry	TRA-1-60-PE	5 µl/test	Thermo Fisher Scientific Cat# 12-8863-82	AB_891602
Nuclear staining	DAPI	1 µg/mL	Merck Cat# D9564	
Secondary antibodies	Goat Anti-Mouse IgG (H + L) Antibody, Alexa Fluor 568	1:1000	Molecular Probes Cat# A-11004	AB_2534072
Secondary antibodies	Conjugated Goat Anti-Rat IgG (H + L) Antibody, Alexa Fluor 488	1:1000	Molecular Probes Cat# A-11006	AB_141373
	Target	Size of band	Forward/Reverse primer (5'-3')	
Pluripotency Markers (TaqMan probe)	<i>NANOG</i>	109 bp	Thermo Fisher Scientific Cat# Hs02387400	
Pluripotency Markers (TaqMan probe)	<i>POU5F1</i>	77 bp	Thermo Fisher Scientific Cat# Hs04260367	
Pluripotency Markers (TaqMan probe)	<i>SOX2</i>	91 bp	Thermo Fisher Scientific Cat# Hs01053049	
Ectoderm Marker (TaqMan probe)	<i>PAX6</i>	86 bp	Thermo Fisher Scientific Cat# Hs01088114	
Endoderm Marker (TaqMan probe)	<i>SOX17</i>	149 bp	Thermo Fisher Scientific Cat# Hs00751752	
Mesoderm Marker (TaqMan probe)	<i>HAND1</i>	65 bp	Thermo Fisher Scientific Cat# Hs02330376	
House-Keeping Genes (TaqMan probe)	<i>ACTB</i>	63 bp	Thermo Fisher Scientific Cat# Hs01060665	
House-Keeping Genes (primer)	<i>GAPDH</i>	81 bp	Forward: 5'-AAT CCC ATC ACC ATC TTC CA-3' Reverse: 5'-TGG ACT CCA CGA CGT ACT CA-3'	
Sendai Virus Detection (primer)	SeV	181 bp	Forward: 5'-GGA TGA CTA GGT GAT ATC GAG C-3' Reverse: 5'-ACC AGA CAA GAG TTT AAG AGA TAT GTA TC-3'	

(Evry, France) for karyotype analysis using HumanCore-24-v1 SNP arrays. Analysis was performed with GenomeStudio software. Chromosome abnormalities were determined by visual inspection of LogR ratio values by comparing parental cells with hiPSC-derived samples. SNP data were used to compute CNV and validate cell line parentality.

4.5. Embryoid body formation

hiPSC colonies were detached (p31-MS573; p67-WT8288; p31-SD378M; p31-202CT) with dispase (STEMCELL™ Technologies), washed and replated in 6-well low-attachment plates (Costar®, Corning). The medium was changed the next day and unsupervised EB differentiation was started a day after, by incubating cells (37 °C, 5% CO₂, 21% O₂) in EB medium: 78% (v/v) KnockOut™ DMEM/F-12 (Gibco™) + 1% (v/v) GlutaMAX™ Supplement (Gibco™) + 20% (v/v) KnockOut™ Serum Replacement (Gibco™) + 1X MEM Non-Essential Amino Acids Solution + 0.1 mM 2-Mercaptoethanol (Gibco™). EB medium was changed every day during 21 days before RT-qPCR analysis.

4.6. RT-qPCR

Total RNA samples were isolated using the NucleoSpin RNA kit (MACHEREY-NAGEL) and their quality assessed by NanoDrop1000. One µg of total RNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems™). PCR amplification was performed using (1) FAM-labeled TaqMan probes (Applied Biosystems™) for pluripotency and differentiation potential, and (2) SYBR-green master mix and primers (Applied Biosystems™) for SeV loss (p25-MS573; p17-WT8288; p25-SD378M). Data were normalized to ACTB or GAPDH.

4.7. Mycoplasma detection

Mycoplasma contamination was assessed (p34-MS573; p83-WT8288; p17-SD378M; p34-202CT) with MycoAlert™ Mycoplasma Detection Kit (Lonza) and was interpreted as negative when < 0.9.

Declaration of Competing Interest

The authors declare that they have no known competing financial

interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the Cytometry core facility (Cyto-Cell) and the iPS core facility of Nantes University, for their technical support. We are most grateful to the Genomics and Bioinformatics Core Facility of Nantes (GenoBiRD, Biogenouest, IFB) for its technical support. We thank the biological resource centre for biobanking (CHU Nantes, Nantes Université, Centre de ressources biologiques (BB-0033-00040)), F-44000 Nantes, France). We thank Estelle Baron for her technical support. This work was funded by grants from Hôpitaux Universitaires du Grand Ouest and Fondation Maladies Rares (HUGO-FMR; RC17_0357 to J.B.G.), the INSERM cross-cutting program GOLD (Genomic variability in health and disease to R.R.), The French National Research Agency (ANR-12-BSC1-0013-01 to F.C. and ANR-15-CE14-0019-01 to N.G.), the Fondation Genavie, La Fédération Française de Cardiologie (to N.G.), and French National Research Project CHOPIN (Cholesterol Personalized Innovation) funded by the French National Research Agency (ANR-16-RHUS-0007 to B.C.) and coordinated by the CHU of Nantes. D.A. was funded by GOLD program.

References

- Belbachir, N., Portero, V., Al Sayed, Z.R., Gourraud, J.-B., Dilasser, F., Jesel, L., Guo, H., Wu, H., Gaborit, N., Guilluy, C., Girardeau, A., Bonnaud, S., Simonet, F., Karakachoff, M., Pattier, S., Scott, C., Burel, S., Marionneau, C., Chariou, C., Gaignerie, A., David, L., Genin, E., Deleuze, J.-F., Dina, C., Sauzeau, V., Loirand, G., Baro, I., Schott, J.-J., Probst, V., Wu, J.C., Redon, R., Charpentier, F., Le Scouarnec, S., 2019. RRAD mutation causes electrical and cytoskeletal defects in cardiomyocytes derived from a familial case of Brugada syndrome. *Eur Heart J.* 40 (37), 3081–3094.
- Volpato, V., Webber, C., 2020. Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. *Dis Model Mech.* 13 (1) <https://doi.org/10.1242/dmm.042317>.

Annexe 3

A consistent arrhythmogenic trait in Brugada syndrome cellular phenotype

Zeina R. Al Sayed, Mariam Jouni, Jean-Baptiste Gourraud, Nadjat Belbachir, Julien Barc, Aurore Girardeau, Virginie Forest, Aude Derevier, Anne Gaignerie, Caroline Chariou, Bastien Cimarosti, **Robin Canac**, Pierre Olchesqui, Eric Charpentier, Jean-Jacques Schott, Richard Redon, Isabelle Baró, Vincent Probst, Flavien Charpentier, Gildas Loussouarn, Kazem Zibara, Guillaume Lamirault, Patricia Lemarchand, Nathalie Gaborit

DOI : 10.1002/ctm2.413 ; PMID : 34185406

Publié dans *Clinical and translational medicine*, 2021

LETTER TO EDITOR

A consistent arrhythmogenic trait in Brugada syndrome cellular phenotype

To the Editor:

Brugada syndrome (BrS) is an inherited arrhythmic disease predisposing to sudden cardiac death (SCD), characterized by a typical electrocardiogram pattern that includes a J point elevation with a coved type ST segment.¹ BrS is a complex genetic disease in which ~20% of patients carry rare variants in *SCN5A* gene, whereas the others remain genetically unresolved.² Despite this genetic complexity, we hypothesize that a common cellular phenotypic trait is at the root of this specific BrS ECG pattern. In this study, we identified a phenotype that is common to human-induced pluripotent stem cell-derived ventricular cardiomyocytes (hiPSC-CMs) generated from six Brugada patients with different genetic backgrounds. Our results unmasked a cellular arrhythmogenic phenotype combining gene expression and electrical abnormalities, including an increase in late sodium current.

Six patients affected by type I BrS (BrS1-6; Figure S1; Tables S1 and S2) with a familial history of SCD or syncope were selected, among whom two carry *SCN5A* variants (marked with a + symbol). An additional individual, not affected by BrS (non-BrS), carrying the same *SCN5A* variant as BrS2+, was also recruited, as well as four control (Ctrl) subjects. Somatic cells from all studied subjects were reprogrammed into hiPSC lines and differentiated into cardiomyocytes (Figure 1).

Transcriptional expression profiling identified 133 differentially expressed genes in BrS hiPSC-CMs (Figure 2A). Gene set enrichment analyses showed that transcripts of transmembrane transporters and channels were significantly overrepresented (Figure 2B), including genes encoding sodium, calcium, and potassium channels (Figure 2C). High-throughput real-time RT-PCR³ on 96 genes related to cardiac electrical function (Table S3) identified 13 differentially expressed genes in BrS, in comparison to Ctrl and non-BrS hiPSC-CMs (Figure 2D). Importantly, the expression of *SCN5A*, the main BrS culprit gene identified to date,⁴ remained unchanged, excluding

SCN5A expression levels as a hallmark for BrS hiPSC-CM phenotype. Conversely, calcium and sodium transporters, playing important roles in membrane depolarization, were differentially expressed. Comparative analysis of hiPSC-CM electrophysiological functions investigated whether these modifications were a consistent trait of BrS phenotype at the cellular level.

Whereas decrease in sodium current is considered as the most frequently associated electrical alteration in BrS pathophysiology,^{5,6} protein expression of Na_v1.5, encoded by *SCN5A*, was decreased in only two BrS, and the non-BrS lines (Figure 2E). Concordantly, reduction in *I*_{Na} density was detected in these same lines (Figure 2F–H). This confirmed previous results, for BrS5+,⁷ and regarding BrS1+, which carries an *SCN5A* rare variant, the reduction was confirmed using conventional transfection in COS-7 cells of this variant (Figure S2). Furthermore, the steady-state activation and inactivation gating properties were not modified in BrS hiPSC-CMs (Figure S3A; Table S4). Therefore, *I*_{Na} reduction is not a common trait of BrS hiPSC-CMs and appears to be solely associated with the presence of variants affecting *SCN5A* expression or function.

Similarly, reduction in *I*_{Ca,L} channel protein expression and current density were not a common trait of BrS hiPSC-CMs (Figure 2I–L, Figure S3B; Table S4).

Global cellular electrophysiological phenotype was then evaluated with action potential (AP) recordings, but no AP basal parameters specifically segregated BrS hiPSC-CMs, and spontaneous beating frequencies did not differ between all cell lines (Figure S4). Noteworthy, ventricular-like AP analysis revealed an arrhythmic phenotype present mostly in BrS hiPSC-CMs, irrespective of their genetic background (Figure 3A). Early afterdepolarizations (EADs) were observed in 39–70% of all six BrS ventricular-like hiPSC-CMs versus only in 4% and 4.7% of Ctrl and non-BrS hiPSC-CMs, respectively (Figure 3B, Figure S5). Thereby, the high EAD occurrence in ventricular-like hiPSC-CMs was associated with the presence of a BrS

This is an open access article under the terms of the [Creative Commons Attribution License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Clinical and Translational Medicine* published by John Wiley & Sons Australia, Ltd on behalf of Shanghai Institute of Clinical Bioinformatics

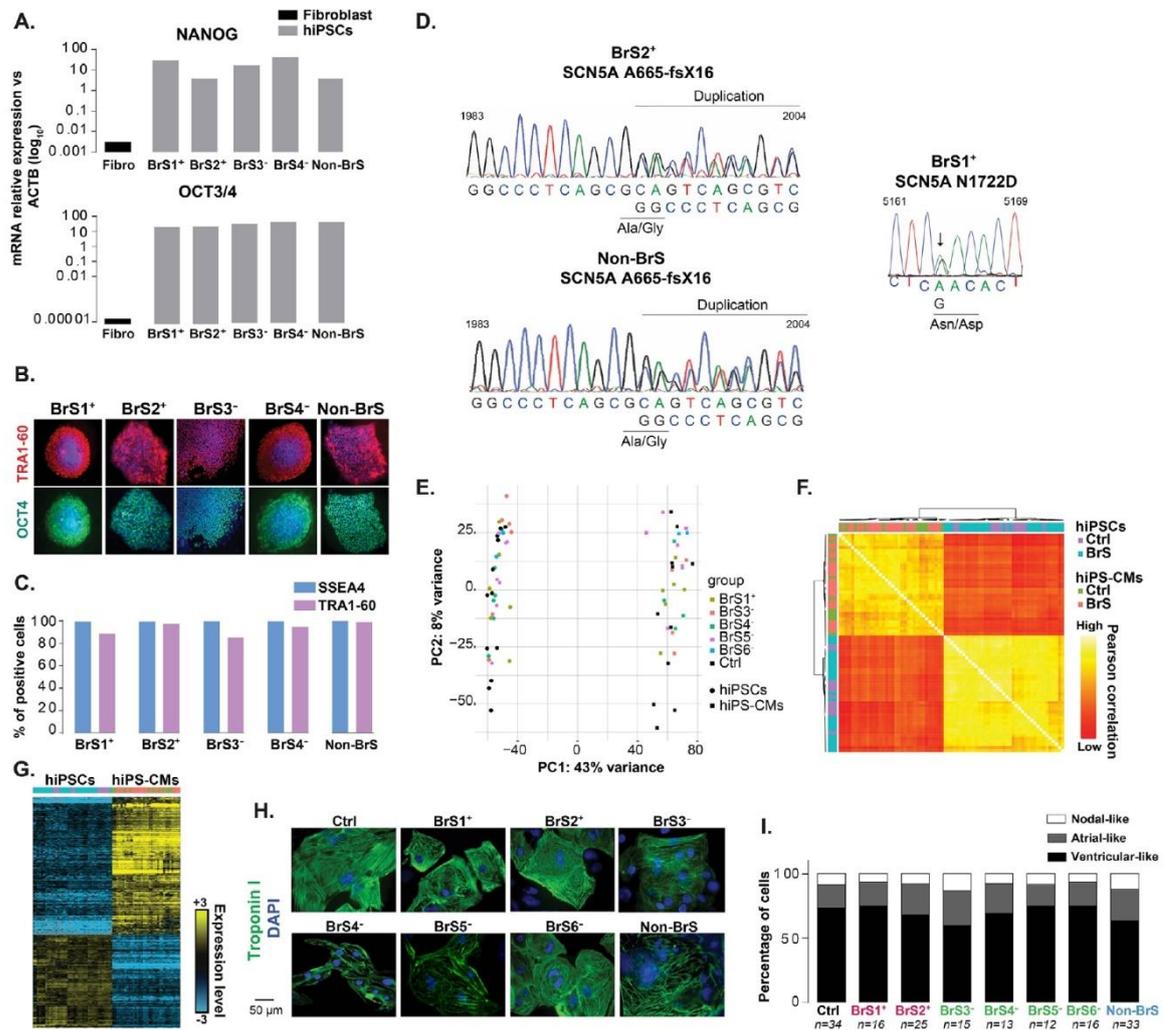


FIGURE 1 Pluripotency and *SCN5A* variant validation in hiPSCs, and characterization of derived cardiomyocytes. (A) Transcript expression of pluripotency markers: NANOG and OCT3/4 in newly described hiPSCs as compared to fibroblasts (Fibro). (B) Representative immunostainings of TRA1-60 (red) and OCT4 (green) in hiPSCs. (C) Percentage of hiPSCs expressing SSEA4 and TRA1-60 evaluated by flow cytometry. (D) Genomic sequence chromatograms validating (right) the 5164A>G *SCN5A* variant carried by BrS1+ and (left) the *SCN5A* 1983-1993 duplication carried by BrS2+ and non-BrS, in the corresponding hiPSCs. (E) Principal component analysis (PCA) of 39 hiPSC samples and their corresponding differentiated hiPSC-CMs, based on their expression pattern of 27106 analyzed transcripts (3'SRP data). All clones of each hiPSC line are highlighted. (F) Correlation matrix of hiPSCs and hiPSC-CMs expression profiles. Yellow and orange indicate high and low correlation, respectively. Samples were clustered using an ascending hierarchical method with Pearson as metric and ward D2 linkage. (G) Heatmap showing expression levels of 9661 differentially expressed genes between hiPSCs and hiPSC-CMs (same samples as in A). Genes were clustered using a hierarchical ascending method with an uncentered correlation metric and complete linkage. Yellow and blue indicate high and low levels, respectively. (H) Illustrative immunostainings of Troponin I (green) in hiPSC-CMs. Nuclei were stained with DAPI (blue). (I) Percentages of nodal-like, atrial-like, and ventricular-like cells classified based on the analysis of spontaneous action potential recordings

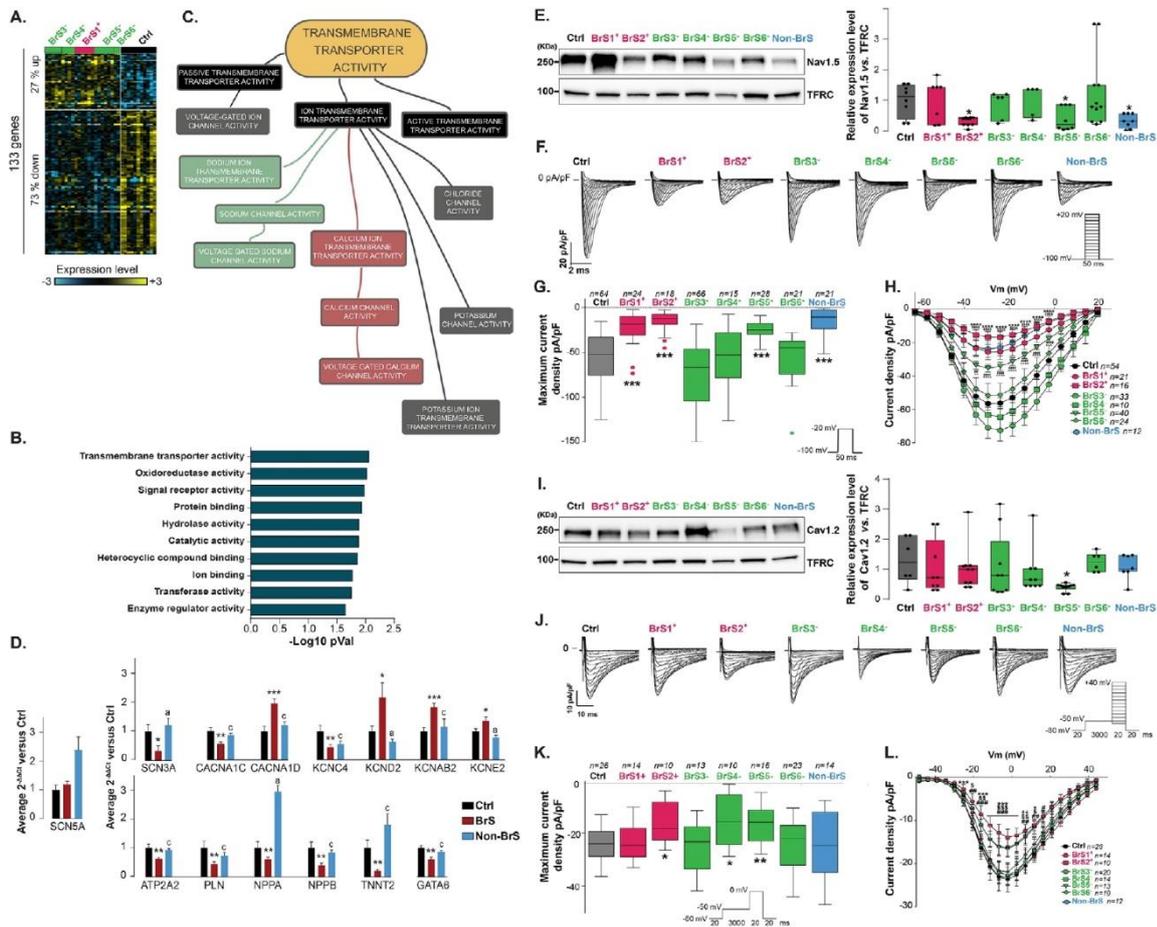


FIGURE 2 Differential gene expression profiles and variations in I_{Na} and I_{CaL} in BrS hiPSC-CMs as compared to controls. (A) Heatmap showing hierarchical clustering of expression profiles of 133 differentially expressed genes obtained by 3'SRP in control (Ctrl) and BrS hiPSC-CMs at day 28 of differentiation. A total of 27% were upregulated, whereas 73% genes were downregulated in BrS hiPSC-CMs. Yellow and blue represent high and low expression levels, respectively. All clones of each hiPSC line are highlighted. (B) Gene set enrichment analysis (GSEA) of gene variations obtained by 3'SRP shows gene sets with statistically altered expression patterns. (C) MindMap describing the transmembrane transporter activity alterations. (D) Expression levels of differentially expressed genes identified using high-throughput TaqMan (TLDA) in BrS hiPSC-CMs ($n = 14$), compared to control hiPSC-CMs ($n = 12$), and in non-BrS hiPSC-CMs ($n = 4$) versus BrS hiPSC-CMs. p -values: *, **, and *** or a, b, and c: $p < .05$, $p < .01$, and $p < .001$ versus Ctrl or BrS, respectively (t -test). (E) Representative immunoblots for $Na_v1.5$ and transferrin receptor (TFRC) in hiPSC-CMs (left panel). Ratios of $Na_v1.5$ expression levels (right panel, Tukey plot, $n = 8$). * $p < .05$ versus control (Mann-Whitney test). $Na_v1.5$ decreases in hiPSC-CMs from three subjects, BrS2+ and non-BrS (both carrying a stop codon in *SCN5A*), as well as BrS5-, harboring *RRAD* variant was observed. (F) Representative superimposed I_{Na} densities (inset: voltage-clamp protocol). Reduction was detected in BrS2+, BrS5-, and non-BrS, as well as in BrS1+ hiPSC-CMs carrying the N1722D-*SCN5A* rare variant. (G) Peak I_{Na} densities measured in control (Ctrl), BrS, and the non-affected carrier of *SCN5A* mutation (non-BrS) hiPSC-CMs determined at -20 mV (Tukey plot). *** $p < .001$ versus control (Mann-Whitney test). (H) Mean peak I_{Na} densities (pA/pF) versus membrane potential (V_m) recorded in hiPSC-CMs. ****, \$\$\$, ###, and **** $p < .0001$ versus control for BrS1+, BrS2+, BrS5-, and non-BrS, respectively (two-way ANOVA with Bonferroni post hoc test). (I) Representative immunoblots for $Ca_v1.2$, the main pore-forming subunit of the cardiac L-type calcium channel, and transferrin receptor (TFRC) in hiPSC-CMs (left panel). Ratios of $Ca_v1.2$ expression levels (right panel, Tukey plot, $n = 8$). A decrease in $Ca_v1.2$ expression was solely observed in BrS5- hiPSC-CMs, carrying an *RRAD*-variant. * $p < .05$ versus control (Mann-Whitney test). (J) Representative superimposed I_{CaL} densities (inset: voltage protocol). (K) Peak I_{CaL} densities measured in control (Ctrl), BrS, and the nonaffected carrier of *SCN5A* mutation (non-BrS) hiPSC-CMs determined at 0 mV (Tukey plot). A decrease in I_{CaL} was observed in BrS2+, BrS4- and, consistently with a previous description, in BrS5-. * $p < .05$ and ** $p < .01$ versus control (Mann-Whitney test). (L) Mean peak I_{CaL} densities (pA/pF) versus membrane potential (V_m) recorded in hiPSC-CMs. *, #, and \$ $p < .05$, **, ##, and \$\$ $p < .01$, and ***, ###, and \$\$\$ $p < .001$ versus control for BrS1+, BrS4-, and BrS5-, respectively (two-way ANOVA with Bonferroni post hoc test)

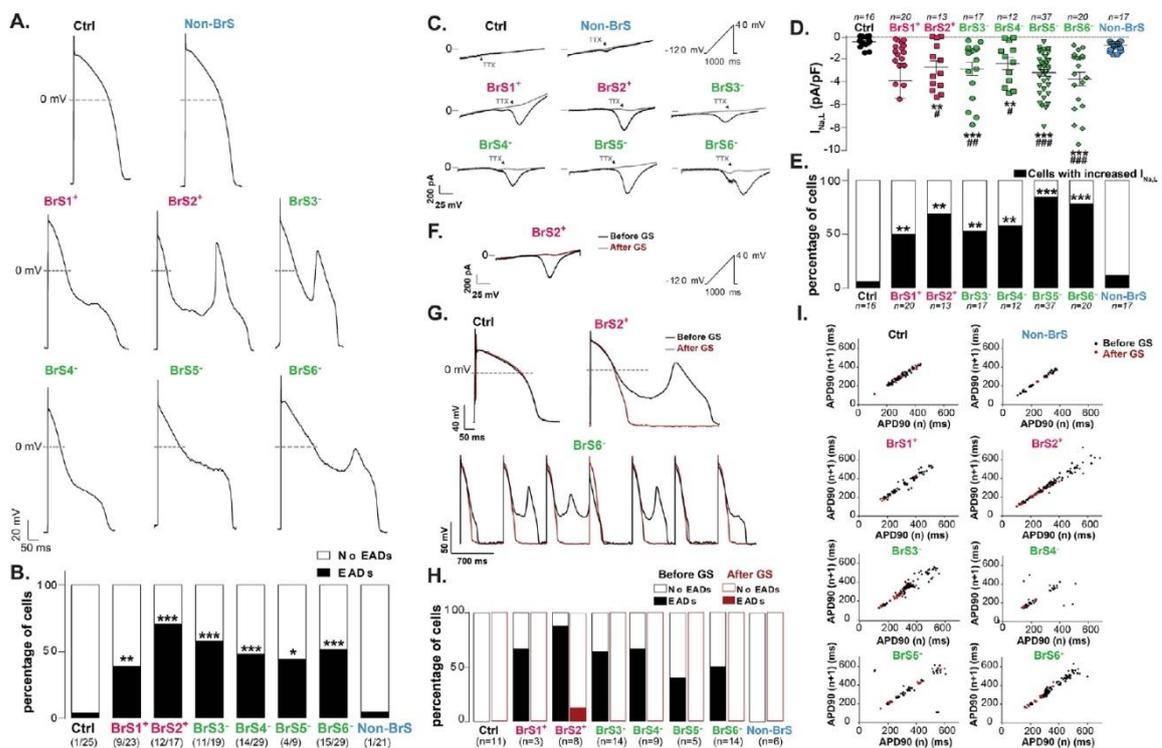


FIGURE 3 Increased early after depolarization (EAD) occurrence in all BrS ventricular-like hiPSC-CM lines, linked to an increase in late sodium current. (A) Representative AP recordings, showing EADs in BrS lines only. Representative ventricular-like AP when paced at 700 ms cycle length and when artificial I_{K1} was injected (dynamic current clamp). APs are defined as ventricular-like when $(APD_{30} - APD_{40}) / (APD_{70} - APD_{80}) > 1.45$. (B) Percentage of ventricular-like hiPSC-CMs presenting at least 1 EAD, irrespective of the current clamp conditions. $*p < .05$, $**p < .01$, and $***p < .001$ versus control (Fisher's exact test). (C) Representative $I_{Na,L}$ recordings from hiPSC-CMs, before (black) and after (grey) TTX application (inset: voltage protocol). (D) $I_{Na,L}$ (TTX-sensitive current) densities at -10 mV. $*p < .05$ and $***p < .001$ versus Ctrl (Mann-Whitney test), and $##p < .01$ and $###p < .001$ versus non-BrS (Mann-Whitney test). (E) Percentage of cells presenting $I_{Na,L}$ density greater than the 97th percentile value of $I_{Na,L}$ in the Ctrl hiPSC-CMs. $**p < .01$ and $***p < .001$ versus control (Fisher's exact test). Indeed, an increase in $I_{Na,L}$ density was defined by values higher than 97th percentile of the Ctrl hiPSC-CMs. (F) Representative example of $I_{Na,L}$ current recorded in BrS hiPSC-CMs before (black) and after (red) application with GS-458967 (300 nM), a specific $I_{Na,L}$ inhibitor (inset: voltage protocol). (G) Representative AP recordings from control and BrS hiPSC-CMs obtained before and after GS-458967 application. (H) Percentage of cells with EADs before and after GS-458967 application. (I) Poincaré plots showing APD_{90} of each AP ($n + 1$) versus APD_{90} of its preceding one, before and after GS-458967 application

phenotype in the investigated cell lines, but not with the presence of a variant in *SCN5A*.

The occurrence of EADs may be linked to an abnormally high density of depolarizing late sodium current ($I_{Na,L}$) during APs repolarizing phase.⁸ Accordingly, BrS hiPSC-CMs presented with a higher density of $I_{Na,L}$ as compared to Ctrl and non-BrS hiPSC-CMs (Figure 3C,D). Moreover, an increase in $I_{Na,L}$ density was observed only in 6% and 12% of Ctrl and non-BrS hiPSC-CMs respectively, in accordance with their low EAD occurrence, whereas increased $I_{Na,L}$ density was present in 50–85% of all BrS ventricular-like hiPSC-CMs, reminiscent of the high EAD occurrence (Figure 3B,E). We then superfused ventricular-like BrS

hiPSC-CMs during AP recording with GS-458967 (6-(4-(trifluoromethoxy)phenyl)-3-(trifluoromethyl)-[1,2,4] triazolo[4,3-a]pyridine, which selectively blocks late sodium current),⁹ causing full inhibition of $I_{Na,L}$ (Figure 3F), and found abolishment of EADs (Figure 3G,H) and reduced APD_{90} dispersion (Figure 3I). Altogether, these data strongly suggested that the abnormal increase of $I_{Na,L}$ in BrS hiPSC-CMs is responsible for EADs.

Further strengthening the role of $I_{Na,L}$ in the electrical cellular phenotype of BrS, when each ECG parameter was tested for its correlation with either $I_{Na,L}$ or I_{Na} measured densities, only $I_{Na,L}$ density correlated significantly with one sole parameter, that is, the J point elevation (Table S5).

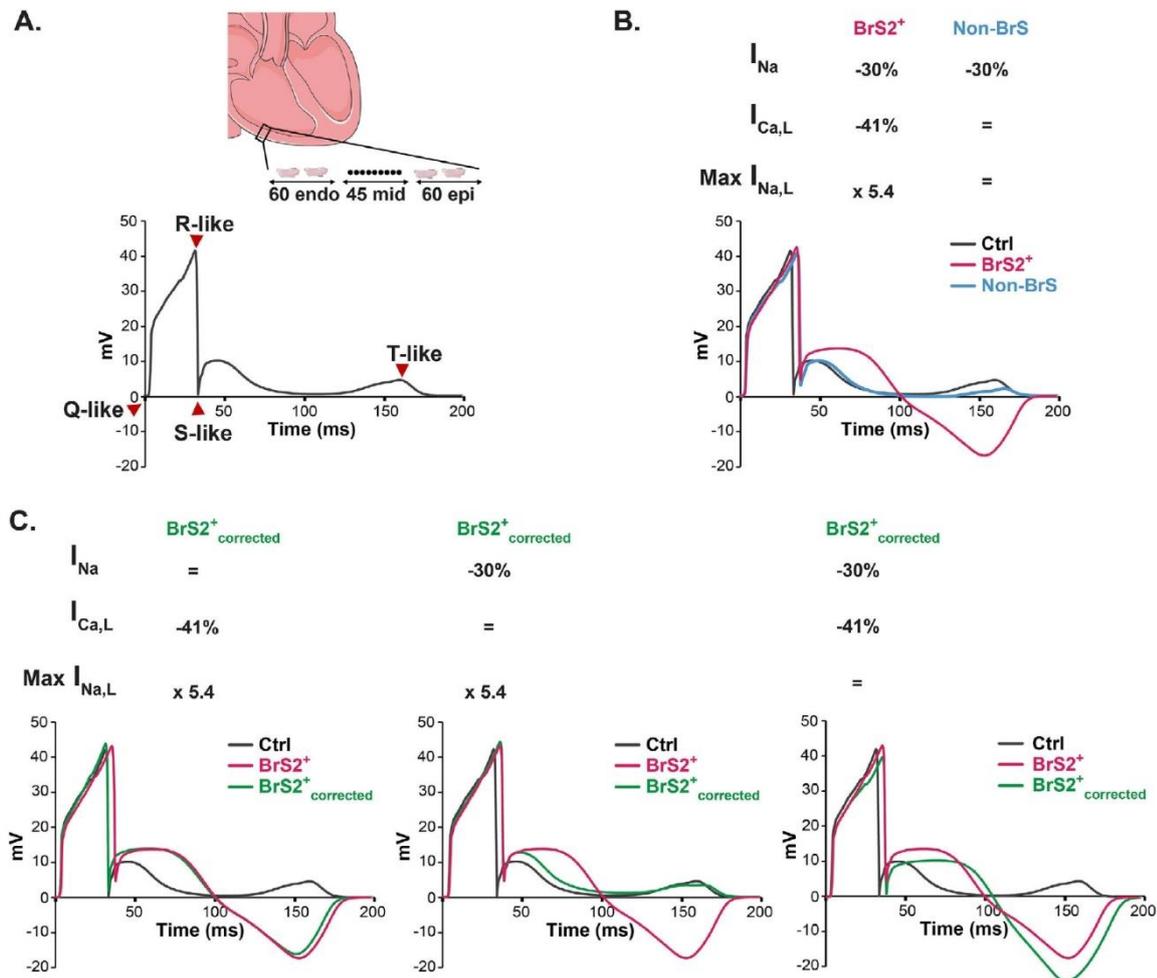


FIGURE 4 Applying depolarizing ion current alterations from BrS hiPSC-CMs on an electrogram model to mimic BrS patient's ECG features. (A) Top: Right ventricle electrogram model simulates the global electrical activity of a transmural wedge comprising 60 subendocardial, 45 midmyocardial, and 60 subepicardial human ventricular cells. Bottom: Representative electrogram showing the Q-like, R-like, S-like and T-like waves. (B) Ventricular transmural electrogram mathematical model of Ctrl (black), BrS2⁺ (pink), and non-BrS (blue) illustrated based on hiPSC-CMs data of the relative variation in ion currents (I_{Na} , $I_{Ca,L}$, and $I_{Na,L}$) mean amplitude as compared to Ctrl lines. In accordance with patient's ECGs, applying BrS2⁺ ionic current changes prolonged the QRS-like complex, elevated and widened the ST-like segment, and inverted the T-like wave; and applying I_{Na} change identified in non-BrS hiPSC-CMs only prolonged the QRS-like complex, similar to non-BrS PCCD ECG. (C) Each of the currents I_{Na} , $I_{Ca,L}$, and $I_{Na,L}$ (from left to right) were sequentially corrected and the resulting electrograms are illustrated in green. Ctrl and BrS2⁺ electrograms are presented in black and pink, respectively

To challenge the pathophysiological relevance of the ion current alterations identified in non-BrS and BrS2⁺ hiPSC-CMs, we applied them to a mathematical human electrogram model that allows visualizing transmural-like electrogram with a QRS-like complex, a ST-like segment, and a T-like wave (Figure 4A).¹⁰ First, in accordance with BrS2⁺ patient's ECG, applying the alterations observed in peak I_{Na} , $I_{Ca,L}$, and in $I_{Na,L}$ in BrS2⁺ hiPSC-CMs was sufficient to induce prolongation of the QRS-like complex, ST-like seg-

ment elevation and widening, and T-like wave inversion (Figure 4B). Then, sequential correction of each altered current in BrS2⁺ hiPSC-CMs was made (BrS2⁺ corrected). Correction of I_{Na} density led to QRS-like complex normalization; correction of $I_{Ca,L}$ density shortened duration of the ST-like segment elevation and normalized the T-like wave orientation; and correction of $I_{Na,L}$ density led to reduction of ST-like segment amplitude toward normalization (Figure 4C, left to right). Overall, these results strongly

suggest that depolarizing current alterations can impact a multicellular electrogram model, mimicking BrS ECG phenotype.

In conclusion, in the present study, a particular cellular electrophysiological phenotype common to six out of six BrS hiPSC-CM lines with various genetic backgrounds has been unveiled. We showed that high EAD occurrence associates with an abnormal increase of $I_{Na,L}$ in all investigated BrS cell lines, and correlates with the corresponding patients' J point elevation on ECG. We focused on the ventricular cell type, at a single-cell level. Implementation of emerging phenotypic technologies, such as single-cell transcriptomics and cardiac tissue engineering, will allow investigation of the potential involvement of other cardiac cell types in the disease phenotype and the role of specific cell-to-cell interactions. Altogether, the obtained results open perspectives to better understand the ventricular arrhythmia occurrence in BrS and to identify a dedicated therapeutic approach to prevent the risk of SCD.

ACKNOWLEDGMENTS

The authors thank Dr. Connie Bezzina and Dr. Isabella Mengarelli for the gift of hiPSCs from BrS6-patient, Dr. Pierre Lindenbaum and Dr. Stephanie Bonnaud for the Haloplex targeted capture and NGS experiments, and Adeline Goudal for her support in the variant annotation. Genomic and bioinformatics analysis, flow cytometry, and iPSCs derivation were performed with the support of GenoBiRD (Biogenouest), CytoCell, and iPSC core facility of Nantes University, respectively. Finally, the authors are grateful to the patients and families who agreed to participate in our research. This work was supported by grants from the Fondation pour la Recherche Médicale (DEQ20140329545), National Research Agency ANR-14-CE10-0001-01 and La Fédération Française de Cardiologie. Dr. Nathalie Gaborit was laureate of grants from Fondation Lefoulon-Delalande and Marie Curie Actions, International Incoming Fellowship FP7-PEOPLE-2012-IIF (PIIF-GA-2012-331436) and from the National Research Agency (HEART-iPS ANR-15-CE14-0019-01). Dr. Zeina R. Al-Sayed was supported by scholarships from the Lebanese University, Eiffel Program of Excellence (Campus France), and Fondation Genavie. Dr. Mariam Jouni was funded by a scholarship from the Association of Scientific Orientation and Specialization (ASOS) and by a grant from the Lebanese University to Dr. Kazem Zibara. Eric Charpentier was supported by Data-Santé (Région Pays de la Loire). Dr. Barc was supported by H2020-MSCA-IF-2014.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

In accordance with the "DFG Guidelines on the Handling of Research Data," the authors declare that all data supporting the findings of this study are available within the article and its supporting information files or from the corresponding author upon reasonable request. The dataset will be archived for at least 10 years after publication.

FUNDING INFORMATION

Fondation pour la Recherche Médicale, Grant Number: DEQ20140329545; National Research Agency: Grant Number: ANR-14-CE10-0001-01 and HEART-iPS Grant Number: ANR-15-CE14-0019-01; La Fédération Française de Cardiologie; Fondation Lefoulon-Delalande; Marie Curie Actions: International Incoming Fellowship, Grant Number: FP7-PEOPLE-2012-IIF (PIIF-GA-2012-331436) and H2020-MSCA-IF-2014 (RISTRAD-661617); Eiffel Program of Excellence; Fondation Genavie; Association of Scientific Orientation and Specialization; Lebanese University; Data-Santé (Région Pays de la Loire)

Zeina R. Al-Sayed¹ 

Mariam Jouni¹

Jean-Baptiste Gourraud^{1,2} 

Nadjet Belbachir¹

Julien Barc¹ 

Aurore Girardeau¹

Virginie Forest¹ 

Aude Derevier³

Anne Gaignerie³

Caroline Chariau³

Bastien Cimarosti¹

Robin Canac¹

Pierre Olchesqui¹ 

Eric Charpentier¹ 

Jean-Jacques Schott^{1,2} 

Richard Redon^{1,2} 

Isabelle Baró¹ 

Vincent Probst^{1,2} 

Flavien Charpentier^{1,2} 

Gildas Loussouarn¹ 

Kazem Zibara⁴ 

Guillaume Lamirault^{1,2} 

Patricia Lemarchand^{1,2} 

Nathalie Gaborit¹ 

¹ *l'institut du thorax, Inserm, CNRS, UNIV Nantes, Nantes, France*

² *l'institut du thorax, CHU Nantes, Nantes, France*

³ *Nantes Université, CHU Nantes, Inserm, CNRS, SFR Santé, Nantes, France*

⁴ *Laboratory of Stem Cells, PRASE, Biology Department, Faculty of Sciences, Lebanese University, Beirut, Lebanon*

Correspondence

Nathalie Gaborit and Patricia Lemarchand, l'institut du thorax, Inserm UMR 1087, CNRS UMR 6291, IRS-UN, 8 quai Moncousu, 44007 Nantes cedex 1, France.

Email: nathalie.gaborit@univ-nantes.fr;
patricia.lemarchand@univ-nantes.fr

ORCID

Zeina R. Al Sayed  <https://orcid.org/0000-0002-0661-2003>

Jean-Baptiste Gourraud  <https://orcid.org/0000-0002-6961-2131>

Julien Barc  <https://orcid.org/0000-0003-4106-5946>

Virginie Forest  <https://orcid.org/0000-0003-2774-5232>

Pierre Olchesqui  <https://orcid.org/0000-0002-4336-2218>

Eric Charpentier  <https://orcid.org/0000-0002-8571-7603>

Jean-Jacques Schott  <https://orcid.org/0000-0002-9578-9475>

Richard Redon  <https://orcid.org/0000-0001-7751-2280>

Isabelle Baró  <https://orcid.org/0000-0003-4850-4171>

Vincent Probst  <https://orcid.org/0000-0002-5492-8619>

Flavien Charpentier  <https://orcid.org/0000-0002-5057-0998>

Gildas Loussouarn  <https://orcid.org/0000-0001-8007-5931>

Kazem Zibara  <https://orcid.org/0000-0002-9887-072X>

Guillaume Lamirault  <https://orcid.org/0000-0001-6490-1279>

Patricia Lemarchand  <https://orcid.org/0000-0002-5330-2008>

Nathalie Gaborit  <https://orcid.org/0000-0002-5231-6555>

REFERENCES

1. Priori SG, Wilde AA, Horie M, et al. HRS/EHRA/APHRS expert consensus statement on the diagnosis and management of

patients with inherited primary arrhythmia syndromes. *Heart Rhythm*. 2013;10:1932-1963.

2. Gourraud J-B, Barc J, Thollet A, et al. The Brugada syndrome: a rare arrhythmia disorder with complex inheritance. *Front Cardiovasc Med*. 2016;3:9.
3. Al Sayed ZR, Canac R, Cimarosti B. Human model of IRX5 mutations reveals key role for this transcription factor in ventricular conduction. *Cardiovasc Res*. 2020;8:259.
4. Kapplinger JD, Tester DJ, Alders M, et al. An international compendium of mutations in the SCN5A-encoded cardiac sodium channel in patients referred for Brugada syndrome genetic testing. *Heart Rhythm*. 2010;7:33-46.
5. Yan GX, Antzelevitch C. Cellular basis for the Brugada syndrome and other mechanisms of arrhythmogenesis associated with ST-segment elevation. *Circulation*. 1999;100:1660-1666.
6. Meregalli PG, Wilde AA, Tan HL. Pathophysiological mechanisms of Brugada syndrome: depolarization disorder, repolarization disorder, or more? *Cardiovasc Res*. 2005;67:367-378.
7. Belbachir N, Portero V, Al Sayed ZR, et al. RRAD mutation causes electrical and cytoskeletal defects in cardiomyocytes derived from a familial case of Brugada syndrome. *Eur Heart J*. 2019;40:3081-3094.
8. Shryock JC, Song Y, Rajamani S, et al. The arrhythmogenic consequences of increasing late INa in the cardiomyocyte. *Cardiovasc Res*. 2013;99:600-611.
9. Potet F, Egecioglu DE, Burridge PW. GS-967 and eleclazine block sodium channels in human induced pluripotent stem cell-derived cardiomyocytes. *Mol Pharmacol*. 2020;98:540-547.
10. Gima K, Rudy Y. Ionic current basis of electrocardiographic waveforms: a model study. *Circ Res*. 2002;90:889-896.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Annexe 4

Generation of human induced pluripotent stem cell lines from three patients affected by Catecholaminergic Polymorphic ventricular tachycardia (CPVT) carrying heterozygous mutations in RYR2 gene

Bastien Cimarosti[#], **Robin Canac**[#], Stephan De Waard[#], Aurore Girardeau, Anne Gaignerie, Aude Derevier, Virginie Forest, Michel Ronjat, Hervé Le Marec, Jean-Baptiste Gourraud, Patricia Lemarchand, Michel De Waard, Guillaume Lamirault*, Nathalie Gaborit*

: Co-premiers auteurs

* : Co-derniers auteurs

DOI: 10.1016/j.scr.2022.102688 ; PMID : 35101670

Publié dans *Stem Cell Research*, 2022



Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Lab Resource: Multiple Cell Lines



Generation of human induced pluripotent stem cell lines from three patients affected by Catecholaminergic Polymorphic ventricular tachycardia (CPVT) carrying heterozygous mutations in *RYR2* gene

Bastien Cimarosti ^{a,1}, Robin Canac ^{a,1}, Stephan De Waard ^{a,b,1}, Aurore Girardeau ^a, Anne Gaignerie ^c, Aude Derevier ^c, Virginie Forest ^a, Michel Ronjat ^{a,b}, Hervé Le Marec ^a, Jean-Baptiste Gourraud ^a, Patricia Lemarchand ^a, Michel De Waard ^{a,b}, Guillaume Lamirault ^{a,2}, Nathalie Gaborit ^{a,2,*}

^a Université de Nantes, CHU Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France

^b LabEx « Ion Channels, Science & Therapeutics », Nice, France

^c Université de Nantes, CHU Nantes, INSERM, CNRS, SFR Santé, FED 4203, INSERM UMS 016, CNRS UMS 3556, F-44000 Nantes, France

ABSTRACT

Catecholaminergic Polymorphic Ventricular Tachycardia (CPVT) is an exercise and emotional stress-induced life-threatening inherited heart rhythm disorder, characterized by an abnormal cellular calcium homeostasis. Most reported cases have been linked to mutations in the gene encoding the type 2 ryanodine receptor gene, *RYR2*. We generated induced pluripotent stem cells (hiPSCs) from peripheral blood mononuclear cells (PBMC) from three CPVT-affected patients, two of them carrying p.R4959Q mutation and one carrying p.Y2476D mutation. These generated hiPSC lines are a useful model to study pathophysiological consequences of *RYR2* dysfunction in humans and the molecular basis of CPVT.

1. Resource table

Unique stem cell lines identifier	ITXi009-A ITXi010-A ITXi011-A
Alternative name(s) of stem cell lines	CPVT033 (ITXi011-A) CPVT068 (ITXi009-A) CPVT249 (ITXi010-A)
Institution	L'institut du thorax (INSERM UMR1087, CNRS UMR6291)
Contact information of distributor	Nathalie Gaborit, nathalie.gaborit@univ-nantes.fr
Type of cell lines	iPSC
Origin	Human
Additional origin info required	CPVT033 (ITXi011-A) Age: 46 Sex: female Ethnicity if known: North-west European CPVT068 (ITXi009-A) Age: 27 Sex: female

(continued on next column)

(continued)

Unique stem cell lines identifier	ITXi009-A ITXi010-A ITXi011-A
Ethnicity if known:	North-west European
CPVT249 (ITXi010-A)	
Age:	48
Sex:	female
Ethnicity if known:	North-west European
Cell Source	Peripheral blood mononuclear cells
Clonality	Clonal
Method of reprogramming	Sendai virus
Genetic Modification	YES
Type of Genetic Modification	Spontaneous mutations
Evidence of the reprogramming transgene loss	RT-qPCR
Associated disease	Catecholaminergic Polymorphic Ventricular Tachycardia type 1 (OMIM Entry - # 604772)
Gene/locus	

(continued on next page)

* Corresponding author at: l'institut du thorax Inserm UMR 1087, CNRS UMR 6291, IRS-UN, 8 quai Moncoussu, 44007 Nantes cedex 1, France.

E-mail address: nathalie.gaborit@univ-nantes.fr (N. Gaborit).

¹ Contributed equally to this work

² Contributed equally to this work

<https://doi.org/10.1016/j.scr.2022.102688>

Received 24 November 2021; Received in revised form 18 January 2022; Accepted 23 January 2022

Available online 27 January 2022

1873-5061/© 2022 The Authors.

Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

(continued)

Unique stem cell lines identifier	ITXi009-A ITXi010-A ITXi011-A
Date archived/stock date	CPVT033 and CPVT068: RYR2 p.Arg4959Glu, c.14876G > A CPVT249: RYR2 p.Tyr2476Asp, c.7426 T > G CPVT033: 07-13-2016 CPVT068: 07-21-2016 CPVT249: 09-13-2016
Cell line repository/bank	CPVT033: https://hpscereg.eu/cell-line/ITXi011-A CPVT068: https://hpscereg.eu/cell-line/ITXi009-A CPVT249: https://hpscereg.eu/cell-line/ITXi010-A
Ethical approval	For all patients, a signed informed consent has been received. Authorization MESR n°DC-2011-1399 and Authorization CPP Ouest II n° 2010-33 have been obtained from competent ethics review board and approval authority (French Health Ministry and French ethical committee) allowing the generation of these lines.

2. Resource utility

The hiPSC lines generated from patients affected by CPVT represent an inexhaustible *in vitro* source of human cardiac cells and therefore an excellent tool to model and better understand the mechanistic bases of CPVT, as well as for drug screening investigations.

3. Resource details

CPVT is an inherited cardiac disease exposing afflicted subjects to high risk of sudden cardiac death (SCD). While patients generally display normal heart structure and electrocardiogram at rest, life-threatening cardiac arrhythmias manifest themselves under exercise or emotional stress (Swan et al., 1999). These arrhythmias are characterized by premature ventricular complexes and polymorphic or bidirectional ventricular tachycardia. More than half of CPVT cases have been linked to mutations located in the gene encoding ryanodine receptor type 2 (*RYR2*) (Priori et al., 2002). *RYR2* is expressed at the sarcoplasmic reticulum membrane and mediates calcium release from its stocks into the cytosol, thus majorly contributing to the cardiac excitation–contraction coupling. In order to provide an inexhaustible source of cardiomyocytes carrying CPVT mutations in *RYR2*, three patients with CPVT were selected: Patients (1) CPVT033 and (2) CPVT068 both harbour a heterozygous c.14876G > A (p.R4959Q) mutation in the region encoding for the C-terminal domain of *RYR2* (hotspot IV for reported CPVT mutations (Xiao et al., 2016); Supplemental Fig. 1A-B). They belong to the same family whose clinical history has previously been published (Allouis et al., 2005; Supplemental Fig. 1C). Both patients were asymptomatic at rest but first degree relatives were victim of unexplained SCD. Exercise tests revealed bidirectional ventricular tachycardia without recurrences under beta blocker therapy. Patient (3) CPVT249 was diagnosed after the occurrence of an unexplained syncope. An implantable cardiac defibrillator was implanted without recurrence of sustained ventricular tachycardia under beta blocker therapy. No family history was found for this particular patient (Supplemental Fig. 1C). In line with this observation, the patient was shown to be affected by a de novo heterozygous c.7426 T > G *RYR2* mutation (p.Y2476D) in a sequence that encodes the HD1 domain (hotspot II for CPVT mutations) of *RYR2* (Supplemental Fig. 1A-B). This mutation, has not been previously described.

hiPSCs harbouring normal morphology were reprogrammed from

peripheral blood mononuclear cells (PBMCs) using the Sendai virus method and maintained on matrigel-coated plates with StemMACS iPS-Brew XF medium (Fig. 1D). Cultures were free of mycoplasma contamination (Supplemental Fig. 1D). Genomic sequencing of the corresponding PBMCs and hiPSCs, compared to unrelated control, confirmed the heterozygous presence of *RYR2* c.14876G > A in CPVT033 and CPVT068, and of *RYR2* c.7426 T > G in CPVT249 (Fig. 1A-C), also confirming the purity of the cell lines. Endogenous expression of OCT3/4 and TRA-1–60 proteins were visualized by immunocytochemistry (Fig. 1E). Expression upregulation of the endogenous pluripotent stem cell marker genes *SOX2*, *NANOG* and *POU5F1* was validated by qRT-PCR (Fig. 1F), and the loss of Sendai virus (SeV) was verified (Fig. 1G). Cytometry analyses showed that > 95 % of cells of each hiPSC lines were positive for the expression of pluripotency markers TRA-1–60 and SSEA4 (Fig. 1H and Supplemental Fig. 1F). The ability of each cell line to differentiate into derivatives of the three germ layers was demonstrated through embryoid body differentiation (Fig. 1I) and cardiac differentiation (Fig. 1J). Molecular karyotyping by SNP analysis revealed normal female karyotypes with no aneuploidies detected (Fig. 1K). The complete characterizations are summarized in Tables 1 and 2.

4. Materials and methods

4.1. Generation and maintenance of hiPSC lines

Blood cells were reprogrammed by the iPSC core facility (PF-iPSC, Nantes) using Sendai virus method (Cytotune reprogramming Kit, Life technologies). hiPSC lines were maintained at 37 °C, 5 %CO₂, 21 %O₂ in StemMACS™ iPS-Brew XF Medium (Miltenyi Biotec) on Matrigel®-coated plates (hESC-Qualified Matrix 0.05 mg/mL, Corning). At 75 % confluency, cells were passaged using Gentle Cell Dissociation Reagent (STEMCELL™ Technologies).

4.2. Genomic DNA extraction

NucleoSpin Tissue kit (MACHEREY-NAGEL) was used for PBMCs and hiPSCs (p33-CPVT033; p33-CPVT068; p33-CPVT249). Regions containing *RYR2* mutations were PCR-amplified using specific primers and sequenced (GenoBIRD, Nantes; Eurofins).

4.3. Flow cytometry analyses

A total of 1x10⁵ cells in suspension (p23-CPVT033; p21-CPVT068; p23-CPVT249), were washed with PBS/BSA 0.1 % and incubated with PE-labelled antibodies for 30 min, 4 °C, in the dark. Data acquisition was performed using FACSDiva™ software with LSR II instrument (Becton Dickinson, BD).

4.4. Immunofluorescence

Cells were dissociated and seeded onto 8-well IBIDI Matrigel-coated plates (p27-CPVT033; p26-CPVT068; p20-CPVT249). Cultured cells were fixed with 4 % paraformaldehyde for 15 min, permeabilized with 0.5 % Triton-X100 for 15 min and blocked with 3 % BSA for 30 min. Immunofluorescence stainings were performed using appropriate primary antibodies (4 °C; overnight) and their corresponding Alexa-conjugated antibodies (room temperature; 1 h). DAPI was used for nuclear staining. Immunostainings were examined using inverted epifluorescent microscope (Zeiss™ Axiovert 200 M).

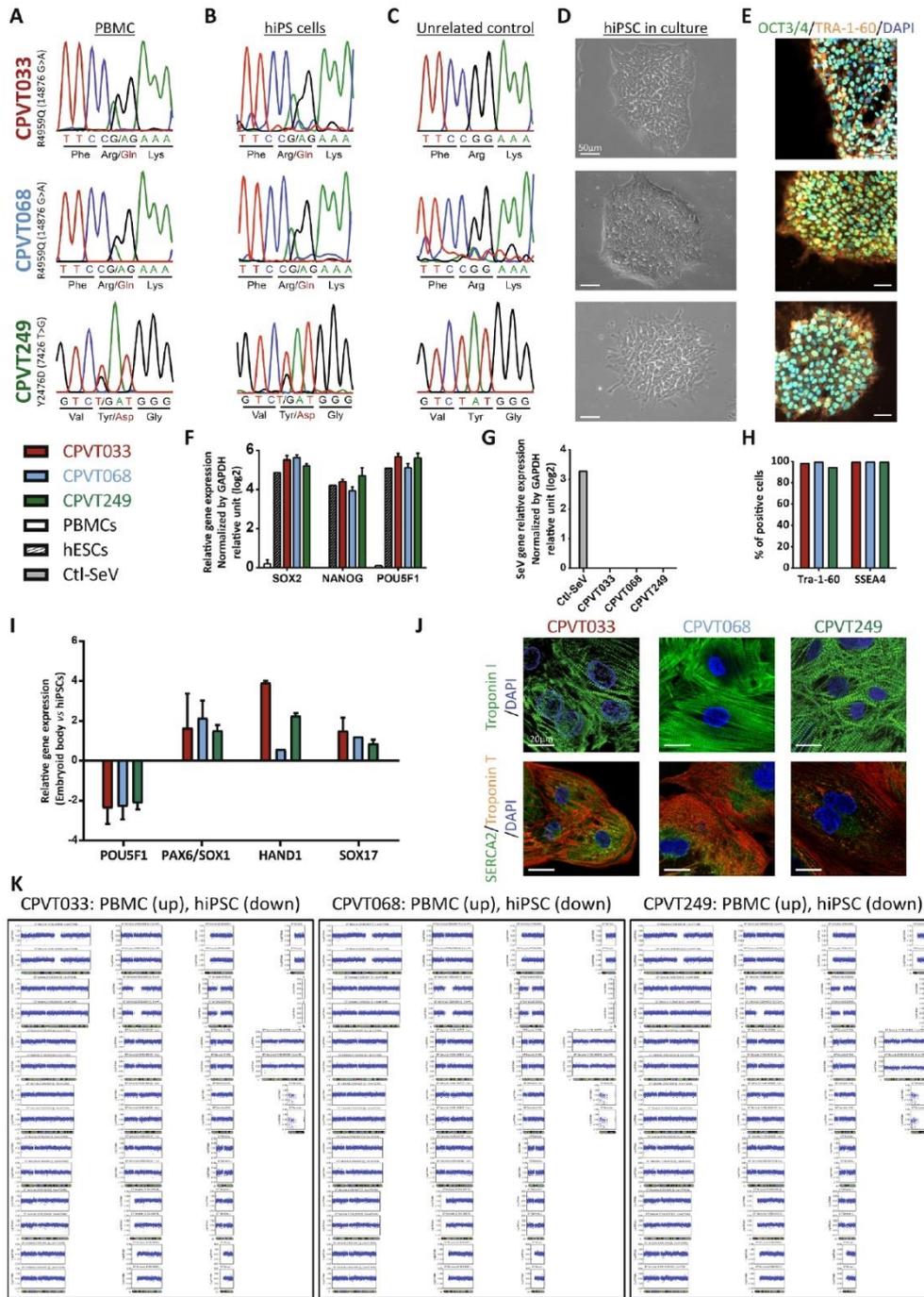


Fig. 1. Characterization of hiPSC lines from CPVT-affected patients carrying mutations in RYR2 gene. (A-C) Sanger sequencing of the RYR2 mutations in patients PBMCs, corresponding hiPSC lines and from an unrelated control, respectively. (D) Brightfield images of each hiPSC line in culture. (E) Immunofluorescence staining of pluripotency markers in each hiPSC line. (F) RT-qPCR analysis of pluripotency genes in parental PBMCs (negative control), reference hESC line (positive control) and each hiPSC line. (G) RT-qPCR analysis of Sendai Virus (SeV) loss in each hiPSC line as compared to a reference line positive for SeV presence. (H) Cytometry analysis of pluripotency markers in each hiPSC line. (I-J) Validation of differentiation potential of each hiPSC line into the 3 germ layers, using unsupervised EB differentiation (I) and directed cardiac differentiation (J). (K) Molecular karyotyping of the hiPSC lines. Above the schematic representation of each chromosome are represented the LogR ratio for PBMCs (up) and for hiPSC line (down).

Table 1
Characterization and validation.

Classification	Test	Result	Data
Morphology Phenotype	Photography Bright field	Normal	Fig 1D
	Qualitative analysis (Immunocytochemistry)	Assess staining/expression of pluripotency markers: OCT3/4, TRA-1-60	Fig 1E
	Quantitative analysis (Flow cytometry and RT-qPCR)	CPVT033: <ul style="list-style-type: none"> • TRA-1-60: 98.7 % • SSEA-4: 99.9 % • Expression of <i>SOX2</i>, <i>POU5F1</i> and <i>NANOG</i> CPVT068: <ul style="list-style-type: none"> • TRA-1-60: 99.9 % • SSEA-4: 100 % • Expression of <i>SOX2</i>, <i>POU5F1</i> and <i>NANOG</i> CPVT249: <ul style="list-style-type: none"> • TRA-1-60: 94.8 % • SSEA-4: 99.9 % • Expression of <i>SOX2</i>, <i>POU5F1</i> and <i>NANOG</i> 	Fig. 1F and H
Genotype	Molecular karyotyping (SNP array; >300000 markers)	CPVT033: 46,XX; Normal karyotype CPVT068: 46,XX; Normal karyotype CPVT249: 46,XX; Normal karyotype	Fig. 1K
Identity	CNV analysis	SNP array to compare CPVT033, CPVT068 and CPVT249 with the blood cells. Matching genotypes confirm the origins of iPSC lines	Supplemental Fig. 1E and archived with the journal
Mutation analysis (IF APPLICABLE)	Sequencing	CPVT033: 74 tested sites CPVT068: 70 tested sites CPVT249: 74 tested sites	Supplemental Fig. 1E and archived with the journal
	Southern Blot OR WGS	hiPSC lines carry the <i>RYR2</i> mutation present in the corresponding parental PBMC	Fig. 1A-C
Microbiology and virology	Mycoplasma	N/A	N/A
Differentiation potential	Embryoid body formation and differentiation followed by RT-qPCR and immunofluorescence analyses	Mycoplasma testing by luminescence: Negative	Supplemental Fig. 1D
List of recommended germ layer markers	Expression of markers at mRNA level (RT-qPCR) and protein level (immunofluorescence)	Expression of genes in embryoid bodies: Ectoderm (<i>PAX6</i> for CPVT033 and CPVT068; <i>SOX1</i> for CPVT249); Mesoderm (<i>HAND1</i> , troponin I, troponin T and <i>SERCA2</i>); Endoderm (<i>SOX17</i>)	Fig. 1I and J
		<ul style="list-style-type: none"> • Ectoderm: <i>PAX6/SOX1</i> • Mesoderm: <i>HAND1</i>, troponin I, troponin T and <i>SERCA2</i> • Endoderm: <i>SOX17</i> 	Fig. 1I and J
Donor screening (OPTIONAL)	HIV 1 + 2 Hepatitis B, Hepatitis C	N/A	N/A
Genotype additional info (OPTIONAL)	Blood group genotyping	N/A	N/A
	HLA tissue typing	N/A	N/A

4.5. SNP and CNV analyses

Single nucleotide polymorphism (SNP) analyses of hiPSC lines compared to parental PBMCs were used to confirm genome integrity. 200 ng of gDNA were outsourced to Integragen Company (Evry, France) for karyotype analyses using HumanCore-24-v1 SNP arrays. Analyses were performed with GenomeStudio software. Chromosome abnormalities were determined by visual inspection of LogR ratios values by comparing parental cells with hiPSC-derived samples. SNP analyses do not detect translocations or inversions, alterations in chromosome structure, mosaicism or polyploidy. SNP data were used to compute CNV and validate cell line parentality.

4.6. Differentiation to three germ layers

Embryonic body formation: hiPSC colonies were detached (p31-CPVT033; p22-CPVT068; p34-CPVT249) with dispase (STEMCELL Technologies), washed and replated in Costar® 6-well low attachment plates (Corning). The medium was changed the next day and unsupervised EB differentiation was started a day after, by incubating cells

(37 °C, 5 %CO₂, 21 %O₂) in EB medium: 78 % (v/v) KnockOut™ DMEM/F-12 (Gibco™) + 1 % (v/v) GlutaMAX™ Supplement (Gibco™) + 20 % (v/v) KnockOut™ Serum Replacement (Gibco™) + 1X MEM Non-Essential Amino Acids Solution + 0.1 mM 2-Mercaptoethanol (Gibco™). EB medium was changed every day during 21 days before RT-qPCR analysis.

Cardiac differentiation protocol: according to Al Sayed et al. (2021).

4.7. RT-qPCR

Total RNA were isolated using NucleoSpin RNA kit (MACHERY-NAGEL) and their quality assessed by NanoDrop1000. One µg RNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems™). PCR amplification was performed using (1) FAM-labelled TaqMan probes (Applied Biosystems™) for pluripotency and differentiation potential, and (2) SYBR-green master mix and primers (Applied Biosystems™) for SeV loss (p10-CPVT033; p9-CPVT068; p10-CPVT249). Data were normalized to ACTB or GAPDH. hESC (H9) expression data were provided by the iPSC core facility.

Table 2
Reagents details.

Antibodies used for immunocytochemistry/flow-cytometry				
	Antibody	Dilution	Company Cat #	RRID
Pluripotency Markers Immunocytochemistry	Rat anti-OCT3/4	1:200	Thermo Fisher Scientific Cat# 14-5841-80	AB_914299
Pluripotency Markers Immunocytochemistry	Mouse anti-TRA-1-60	1:200	Thermo Fisher Scientific Cat# 14-8863-80	AB_891612
Cardiomyocyte (mesoderm) Marker Immunocytochemistry	Rabbit anti-Troponin I	1:250	Santa Cruz Cat# sc-15368	AB_793465
Cardiomyocyte (mesoderm) Marker Immunocytochemistry	Mouse anti-Cardiac Troponin T	1:250	Abcam Cat# ab8295	AB_306445
Cardiomyocyte (mesoderm) Marker Immunocytochemistry	Anti-SERCA2		Santa Cruz Cat# sc-8094	AB_2061440
Pluripotency Markers Flow cytometry	TRA-1-60-PE	5 µl/test	Thermo Fisher Scientific Cat# 12-8863-82	AB_891602
Nuclear staining	DAPI	1 µg/mL	Merck Cat# D9564	
Secondary antibodies	Goat anti-Mouse IgG (H + L) Antibody, Alexa Fluor 568 Conjugated	1:1000	Molecular Probes Cat# A-11004	AB_2534072
Secondary antibodies	Goat anti-Rat IgG (H + L) Antibody, Alexa Fluor 488 Conjugated	1:1000	Molecular Probes Cat# A-11006	AB_141373
	Primers			
	Target	Size of band	Forward/Reverse primer (5'-3')	
Pluripotency Markers (TaqMan probe)	NANOG	109 bp	Thermo Fisher Scientific Cat# Hs02387400	
Pluripotency Markers (TaqMan probe)	POU5F1	77 bp	Thermo Fisher Scientific Cat# Hs04260367	
Pluripotency Markers (TaqMan probe)	SOX2	91 bp	Thermo Fisher Scientific Cat# Hs01053049	
Ectoderm Marker (TaqMan probe)	PAX6	86 bp	Thermo Fisher Scientific Cat# Hs01088114	
Ectoderm Marker (TaqMan probe)	SOX1	96 bp	Thermo Fisher Scientific Cat# Hs01057642	
Endoderm Marker (TaqMan probe)	SOX17	149 bp	Thermo Fisher Scientific Cat# Hs00751752	
Mesoderm Marker (TaqMan probe)	HAND1	65 bp	Thermo Fisher Scientific Cat# Hs02330376	
House-Keeping Genes (TaqMan probe)	ACTB	63 bp	Thermo Fisher Scientific Cat# Hs01060665	
House-Keeping Genes (primer)	GAPDH	81 bp	Forward: 5'-AAT CCC ATC ACC ATC TTC CA-3' Reverse: 5'-TGG ACT CGA CGA CGT ACT CA-3'	
Sendai Virus Detection (primer)	SeV	181 bp	Forward: 5'-GGA TCA CTA GGT GAT ATC GAG C-3' Reverse: 5'-ACC AGA CAA GAG TTT AAG AGA TAT GTA TC-3'	
Mutation sequencing	RYR2 p.Y2476D	186 bp	Forward: 5'-ACA CAG CCA TTG ACA CCA AA-3' Reverse: 5'-CGC TCC AAT CTC ACC GTA TC-3'	
Mutation sequencing	RYR2 p.R4959Q	350 bp	Forward: 5'-GCA CAC ACT TTG GGG AAA AT-3' Reverse: 5'-GGC AAT TCA CAA AAT CAG CA-3'	

4.8. Mycoplasma detection

Mycoplasma contamination was assessed (p28-CPVT033; p30-CPVT068; p29-CPVT249) with MycoAlert™ Mycoplasma Detection Kit (Lonza).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the Genomic Core Facility (Geno-BiRD, Biogenouest), the Cytometry facility (CytoCell) and the iPS core facility of Nantes University, for their technical support. This work was funded by grants from The French National Research Agency (HEART iPS ANR-15-CE14-0019-01), La Fédération Française de Cardiologie and the Fondation Genavie to M.D.W.. B.C. and R.C. are laureate of fellowships from the Fondation pour la Recherche Médicale (FRM, PBR201810007614) and the Fonds Marion Elizabeth Brancher, respectively.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.scr.2022.102688>.

References

- Allouis, M., Probst, V., Jaafar, P., Schott, J.-J., Le Marec, H., 2005. Unusual clinical presentation in a family with catecholaminergic polymorphic ventricular tachycardia due to a G14876A ryanodine receptor gene mutation. *Am. J. Cardiol.* 95 (5), 700–702.
- Al Sayed, Z.R., Canac, R., Cimarosti, B., Bonnard, C., Gourraud, J.-B., Hamamy, H., Kayserli, H., Girardeau, A., Jouni, M., Jacob, N., Gaignerie, A., Chariau, C., David, L., Forest, V., Marionneau, C., Charpentier, F., Loussouarn, G., Lamirault, G., Reversade, B., Zibara, K., Lemarchand, P., Gaborit, N., 2021. Human model of IRX5 mutations reveals key role for this transcription factor in ventricular conduction. *Cardiovasc. Res.* 117 (9), 2092–2107.
- Priori, S.G., Napolitano, C., Memmi, M., Colombi, B., Drago, F., Gasparini, M., DeSimone, L., Coltorti, F., Bloise, R., Keegan, R., Cruz Filho, F.E.S., Vignati, G., Benatar, A., DeLogu, A., 2002. Clinical and molecular characterization of patients with catecholaminergic polymorphic ventricular tachycardia. *Circulation* 106 (1), 69–74.
- Swan, H., Piippo, K., Viitasalo, M., Heikkilä, P., Paavonen, T., Kainulainen, K., Kere, J., Keto, P., Kontula, K., Toivonen, L., 1999. Arrhythmic disorder mapped to chromosome 1q42-q43 causes malignant polymorphic ventricular tachycardia in structurally normal hearts. *J. Am. Coll. Cardiol.* 34 (7), 2035–2042.
- Xiao, Z., Guo, W., Sun, B.O., Hunt, D.J., Wei, J., Liu, Y., Wang, Y., Wang, R., Jones, P.P., Back, T.G., Chen, S.R.W., 2016. Enhanced Cytosolic Ca²⁺ Activation Underlies a Common Defect of Central Domain Cardiac Ryanodine Receptor Mutations Linked to Arrhythmias. *J. Biol. Chem.* 291 (47), 24528–24537.

Annexe 5

Generation of human induced pluripotent stem cell lines from two patients affected by catecholamine-induced QT prolongation (CIQTP)

Bastien Cimarosti, Robin Canac, Aurore Girardeau, Marine Arnaud, Quentin Francheteau, Vincent Probst, Patricia Lemarchand, Richard Redon, Jean-Baptiste Gourraud, Nathalie Gaborit*, Guillaume Lamirault*

* : Co-derniers auteurs

DOI : 10.1016/j.scr.2021.102649 ; PMID : 34995842

Publié dans *Stem Cell Research*, 2022



Contents lists available at ScienceDirect

Stem Cell Research

journal homepage: www.elsevier.com/locate/scr

Lab Resource: Multiple Cell Lines



Generation of human induced pluripotent stem cell lines from two patients affected by catecholamine-induced QT prolongation (CIQTP)

Bastien Cimarosti^a, Robin Canac^a, Aurore Girardeau^a, Marine Arnaud^a, Quentin Francheteau^b, Vincent Probst^a, Patricia Lemarchand^a, Richard Redon^a, Jean-Baptiste Gourraud^a, Nathalie Gaborit^{a,1}, Guillaume Lamirault^{a,1,*}

^a Université de Nantes, CHU Nantes, CNRS, INSERM, l'institut du thorax, F-44000 Nantes, France

^b Université de Nantes, CHU Nantes, INSERM, CNRS, SFR Santé, FED 4203, INSERM UMS 016, CNRS UMS 3556, F-44000 Nantes, France

A B S T R A C T

Catecholamine-induced QT prolongation (CIQTP) is an inherited cardiac disease characterized by a normal baseline ECG and a risk of sudden cardiac death by ventricular arrhythmia due to a QT prolongation that only appears during catecholergic stimulation, especially mental stress. Induced pluripotent stem cells (hiPSCs) were generated from peripheral blood mononuclear cells collected from two CIQTP-affected patients from two different families. These two hiPSC lines are a valuable model to study biological alterations due to CIQTP.

1. Resource table

Unique stem cell lines identifier	ITXi003-A ITXi004-A
Alternative name(s) of stem cell lines	MS368 (ITXi003-A) MS381 (ITXi004-A)
Institution	l'institut du thorax (INSERM UMR1087, CNRS UMR6291)
Contact information of distributor	Guillaume Lamirault, guillaume.lamirault@univ-nantes.fr
Type of cell lines	iPSC
Origin	Human
Additional origin info required	MS368 (ITXi003-A) <ul style="list-style-type: none"> • Age: 13 • Male • Ethnicity: European MS381 (ITXi004-A) <ul style="list-style-type: none"> • Age: 46 • Female • Ethnicity: European
Cell Source	Peripheral blood mononuclear cells
Clonality	Clonal (molecular karyotyping)
Method of reprogramming	Sendai virus
Genetic Modification	No
Type of Genetic Modification	N/A
Evidence of the reprogramming transgene loss (including genomic copy if applicable)	RT-qPCR

(continued on next column)

(continued)

Unique stem cell lines identifier	ITXi003-A ITXi004-A
Associated disease	Catecholamine-induced QT prolongation (CIQTP)
Gene/locus	N/A
Date archived/stock date	MS368: 11-02-2016 MS381: 12-20-2016
Cell line repository/bank	MS368: https://hpscereg.eu/cell-line/ITXi003-A MS381: https://hpscereg.eu/cell-line/ITXi004-A
Ethical approval	For all patients, a signed informed consent has been received. Authorization MESR n°DC-2011-1399 and Authorization CPP Ouest II n°2010-33 have been obtained from competent ethics review board and approval authority (French Health Ministry and French ethical committee) allowing the generation of these lines.

2. Resource utility

The two generated hiPSC lines, reprogrammed from somatic cells of two CIQTP patients, provide a cellular model to decipher biological

* Corresponding author at: l'institut du thorax Inserm UMR 1087, CNRS UMR 6291, IRS-UN, 8 quai Moncoussu, 44007 Nantes cedex 1, France.

E-mail address: guillaume.lamirault@univ-nantes.fr (G. Lamirault).

¹ Contributed equally to this work.

<https://doi.org/10.1016/j.scr.2021.102649>

Received 26 November 2021; Accepted 29 December 2021

Available online 30 December 2021

1873-5061/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

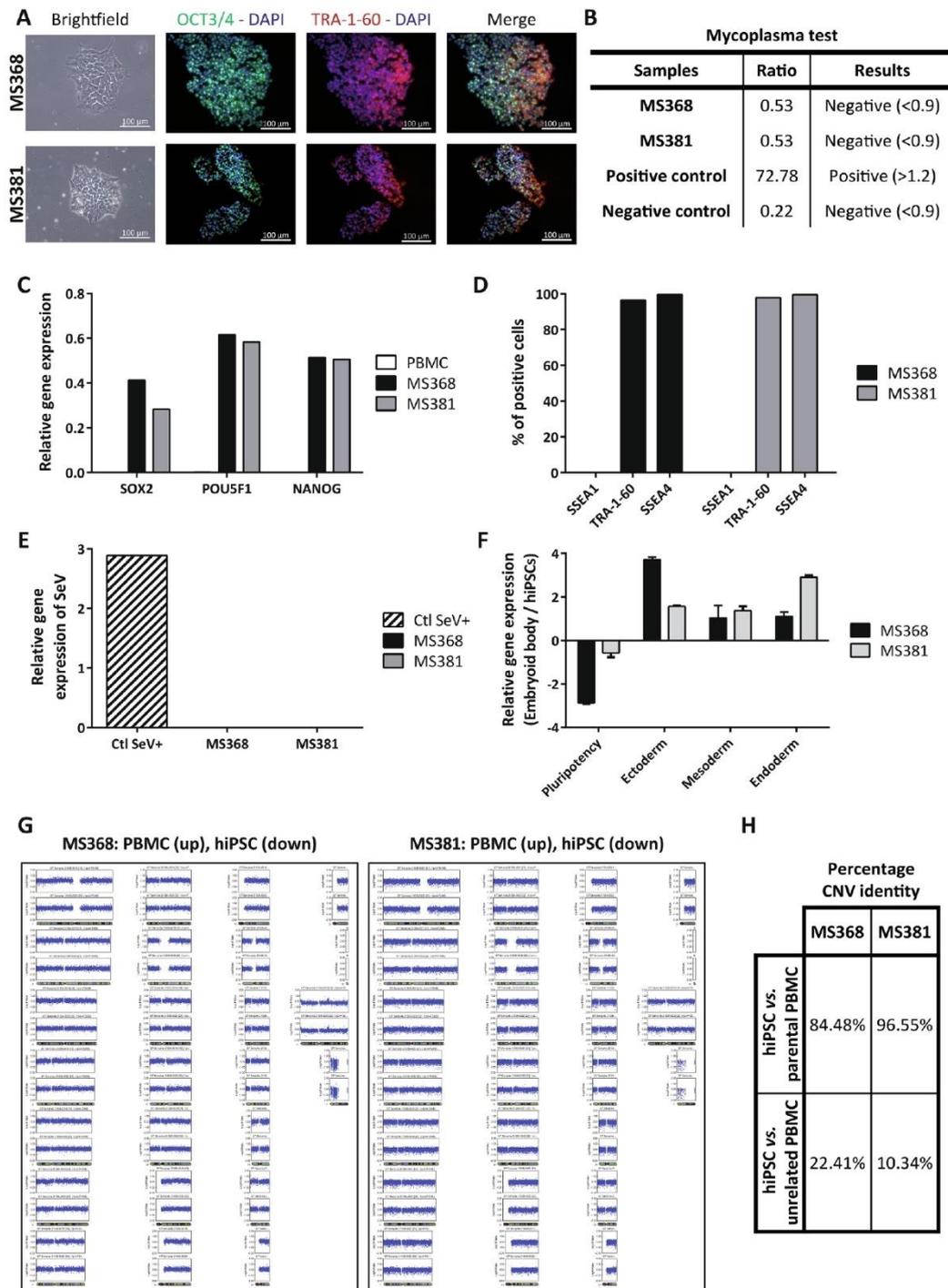


Fig. 1.

Table 1
Characterization and validation.

Classification	Test	Result	Data
Morphology	Photography Bright field	Normal	Fig. 1A left panels
Phenotype	Qualitative analysis (Immunocytochemistry)	Assess staining/ expression of pluripotency markers: OCT3/4, TRA-1-60	Fig. 1A right panels
	Quantitative analysis (Flow cytometry and RT-qPCR)	MS368: <ul style="list-style-type: none"> • SSEA-1: 0% • TRA-1-60: 96% • SSEA-4: 99.8% • Expression of <i>SOX2</i>, <i>POU5F1</i> and <i>NANOG</i> MS381: <ul style="list-style-type: none"> • SSEA-1: 0.1% • TRA-1-60: 97.5% • SSEA-4: 99.4% • Expression of <i>SOX2</i>, <i>POU5F1</i> and <i>NANOG</i> 	Fig. 1C and 1D
Genotype	Molecular karyotyping (SNP array; greater than 300000 markers)	MS368: 46,XY; Normal karyotype; MS381: 46,XX; Normal karyotype	Fig. 1G
Identity	CNV analysis	SNP array to compare MS368 and MS381 with the blood cells. Matching genotypes confirm the origins of iPSC lines MS368: 58 tested sites; MS381: 58 tested sites	Fig. 1H and archived with the journal Fig. 1H and archived with the journal
Mutation analysis (IF APPLICABLE)	Sequencing Southern Blot OR WGS	N/A N/A	N/A N/A
Microbiology and virology	Mycoplasma	Mycoplasma testing by luminescence: Negative	Fig. 1B
Differentiation potential	Embryoid body formation and differentiation followed by RT-qPCR analysis	Expression of genes in embryoid bodies: Ectoderm (<i>PAX6</i>); Mesoderm (<i>HAND1</i>); Endoderm (<i>SOX17</i>)	Fig. 1F
List of recommended germ layer markers	Expression of markers at mRNA level (RT-qPCR)	<ul style="list-style-type: none"> • Ectoderm: <i>PAX6</i> • Mesoderm: <i>HAND1</i> • Endoderm: <i>SOX17</i> 	Fig. 1F
Donor screening (OPTIONAL)	HIV 1 + 2 Hepatitis B, Hepatitis C	N/A	N/A
Genotype additional info (OPTIONAL)	Blood group genotyping HLA tissue typing	N/A N/A	N/A N/A

processes associated with the disease pathophysiology, notably through their differentiation into cardiac cells.

3. Resource details

Catecholamine-induced QT prolongation (CIQTP) is a recently

described inherited cardiac arrhythmia characterized by an abnormal QT interval prolongation on the ECG during catecholergic stimulation, which has been shown to be responsible for the occurrence of life-threatening ventricular arrhythmias and sudden cardiac death (SCD; Finlay et al., 2016; Huchet et al., 2017). CIQTP is distinguished from long QT syndrome by a normal ECG at rest but a major and abnormal QT interval prolongation during catecholergic stimulation, especially mental stress (Priori et al., 2015; Etienne et al., 2018). While familial investigations demonstrated the hereditary character of CIQTP, no genetic data is available so far. In order to provide an inexhaustible source of cardiomyocytes carrying CIQTP genetic backgrounds, two patients were selected from two families (Huchet et al., 2017). Patient MS368 is the propositus of the first family. He suffered at age 10, during a swimming lesson, from a resuscitated SCD. Corrected QT (QTc) measurement at rest was normal (420 ms) but abnormally prolonged under mental stress test (516 ms), exercise test (510 ms) and epinephrine challenge (555 ms). Patient MS381 is the twin sister of the second family propositus who suddenly died at age 30 while jogging. QTc measurement of patient MS381 was considered as normal at rest (420 ms) but increased to 600 ms during mental stress test but not during exercise test (epinephrine challenge not performed).

Using the Sendai virus (SeV) method, peripheral blood mononuclear cells (PBMCs) were reprogrammed into induced pluripotent stem cells (hiPSCs) at the iPSC core facility of Nantes. hiPSCs were maintained on Matrigel®-coated plates with StemMACS™ iPS-Brew XF medium (Fig. 1A left panels). Immunocytochemistry proved expression and proper cellular localization of pluripotency markers OCT3/4 and TRA-1-60 (Fig. 1A right panels). No mycoplasma contamination was found in the hiPSC culture media (Fig. 1B). For each hiPSC line, RT-qPCR showed expression of pluripotency markers (Fig. 1C) and flow cytometry analysis showed that more than 95% of cells were positive for TRA-1-60 and SSEA4 pluripotency markers, while 0% for SSEA1, as expected (Fig. 1D). Loss of SeV expression was verified by RT-qPCR (Fig. 1E). Single nucleotide polymorphism (SNP) genotyping confirmed (1) genomic integrity of both hiPSC lines by molecular karyotyping (Fig. 1G) and (2) parentality of each line to their corresponding PBMC samples by CNV analysis (Fig. 1H). Finally, differentiation capacity in derivatives of the 3 germ layers was demonstrated by RT-qPCR on spontaneously differentiated embryoid bodies. It showed increase in expression of specific differentiation markers and, concordantly, decrease in expression of pluripotency marker (Fig. 1F). Characterization of hiPSC lines is summarized in Table 1 and Table 2.

4. Materials and methods

4.1. Generation and maintenance of hiPSC lines

Collected blood cells were reprogrammed by the iPSC core facility (PF-iPSC, Nantes) using Sendai virus method (Cytotune reprogramming Kit, Life technologies). hiPSC lines were maintained at 37 °C, 5% CO₂, 21% O₂ in StemMACS™ iPS Brew XF Medium (Miltenyi Biotec) on culture plates coated with Matrigel® hESC-Qualified Matrix (0.05 mg/mL, Corning). At 75% confluency, cells were passaged using Gentle Cell Dissociation Reagent (STEMCELL Technologies).

4.2. Flow cytometry analysis

A total of 1x10⁵ cells in suspension (p22-MS368; p21-MS381) were washed 3 times with PBS/BSA 0.1% and then incubated with PE-labelled antibodies for 30 min at 4 °C in the dark. Data acquisition was performed using FACSDiva software with LSR II instrument (Becton Dickinson, BD).

4.3. Immunofluorescence

Cells were dissociated and seeded onto 8-well IBIDI plates coated

Table 2
Reagents details.

Antibodies used for immunocytochemistry/flow-cytometry				
	Antibody	Dilution	Company Cat #	RRID
Pluripotency MarkersImmunocytochemistry	Rat anti-OCT3/4	1:200	Thermo Fisher Scientific Cat# 14-5841-80	AB_914299
Pluripotency MarkersImmunocytochemistry	Mouse anti-TRA-1-60	1:200	Thermo Fisher Scientific Cat# 14-8863-80	AB_891_612
Non-pluripotency MarkersFlow cytometry	SSEA-1-PE	10 µl/test	Miltenyi Biotec Cat#130-117-801	AB_2728020
Pluripotency MarkersFlow cytometry	SSEA-4-PE	5 µl/test	Thermo Fisher Scientific Cat# 12-8843-42	AB_11151520
Pluripotency MarkersFlow cytometry	TRA-1-60-PE	5 µl/test	Thermo Fisher Scientific Cat# 12-8863-82	AB_891602
Nuclear staining	DAPI	1 µg/mL	Merck Cat# D9564	
Secondary antibodies	Goat Anti-Mouse IgG (H + L) Antibody, Alexa Fluor 568 Conjugated	1:1000	Molecular Probes Cat# A-11004	AB_2534072
Secondary antibodies	Goat Anti-Rat IgG (H + L) Antibody, Alexa Fluor 488 Conjugated	1:1000	Molecular Probes Cat# A-11006	AB_141373
Primers				
	Target	Size of band	Forward/Reverse primer (5'-3')	
Pluripotency Markers (TaqMan probe)	<i>NANOG</i>	109 bp	Thermo Fisher Scientific Cat# Hs02387400	
Pluripotency Markers (TaqMan probe)	<i>POU5F1</i>	77 bp	Thermo Fisher Scientific Cat# Hs04260367	
Pluripotency Markers (TaqMan probe)	<i>SOX2</i>	91 bp	Thermo Fisher Scientific Cat# Hs01053049	
Ectoderm Marker (TaqMan probe)	<i>PAX6</i>	86 bp	Thermo Fisher Scientific Cat# Hs01088114	
Endoderm Marker (TaqMan probe)	<i>SOX17</i>	149 bp	Thermo Fisher Scientific Cat# Hs00751752	
Mesoderm Marker (TaqMan probe)	<i>HAND1</i>	65 bp	Thermo Fisher Scientific Cat# Hs02330376	
House-Keeping Genes (TaqMan probe)	<i>ACTB</i>	63 bp	Thermo Fisher Scientific Cat# Hs01060665	
House-Keeping Genes (primer)	<i>GAPDH</i>	81 bp	Forward: 5'-AAT CCC ATC ACC ATC TTC CA-3' Reverse: 5'-TGG ACT CCA CGA CGT ACT CA-3'	
Sendai Virus Detection (primer)	SeV	181 bp	Forward: 5'-GGA TCA CTA GGT GAT ATC GAG C-3' Reverse: 5'-ACC AGA CAA GAG TTT AAG AGA TAT GTA TC-3'	

with Matrigel® (p21-MS368; p22-MS381). Cultured cells were fixed with 4% paraformaldehyde for 15 min, permeabilized with 0.5% Triton-X100 for 15 min and blocked with 3% BSA for 30 min. Immunofluorescence stainings were performed using appropriate primary antibodies (4 °C; overnight) and their corresponding Alexa conjugated antibodies (room temperature; 1 h). DAPI was used for nuclear staining. Immunostainings were examined using inverted epifluorescent microscope (Zeiss Axiovert 200 M).

4.4. SNP and CNV analysis

Single nucleotide polymorphism (SNP) analysis of hiPSC lines, compared to parental blood cells, was used to confirm genome integrity. DNA was extracted from somatic and hiPSC samples (p19-MS368; p18-MS381) using the QIAGEN QiaAmp kit. 200 ng of gDNA were outsourced to Integrigen Company (Evry, France) for karyotype analysis using HumanCore-24-v1 SNP arrays. Analysis was performed with GenomeStudio software. Chromosome abnormalities were determined by visual inspection of logR ratios values and comparing parental cells with hiPSC-derived samples. SNP data were used to compute CNV and validate cell lines parentality.

4.5. Embryoid body formation

hiPSC colonies were detached (p20-MS368; p19-MS381) with dispase (STEMCELL Technologies), washed and replated in Costar® 6 Well Plate Low Attachment (Corning). Next day the medium was changed and the unsupervised EB differentiation was started a day after by incubating cells (37 °C, 5% CO₂, 21% O₂) in EB medium: 78% (v/v) KnockOut™ DMEM/F-12 (Gibco) + 1% (v/v) GlutaMAX™ Supplement (Gibco) + 20% (v/v) KnockOut™ Serum Replacement (Gibco) + 1X MEM Non-Essential Amino Acids Solution + 0.1 mM 2-Mercaptoethanol (Gibco). EB medium was changed every day during 21 days prior RT-qPCR analysis.

4.6. RT-qPCR

Total RNA samples were isolated using the NucleoSpin RNA kit (MACHEREY-NAGEL) and their quality assessed by NanoDrop1000. 1 µg of tRNA was reverse transcribed using High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems). PCR amplification was performed using (1) FAM-labeled TaqMan probes (Applied Biosystems) for pluripotency and differentiation potential, and (2) SYBR-green master mix and primers (Applied Biosystems) for SeV loss (p23-MS368; p25-MS381). Data were normalized to ACTB or GAPDH.

4.7. Mycoplasma detection

Mycoplasma contamination was assessed (p19-MS368; p18-MS381) with MycoAlert™ Mycoplasma Detection Kit (Lonza).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the Cytometry core facility (Cyto-Cell) and the iPSC core facility of Nantes University, for their technical support. This work was funded by grants from Hôpitaux Universitaires du Grand Ouest and Fondation Maladies Rares (HUGO-FMR; RC17_0357), The French National Research Agency (HEART iPSC ANR-15-CE14-0019-01) and La Fédération Française de Cardiologie. This work was supported by the Fondation pour la Recherche Médicale, grant number PBR201810007614, to B.C. R.C. is laureate of fellowship from the Fonds Marion Elizabeth Brancher.

References

- Etienne, P., Huchet, F., Gaborit, N., Barc, J., Thollet, A., Kyndt, F., Guyomarch, B., Le Marec, H., Charpentier, F., Schott, J.-J., Redon, R., Probst, V., Gourraud, J.-B., 2018. Mental stress test: a rapid, simple, and efficient test to unmask long QT syndrome. *Europace* 20 (12), 2014–2020.
- Finlay, M.C., Lambiase, P.D., Ben-Simon, R., Taggart, P., 2016. Effect of mental stress on dynamic electrophysiological properties of the endocardium and epicardium in humans. *Heart Rhythm*. 13 (1), 175–182.
- Huchet, F., Kyndt, F., Barc, J., Thollet, A., Charpentier, F., Redon, R., Schott, J.J., le Marec, H., Probst, V., Gourraud, J.B., 2017. Familial catecholamine-induced QT prolongation in unexplained sudden cardiac death. *J Am Coll Cardiol*. 69 (12), 1642–1643.
- Priori, S.G., Blomström-Lundqvist, C., Mazzanti, A., Blom, N., Borggrefe, M., Camm, J., Elliott, P.M., Fitzsimons, D., Hatala, R., Hindricks, G., Kirchhof, P., Kjeldsen, K., Kuck, K.H., Hernandez-Madrid, A., Nikolaou, N., Norekvål, T.M., Spaulding, C., Van Veldhuisen, D.J., 2015. 2015 ESC Guidelines for the management of patients with ventricular arrhythmias and the prevention of sudden cardiac death. *Europace* 17, 1601–1687.

VII

Bibliographie

Acimovic, I., Vilotic, A., Pesl, M., Lacampagne, A., Dvorak, P., Rotrekl, V., and Meli, A.C. (2014). Human pluripotent stem cell-derived cardiomyocytes as research and therapeutic tools. *BioMed Res. Int.* 2014, 512831. <https://doi.org/10.1155/2014/512831>.

Afzal, M.Z., Gartz, M., Klyachko, E.A., Khan, S.S., Shah, S.J., Gupta, S., Shapiro, A.D., Vaughan, D.E., and Strande, J.L. (2017). Generation of human iPSCs from urine derived cells of a non-affected control subject. *Stem Cell Res.* 18, 33–36. <https://doi.org/10.1016/j.scr.2016.12.008>.

Agrawal, P., Heimbruch, K.E., and Rao, S. (2018). Genome-wide maps of transcription regulatory elements and transcription enhancers in development and disease. *Compr. Physiol.* 9, 439–455. <https://doi.org/10.1002/cphy.c180028>.

Agu, C.A., Soares, F.A.C., Alderton, A., Patel, M., Ansari, R., Patel, S., Forrest, S., Yang, F., Lineham, J., Vallier, L., et al. (2015). Successful Generation of Human Induced Pluripotent Stem Cell Lines from Blood Samples Held at Room Temperature for up to 48 hr. *Stem Cell Rep.* 5, 660–671. <https://doi.org/10.1016/j.stemcr.2015.08.012>.

Ahmed, R.E., Anzai, T., Chanthra, N., and Uosaki, H. (2020). A Brief Review of Current Maturation Methods for Human Induced Pluripotent Stem Cells-Derived Cardiomyocytes. *Front. Cell Dev. Biol.* 8. .

Akar, F.G., and Rosenbaum, D.S. (2003). Transmural Electrophysiological Heterogeneities Underlying Arrhythmogenesis in Heart Failure. *Circ. Res.* 93, 638–645. <https://doi.org/10.1161/01.RES.0000092248.59479.AE>.

Akerberg, B.N., Gu, F., VanDusen, N.J., Zhang, X., Dong, R., Li, K., Zhang, B., Zhou, B., Sethi, I., Ma, Q., et al. (2019). A reference map of murine cardiac transcription factor chromatin occupancy identifies dynamic and conserved enhancers. *Nat. Commun.* 10, 4907. <https://doi.org/10.1038/s41467-019-12812-3>.

Al Sayed, Z.R., Canac, R., Cimarosti, B., Bonnard, C., Gourraud, J.-B., Hamamy, H., Kayserili, H., Girardeau, A., Jouni, M., Jacob, N., et al. (2021a). Human model of IRX5 mutations reveals key role for this transcription factor in ventricular conduction. *Cardiovasc. Res.* 117, 2092–2107. <https://doi.org/10.1093/cvr/cvaa259>.

Al Sayed, Z.R., Jouni, M., Gourraud, J.-B., Belbachir, N., Barc, J., Girardeau, A., Forest, V., Derevier, A., Gaignerie, A., Chariou, C., et al. (2021b). A consistent arrhythmogenic trait in Brugada syndrome cellular phenotype. *Clin. Transl. Med.* 11, e413. <https://doi.org/10.1002/ctm2.413>.

Alam, M.S. (2018). Proximity Ligation Assay (PLA). *Curr. Protoc. Immunol.* 123, e58. <https://doi.org/10.1002/cpim.58>.

Albini, S., Zakharova, V., and Ait-Si-Ali, S. (2019). Histone Modifications. In *Epigenetics and Regeneration*, (Elsevier), pp. 47–72.

Alexander, J.M., Guan, J., Li, B., Maliskova, L., Song, M., Shen, Y., Huang, B., Lomvardas, S., and Weiner, O.D. (2019). Live-cell imaging reveals enhancer-dependent Sox2 transcription in the absence of enhancer proximity. *ELife* 8, e41769. <https://doi.org/10.7554/eLife.41769>.

- Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat. Rev. Genet.* *17*, 487–500. <https://doi.org/10.1038/nrg.2016.59>.
- Allshire, R.C., and Madhani, H.D. (2017). Ten principles of heterochromatin formation and function. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/nrm.2017.119>.
- Altschuler, S.J., and Wu, L.F. (2010). Cellular heterogeneity: do differences make a difference? *Cell* *141*, 559–563. <https://doi.org/10.1016/j.cell.2010.04.033>.
- Altshuler, D.M., Gibbs, R.A., Peltonen, L., Altshuler, D.M., Gibbs, R.A., Peltonen, L., Dermitzakis, E., Schaffner, S.F., Yu, F., Peltonen, L., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature* *467*, 52–58. <https://doi.org/10.1038/nature09298>.
- Anderson, D.J., Kaplan, D.I., Bell, K.M., Koutsis, K., Haynes, J.M., Mills, R.J., Phelan, D.G., Qian, E.L., Leitoguinho, A.R., Arasaratnam, D., et al. (2018). NKX2-5 regulates human cardiomyogenesis via a HEY2 dependent transcriptional network. *Nat. Commun.* *9*, 1373. <https://doi.org/10.1038/s41467-018-03714-x>.
- Anderson, D.M., George, R., Noyes, M.B., Rowton, M., Liu, W., Jiang, R., Wolfe, S.A., Wilson-Rawls, J., and Rawls, A. (2012). Characterization of the DNA-binding Properties of the Mohawk Homeobox Transcription Factor. *J. Biol. Chem.* *287*, 35351–35359. <https://doi.org/10.1074/jbc.M112.399386>.
- Andersson, R., and Sandelin, A. (2020). Determinants of enhancer and promoter activities of regulatory elements. *Nat. Rev. Genet.* *21*, 71–87. <https://doi.org/10.1038/s41576-019-0173-8>.
- Andersson, C., Lin, H., Liu, C., Levy, D., Mitchell, G.F., Larson, M.G., and Vasan, R.S. (2019). Integrated Multiomics Approach to Identify Genetic Underpinnings of Heart Failure and Its Echocardiographic Precursors. *Circ. Genomic Precis. Med.* *12*, e002489. <https://doi.org/10.1161/CIRCGEN.118.002489>.
- Andrysiak, K., Stępniewski, J., and Dulak, J. (2021). Human-induced pluripotent stem cell-derived cardiomyocytes, 3D cardiac structures, and heart-on-a-chip as tools for drug research. *Pflugers Arch.* *473*, 1061–1085. <https://doi.org/10.1007/s00424-021-02536-z>.
- Ang, Y.-S., Rivas, R.N., Ribeiro, A.J.S., Srivas, R., Rivera, J., Stone, N.R., Pratt, K., Mohamed, T.M.A., Fu, J.-D., Spencer, C.I., et al. (2016). Disease Model of GATA4 Mutation Reveals Transcription Factor Cooperativity in Human Cardiogenesis. *Cell* *167*, 1734–1749.e22. <https://doi.org/10.1016/j.cell.2016.11.033>.
- Antzelevitch, C. (2005). Modulation of Transmural Repolarization. *Ann. N. Y. Acad. Sci.* *1047*, 314–323. <https://doi.org/10.1196/annals.1341.028>.
- Anzai, T., Yamagata, T., and Uosaki, H. (2020). Comparative Transcriptome Landscape of Mouse and Human Hearts. *Front. Cell Dev. Biol.* *8*, 268. <https://doi.org/10.3389/fcell.2020.00268>.

Argelaguet, R., Lohoff, T., Li, J.G., Nakhuda, A., Drage, D., Krueger, F., Velten, L., Clark, S.J., and Reik, W. (2022). Decoding gene regulation in the mouse embryo using single-cell multi-omics. 2022.06.15.496239. <https://doi.org/10.1101/2022.06.15.496239>.

Arslan, U., Orlova, V.V., and Mummery, C.L. (2022). Perspectives for Future Use of Cardiac Microtissues from Human Pluripotent Stem Cells. *ACS Biomater. Sci. Eng.* <https://doi.org/10.1021/acsbiomaterials.1c01296>.

Asp, M., Giacomello, S., Larsson, L., Wu, C., Fürth, D., Qian, X., Wärdell, E., Custodio, J., Reimegård, J., Salmén, F., et al. (2019). A Spatiotemporal Organ-Wide Gene Expression and Cell Atlas of the Developing Human Heart. *Cell* *179*, 1647–1660.e19. <https://doi.org/10.1016/j.cell.2019.11.025>.

Atkins, M., Potier, D., Romanelli, L., Jacobs, J., Mach, J., Hamaratoglu, F., Aerts, S., and Halder, G. (2016). An Ectopic Network of Transcription Factors Regulated by Hippo Signaling Drives Growth and Invasion of a Malignant Tumor Model. *Curr. Biol. CB* *26*, 2101–2113. <https://doi.org/10.1016/j.cub.2016.06.035>.

Atlasi, Y., and Stunnenberg, H.G. (2017). The interplay of epigenetic marks during stem cell differentiation and development. *Nat. Rev. Genet.* *18*, 643–658. <https://doi.org/10.1038/nrg.2017.57>.

Ávila-Flores, A., Arranz-Nicolás, J., and Mérida, I. (2019). Transcriptional Activity of FOXO Transcription Factors Measured by Luciferase Assays. *Methods Mol. Biol. Clifton NJ* *1890*, 91–102. https://doi.org/10.1007/978-1-4939-8900-3_8.

Ávila-González, D., Portillo, W., García-López, G., Molina-Hernández, A., Díaz-Martínez, N.E., and Díaz, N.F. (2021). Unraveling the Spatiotemporal Human Pluripotency in Embryonic Development. *Front. Cell Dev. Biol.* *9*, 676998. <https://doi.org/10.3389/fcell.2021.676998>.

Azpeitia, E., and Wagner, A. (2020). Short Residence Times of DNA-Bound Transcription Factors Can Reduce Gene Expression Noise and Increase the Transmission of Information in a Gene Regulation System. *Front. Mol. Biosci.* *7*.

Bae, S., and Lesch, B.J. (2020). H3K4me1 Distribution Predicts Transcription State and Poising at Promoters. *Front. Cell Dev. Biol.* *8*.

Bakker, M.L., Moorman, A.F.M., and Christoffels, V.M. (2010). The atrioventricular node: origin, development, and genetic program. *Trends Cardiovasc. Med.* *20*, 164–171. <https://doi.org/10.1016/j.tcm.2011.02.001>.

Balsalobre, A., and Drouin, J. (2022). Pioneer factors as master regulators of the epigenome and cell fate. *Nat. Rev. Mol. Cell Biol.* *23*, 449–464. <https://doi.org/10.1038/s41580-022-00464-z>.

Ban, K., Wile, B., Cho, K.-W., Kim, S., Song, M.-K., Kim, S.Y., Singer, J., Syed, A., Yu, S.P., Wagner, M., et al. (2015). Non-genetic Purification of Ventricular Cardiomyocytes from Differentiating Embryonic Stem Cells through Molecular Beacons Targeting IRX-4. *Stem Cell Rep.* *5*, 1239–1249. <https://doi.org/10.1016/j.stemcr.2015.10.021>.

- Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* *21*, 381–395. <https://doi.org/10.1038/cr.2011.22>.
- Bao, Z.Z., Bruneau, B.G., Seidman, J.G., Seidman, C.E., and Cepko, C.L. (1999). Regulation of chamber-specific gene expression in the developing heart by *Irx4*. *Science* *283*, 1161–1164. <https://doi.org/10.1126/science.283.5405.1161>.
- Barc, J., Tadros, R., Glinge, C., Chiang, D.Y., Jouni, M., Simonet, F., Jurgens, S.J., Baudic, M., Nicastro, M., Potet, F., et al. (2022). Genome-wide association analyses identify new Brugada syndrome risk loci and highlight a new mechanism of sodium channel regulation in disease susceptibility. *Nat. Genet.* *54*, 232–239. <https://doi.org/10.1038/s41588-021-01007-6>.
- Barral, A., Pozo, G., Ducrot, L., Papadopoulos, G.L., Sauzet, S., Oldfield, A.J., Cavalli, G., and Déjardin, J. (2022). SETDB1/NSD-dependent H3K9me3/H3K36me3 dual heterochromatin maintains gene expression profiles by bookmarking poised enhancers. *Mol. Cell* *82*, 816–832.e12. <https://doi.org/10.1016/j.molcel.2021.12.037>.
- Barrios, N., and Campuzano, S. (2016). Expanding the Iroquois genes repertoire: a non-transcriptional function in cell cycle progression. *Fly (Austin)* *9*, 126–131. <https://doi.org/10.1080/19336934.2016.1139654>.
- Barrios, N., González-Pérez, E., Hernández, R., and Campuzano, S. (2015). The Homeodomain Iroquois Proteins Control Cell Cycle Progression and Regulate the Size of Developmental Fields. *PLOS Genet.* *11*, e1005463. <https://doi.org/10.1371/journal.pgen.1005463>.
- Barth, T.K., and Imhof, A. (2010). Fast signals and slow marks: the dynamics of histone modifications. *Trends Biochem. Sci.* *35*, 618–626. <https://doi.org/10.1016/j.tibs.2010.05.006>.
- Bekhite, M.M., and Schulze, P.C. (2021). Human Induced Pluripotent Stem Cell as a Disease Modeling and Drug Development Platform-A Cardiac Perspective. *Cells* *10*, 3483. <https://doi.org/10.3390/cells10123483>.
- Benabdallah, N.S., Williamson, I., Illingworth, R.S., Kane, L., Boyle, S., Sengupta, D., Grimes, G.R., Therizols, P., and Bickmore, W.A. (2019). Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Mol. Cell* *76*, 473–484.e7. <https://doi.org/10.1016/j.molcel.2019.07.038>.
- Bentsen, M., Goymann, P., Schultheis, H., Klee, K., Petrova, A., Wiegandt, R., Fust, A., Preussner, J., Kuenne, C., Braun, T., et al. (2020). ATAC-seq footprinting unravels kinetics of transcription factor binding during zygotic genome activation. *Nat. Commun.* *11*, 4267. <https://doi.org/10.1038/s41467-020-18035-1>.
- Benveniste, D., Sonntag, H.-J., Sanguinetti, G., and Sproul, D. (2014). Transcription factor binding predicts histone modifications in human cell lines. *Proc. Natl. Acad. Sci.* *111*, 13367–13372. <https://doi.org/10.1073/pnas.1412081111>.
- Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Peña-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al. (2008). Variation in homeodomain DNA

binding revealed by high-resolution analysis of sequence preferences. *Cell* 133, 1266–1276. <https://doi.org/10.1016/j.cell.2008.05.024>.

Bergman, D.T., Jones, T.R., Liu, V., Ray, J., Jagoda, E., Siraj, L., Kang, H.Y., Nasser, J., Kane, M., Rios, A., et al. (2022). Compatibility rules of human enhancer and promoter sequences. *Nature* <https://doi.org/10.1038/s41586-022-04877-w>.

Bergmann, O., Bhardwaj, R.D., Bernard, S., Zdunek, S., Barnabé-Heider, F., Walsh, S., Zupicich, J., Alkass, K., Buchholz, B.A., Druid, H., et al. (2009). Evidence for Cardiomyocyte Renewal in Humans. *Science* 324, 98–102. <https://doi.org/10.1126/science.1164680>.

Bergmann, O., Zdunek, S., Felker, A., Salehpour, M., Alkass, K., Bernard, S., Sjöstrom, S.L., Szewczykowska, M., Jackowska, T., dos Remedios, C., et al. (2015). Dynamics of Cell Generation and Turnover in the Human Heart. *Cell* 161, 1566–1575. <https://doi.org/10.1016/j.cell.2015.05.026>.

Bers, D.M. (2002). Cardiac excitation–contraction coupling. *Nature* 415, 198–205. <https://doi.org/10.1038/415198a>.

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M., and Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Dev. Biol.* 353, 266–274. <https://doi.org/10.1016/j.ydbio.2011.02.029>.

Bevilacqua, A., Willis, M.S., and Bultman, S.J. (2014). SWI/SNF chromatin-remodeling complexes in cardiovascular development and disease. *Cardiovasc. Pathol. Off. J. Soc. Cardiovasc. Pathol.* 23, 85–91. <https://doi.org/10.1016/j.carpath.2013.09.003>.

Bhattacharyya, S., and Munshi, N.V. (2020). Development of the Cardiac Conduction System. *Cold Spring Harb. Perspect. Biol.* 12, a037408. <https://doi.org/10.1101/cshperspect.a037408>.

Bilioni, A., Craig, G., Hill, C., and McNeill, H. (2005). Iroquois transcription factors recognize a unique motif to mediate transcriptional repression in vivo. *Proc. Natl. Acad. Sci.* 102, 14671–14676. <https://doi.org/10.1073/pnas.0502480102>.

Bird, A. (2007). Perceptions of epigenetics. *Nature* 447, 396–398. <https://doi.org/10.1038/nature05913>.

Bjune, J.-I., Dyer, L., Røslund, G.V., Tronstad, K.J., Njølstad, P.R., Sagen, J.V., Dankel, S.N., and Mellgren, G. (2020). The homeobox factor *Irx3* maintains adipogenic identity. *Metabolism.* 103, 154014. <https://doi.org/10.1016/j.metabol.2019.154014>.

Boeva, V. (2016). Analysis of Genomic Sequence Motifs for Deciphering Transcription Factor Binding and Transcriptional Regulation in Eukaryotic Cells. *Front. Genet.* 7. .

Bogliotti, Y.S., and Ross, P.J. (2012). Mechanisms of histone H3 lysine 27 trimethylation remodeling during early mammalian development. *Epigenetics* 7, 976–981. <https://doi.org/10.4161/epi.21615>.

Boija, A., Klein, I.A., Sabari, B.R., Dall’Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through

the Phase-Separation Capacity of Their Activation Domains. *Cell* 175, 1842–1855.e16. <https://doi.org/10.1016/j.cell.2018.10.042>.

Bolt, C.C., and Duboule, D. (2020). The regulatory landscapes of developmental genes. *Dev. Camb. Engl.* 147, dev171736. <https://doi.org/10.1242/dev.171736>.

Bonnard, C., Strobl, A.C., Shboul, M., Lee, H., Merriman, B., Nelson, S.F., Ababneh, O.H., Uz, E., Güran, T., Kayserili, H., et al. (2012). Mutations in IRX5 impair craniofacial development and germ cell migration via SDF1. *Nat. Genet.* 44, 709–713. <https://doi.org/10.1038/ng.2259>.

Bosse, A., Zülch, A., Becker, M.-B., Torres, M., Gómez-Skarmeta, J.L., Modolell, J., and Gruss, P. (1997). Identification of the vertebrate Iroquois homeobox gene family with overlapping expression during early development of the nervous system. *Mech. Dev.* 69, 169–181. [https://doi.org/10.1016/S0925-4773\(97\)00165-2](https://doi.org/10.1016/S0925-4773(97)00165-2).

Boukens, B.J., Walton, R., Meijborg, V.M., and Coronel, R. (2016). Transmural electrophysiological heterogeneity, the T-wave and ventricular arrhythmias. *Prog. Biophys. Mol. Biol.* 122, 202–214. <https://doi.org/10.1016/j.pbiomolbio.2016.05.009>.

Boukens, B.J.D., Christoffels, V.M., Coronel, R., and Moorman, A.F.M. (2009). Developmental Basis for Electrophysiological Heterogeneity in the Ventricular and Outflow Tract Myocardium As a Substrate for Life-Threatening Ventricular Arrhythmias. *Circ. Res.* 104, 19–31. <https://doi.org/10.1161/CIRCRESAHA.108.188698>.

Broussard, J.A., Rappaz, B., Webb, D.J., and Brown, C.M. (2013). Fluorescence resonance energy transfer microscopy as demonstrated by measuring the activation of the serine/threonine kinase Akt. *Nat. Protoc.* 8, 265–281. <https://doi.org/10.1038/nprot.2012.147>.

Bruneau, B.G. (2013). Signaling and transcriptional networks in heart development and regeneration. *Cold Spring Harb. Perspect. Biol.* 5, a008292. <https://doi.org/10.1101/cshperspect.a008292>.

Bruneau, B.G., Bao, Z.Z., Tanaka, M., Schott, J.J., Izumo, S., Cepko, C.L., Seidman, J.G., and Seidman, C.E. (2000). Cardiac expression of the ventricle-specific homeobox gene *Irx4* is modulated by *Nkx2-5* and *dHand*. *Dev. Biol.* 217, 266–277. <https://doi.org/10.1006/dbio.1999.9548>.

Bruneau, B.G., Bao, Z.Z., Fatkin, D., Xavier-Neto, J., Georgakopoulos, D., Maguire, C.T., Berul, C.I., Kass, D.A., Kuroski-de Bold, M.L., de Bold, A.J., et al. (2001). Cardiomyopathy in *Irx4*-deficient mice is preceded by abnormal ventricular gene expression. *Mol. Cell. Biol.* 21, 1730–1736. <https://doi.org/10.1128/MCB.21.5.1730-1736.2001>.

Buijtenlijk, M.F.J., Barnett, P., and van den Hoff, M.J.B. (2020). Development of the human heart. *Am. J. Med. Genet. C Semin. Med. Genet.* 184, 7–22. <https://doi.org/10.1002/ajmg.c.31778>.

Bürglin, T.R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. *Nucleic Acids Res.* *25*, 4173–4180. <https://doi.org/10.1093/nar/25.21.4173>.

Bürglin, T.R. (2011). Homeodomain subtypes and functional diversity. *Subcell. Biochem.* *52*, 95–122. https://doi.org/10.1007/978-90-481-9069-0_5.

Cahan, P., and Daley, G.Q. (2013). Origins and implications of pluripotent stem cell variability and heterogeneity. *Nat. Rev. Mol. Cell Biol.* *14*, 357–368. <https://doi.org/10.1038/nrm3584>.

Cai, Y., Zhang, Y., Loh, Y.P., Tng, J.Q., Lim, M.C., Cao, Z., Raju, A., Lieberman Aiden, E., Li, S., Manikandan, L., et al. (2021). H3K27me3-rich genomic regions can function as silencers to repress gene expression via chromatin interactions. *Nat. Commun.* *12*, 719. <https://doi.org/10.1038/s41467-021-20940-y>.

Calle-Mustienes, E. de la, Feijóo, C.G., Manzanares, M., Tena, J.J., Rodríguez-Seguel, E., Letizia, A., Allende, M.L., and Gómez-Skarmeta, J.L. (2005). A functional survey of the enhancer activity of conserved non-coding sequences from vertebrate Iroquois cluster gene deserts. *Genome Res.* *15*, 1061–1072. <https://doi.org/10.1101/gr.4004805>.

Campione, M., Aranega, A., and Franco, D. (2018). Cardiac looping and laterality. *ESC Textb. Cardiovasc. Dev.* *124*. .

Carlson, B.M. (2018). *Human Embryology and Developmental Biology - Inkling Enhanced E-Book* (Elsevier Health Sciences).

Cattaneo, P., Kunderfranco, P., Greco, C., Guffanti, A., Stirparo, G.G., Rusconi, F., Rizzi, R., Di Pasquale, E., Locatelli, S.L., Latronico, M.V.G., et al. (2016). DOT1L-mediated H3K79me2 modification critically regulates gene expression during cardiomyocyte differentiation. *Cell Death Differ.* *23*, 555–564. <https://doi.org/10.1038/cdd.2014.199>.

Cavodeassi, F., Modolell, J., and Gómez-Skarmeta, J.L. (2001). The Iroquois family of genes: from body building to neural patterning. *Dev. Camb. Engl.* *128*, 2847–2855. <https://doi.org/10.1242/dev.128.15.2847>.

Chahal, G., Tyagi, S., and Ramialison, M. (2019). Navigating the non-coding genome in heart development and Congenital Heart Disease. *Differ. Res. Biol. Divers.* *107*, 11–23. <https://doi.org/10.1016/j.diff.2019.05.001>.

Chantalat, S., Depaux, A., Héry, P., Barral, S., Thuret, J.-Y., Dimitrov, S., and Gérard, M. (2011). Histone H3 trimethylation at lysine 36 is associated with constitutive and facultative heterochromatin. *Genome Res.* *21*, 1426–1437. <https://doi.org/10.1101/gr.118091.110>.

Charpentier, E., Cornec, M., Dumont, S., Meistermann, D., Bordron, P., David, L., Redon, R., Bonnaud, S., and Bihouée, A. (2021). 3'RNA sequencing for robust and low-cost gene expression profiling. *Protoc. Exch.* <https://doi.org/10.21203/rs.3.pex-1336/v1>.

- Cheatle Jarvela, A.M., and Hinman, V.F. (2015). Evolution of transcription factor function as a mechanism for changing metazoan developmental gene regulatory networks. *EvoDevo* 6, 3. <https://doi.org/10.1186/2041-9139-6-3>.
- Chen, H., and Pugh, B.F. (2021). What do Transcription Factors Interact With? *J. Mol. Biol.* 433, 166883. <https://doi.org/10.1016/j.jmb.2021.166883>.
- Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer-promoter topology and gene activity. *Nat. Genet.* 50, 1296–1303. <https://doi.org/10.1038/s41588-018-0175-z>.
- Cheng, C.W., Chow, R.L., Lebel, M., Sakuma, R., Cheung, H.O.-L., Thanabalasingham, V., Zhang, X., Bruneau, B.G., Birch, D.G., Hui, C., et al. (2005). The Iroquois homeobox gene, *Ir5*, is required for retinal cone bipolar cell development. *Dev. Biol.* 287, 48–60. <https://doi.org/10.1016/j.ydbio.2005.08.029>.
- Cheng, Z., Wang, J., Su, D., Pan, H., Huang, G., Li, X., Li, Z., Shen, A., Xie, X., Wang, B., et al. (2011). Two novel mutations of the *IRX4* gene in patients with congenital heart disease. *Hum. Genet.* 130, 657–662. <https://doi.org/10.1007/s00439-011-0996-7>.
- Chiapparo, G., Lin, X., Lescroart, F., Chabab, S., Paulissen, C., Pitisci, L., Bondue, A., and Blanpain, C. (2016). *Mesp1* controls the speed, polarity, and directionality of cardiovascular progenitor migration. *J. Cell Biol.* 213, 463–477. <https://doi.org/10.1083/jcb.201505082>.
- Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., Ambartsumyan, G., Aimiwu, O., Richter, L., Zhang, J., et al. (2009). Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell* 5, 111–123. <https://doi.org/10.1016/j.stem.2009.06.008>.
- Cho, J., Lee, H., Rah, W., Chang, H.J., and Yoon, Y.-S. (2022). From engineered heart tissue to cardiac organoid. *Theranostics* 12, 2758–2772. <https://doi.org/10.7150/thno.67661>.
- Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., Choi, I.Y., Ferrari, F., Tsankov, A.M., Pop, R., et al. (2015). A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nat. Biotechnol.* 33, 1173–1181. <https://doi.org/10.1038/nbt.3388>.
- Choquet, C., Boulgakoff, L., Kelly, R.G., and Miquerol, L. (2021). New Insights into the Development and Morphogenesis of the Cardiac Purkinje Fiber Network: Linking Architecture and Function. *J. Cardiovasc. Dev. Dis.* 8, 95. <https://doi.org/10.3390/jcdd8080095>.
- Christoffels, V.M., Keijser, A.G., Houweling, A.C., Clout, D.E., and Moorman, A.F. (2000). Patterning the embryonic heart: identification of five mouse Iroquois homeobox genes in the developing heart. *Dev. Biol.* 224, 263–274. <https://doi.org/10.1006/dbio.2000.9801>.
- Christoffels, V.M., Smits, G.J., Kispert, A., and Moorman, A.F.M. (2010). Development of the pacemaker tissues of the heart. *Circ. Res.* 106, 240–254. <https://doi.org/10.1161/CIRCRESAHA.109.205419>.

Chua, E.H.Z., Yasar, S., and Harmston, N. (2022). The importance of considering regulatory domains in genome-wide analyses - the nearest gene is often wrong! *Biol. Open* *11*, bio059091. <https://doi.org/10.1242/bio.059091>.

Chun, Y.W., Durbin, M.D., and Hong, C.C. (2018). Genome Editing and Induced Pluripotent Stem Cell Technologies for Personalized Study of Cardiovascular Diseases. *Curr. Cardiol. Rep.* *20*, 38. <https://doi.org/10.1007/s11886-018-0984-9>.

Churko, J.M., Garg, P., Treutlein, B., Venkatasubramanian, M., Wu, H., Lee, J., Wessells, Q.N., Chen, S.-Y., Chen, W.-Y., Chetal, K., et al. (2018). Defining human cardiac transcription factor hierarchies using integrated single-cell heterogeneity analysis. *Nat. Commun.* *9*. <https://doi.org/10.1038/s41467-018-07333-4>.

Cohen, D.R., Cheng, C.W., Cheng, S.H., and Hui, C. (2000). Expression of two novel mouse Iroquois homeobox genes during neurogenesis. *Mech. Dev.* *91*, 317–321. [https://doi.org/10.1016/S0925-4773\(99\)00263-4](https://doi.org/10.1016/S0925-4773(99)00263-4).

Condic, M.L. (2014). Totipotency: What It Is and What It Is Not. *Stem Cells Dev.* *23*, 796–812. <https://doi.org/10.1089/scd.2013.0364>.

Core, L., and Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* *33*, 960–982. <https://doi.org/10.1101/gad.325142.119>.

Costantini, D.L., Arruda, E.P., Agarwal, P., Kim, K.-H., Zhu, Y., Zhu, W., Lebel, M., Cheng, C.W., Park, C.Y., Pierce, S.A., et al. (2005). The Homeodomain Transcription Factor *Irx5* Establishes the Mouse Cardiac Ventricular Repolarization Gradient. *Cell* *123*, 347–358. <https://doi.org/10.1016/j.cell.2005.08.004>.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 21931–21936. <https://doi.org/10.1073/pnas.1016071107>.

Crocker, J., Noon, E.P.-B., and Stern, D.L. (2016). The Soft Touch: Low-Affinity Transcription Factor Binding Sites in Development and Evolution. *Curr. Top. Dev. Biol.* *117*, 455–469. <https://doi.org/10.1016/bs.ctdb.2015.11.018>.

Cui, Y., Zhang, X., Yu, M., Zhu, Y., Xing, J., and Lin, J. (2019). Techniques for detecting protein-protein interactions in living cells: principles, limitations, and recent progress. *Sci. China Life Sci.* *62*, 619–632. <https://doi.org/10.1007/s11427-018-9500-7>.

Cunningham, T.J., Yu, M.S., McKeithan, W.L., Spiering, S., Carrette, F., Huang, C.-T., Bushway, P.J., Tierney, M., Albin, S., Giacca, M., et al. (2017). *Id* genes are essential for early heart formation. *Genes Dev.* *31*, 1325–1338. <https://doi.org/10.1101/gad.300400.117>.

D'Amato, G., Luxán, G., and de la Pompa, J.L. (2018). Defining cardiac domains from the inside: NOTCH in endocardial–myocardial interactions. *ESC Textb. Cardiovasc. Dev.* *94*. .

- Dambly-Chaudière, C., and Leyns, L. (1992). The determination of sense organs in *Drosophila*: a search for interacting genes. *Int. J. Dev. Biol.* *36*, 85–91. .
- Davidson, K.C., Mason, E.A., and Pera, M.F. (2015). The pluripotent state in mouse and human. *Development* *142*, 3090–3099. <https://doi.org/10.1242/dev.116061>.
- De Bono, C., Théveniau-Ruissy, M., and Kelly, R.G. (2018). Cardiac fields and myocardial cell lineages. *ESC Textb. Cardiovasc. Dev.* *23*. .
- De Masi, C., Spitalieri, P., Murdocca, M., Novelli, G., and Sangiuolo, F. (2020). Application of CRISPR/Cas9 to human-induced pluripotent stem cells: from gene editing to drug discovery. *Hum. Genomics* *14*, 25. <https://doi.org/10.1186/s40246-020-00276-2>.
- Deans, C., and Maggert, K.A. (2015). What Do You Mean, “Epigenetic”? *Genetics* *199*, 887–896. <https://doi.org/10.1534/genetics.114.173492>.
- Dees, E., and Baldwin, S. (2018). Heart Development. In *Encyclopedia of Cardiovascular Research and Medicine*, R.S. Vasan, and D.B. Sawyer, eds. (Oxford: Elsevier), pp. 380–398.
- Deichmann, U. (2020). The social construction of the social epigenome and the larger biological context. *Epigenetics Chromatin* *13*, 37. <https://doi.org/10.1186/s13072-020-00360-w>.
- Desgrange, A., Le Garrec, J.-F., and Meilhac, S.M. (2018). Left-right asymmetry in heart development and disease: forming the right loop. *Dev. Camb. Engl.* *145*, dev162776. <https://doi.org/10.1242/dev.162776>.
- Di Cerbo, V., Mohn, F., Ryan, D.P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., et al. (2014). Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *ELife* *3*, e01632. <https://doi.org/10.7554/eLife.01632>.
- Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* *485*, 376–380. <https://doi.org/10.1038/nature11082>.
- Dodou, E., Verzi, M.P., Anderson, J.P., Xu, S.-M., and Black, B.L. (2004). Mef2c is a direct transcriptional target of ISL1 and GATA factors in the anterior heart field during mouse embryonic development. *Development* *131*, 3931–3942. <https://doi.org/10.1242/dev.01256>.
- Doğan, E.S., and Liu, C. (2018). Three-dimensional chromatin packing and positioning of plant genomes. *Nat. Plants* *4*, 521–529. <https://doi.org/10.1038/s41477-018-0199-5>.
- Doi, T., Lukošiušė, A., Ruttenstock, E., Dingemann, J., and Puri, P. (2011). Expression of Iroquois genes is up-regulated during early lung development in the nitrofen-induced pulmonary hypoplasia. *J. Pediatr. Surg.* *46*, 62–66. <https://doi.org/10.1016/j.jpedsurg.2010.09.059>.
- Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* *489*, 57–74. <https://doi.org/10.1038/nature11247>.

- Eaton, J.D., and West, S. (2020). Termination of Transcription by RNA Polymerase II: BOOM! Trends Genet. TIG 36, 664–675. <https://doi.org/10.1016/j.tig.2020.05.008>.
- Efrat, S. (2021). Epigenetic Memory: Lessons From iPS Cells Derived From Human β Cells. Front. Endocrinol. 11. .
- Eisner, D.A., Caldwell, J.L., Kistamás, K., and Trafford, A.W. (2017). Calcium and Excitation-Contraction Coupling in the Heart. Circ. Res. 121, 181–195. <https://doi.org/10.1161/CIRCRESAHA.117.310230>.
- Epstein, J.A., and Parmacek, M.S. (2005). Recent Advances in Cardiac Development With Therapeutic Implications for Adult Cardiovascular Disease. Circulation 112, 592–597. <https://doi.org/10.1161/CIRCULATIONAHA.104.479857>.
- Ernst, J., and Kellis, M. (2017). Chromatin-state discovery and genome annotation with ChromHMM. Nat. Protoc. 12, 2478–2492. <https://doi.org/10.1038/nprot.2017.124>.
- Es-Salah-Lamoureux, Z., Jouni, M., Malak, O.A., Belbachir, N., Al Sayed, Z.R., Gandon-Renard, M., Lamirault, G., Gauthier, C., Baró, I., Charpentier, F., et al. (2016). HIV-Tat induces a decrease in IKr and IKs via reduction in phosphatidylinositol-(4,5)-bisphosphate availability. J. Mol. Cell. Cardiol. 99, 1–13. <https://doi.org/10.1016/j.yjmcc.2016.08.022>.
- Even-Faitelson, L., Hassan-Zadeh, V., Baghestani, Z., and Bazett-Jones, D.P. (2016). Coming to terms with chromatin structure. Chromosoma 125, 95–110. <https://doi.org/10.1007/s00412-015-0534-9>.
- Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature 421, 448–453. <https://doi.org/10.1038/nature01411>.
- Fisch, S., Gray, S., Heymans, S., Haldar, S.M., Wang, B., Pfister, O., Cui, L., Kumar, A., Lin, Z., Sen-Banerjee, S., et al. (2007). Kruppel-like factor 15 is a regulator of cardiomyocyte hypertrophy. Proc. Natl. Acad. Sci. 104, 7074–7079. <https://doi.org/10.1073/pnas.0701981104>.
- Flaus, A. (2011). Principles and practice of nucleosome positioning in vitro. Front. Life Sci. 5, 5–27. .
- Frangogiannis, N.G. (2021). Cardiac fibrosis. Cardiovasc. Res. 117, 1450–1488. <https://doi.org/10.1093/cvr/cvaa324>.
- Fu, A., Oberholtzer, S.M., Bagheri-Fam, S., Rastetter, R.H., Holdreith, C., Caceres, V.L., John, S.V., Shaw, S.A., Krentz, K.J., Zhang, X., et al. (2018). Dynamic expression patterns of Irx3 and Irx5 during germline nest breakdown and primordial follicle formation promote follicle survival in mouse ovaries. PLOS Genet. 14, e1007488. <https://doi.org/10.1371/journal.pgen.1007488>.
- Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. 31, 822–826. <https://doi.org/10.1038/nbt.2623>.

- Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* 15, 2038–2049. <https://doi.org/10.1016/j.celrep.2016.04.085>.
- Gaborit, N., Le Bouter, S., Szuts, V., Varro, A., Escande, D., Nattel, S., and Demolombe, S. (2007). Regional and tissue specific transcript signatures of ion channel genes in the non-diseased human heart. *J. Physiol.* 582, 675–693. <https://doi.org/10.1113/jphysiol.2006.126714>.
- Gaborit, N., Varro, A., Le Bouter, S., Szuts, V., Escande, D., Nattel, S., and Demolombe, S. (2010). Gender-related differences in ion-channel and transporter subunit expression in non-diseased human hearts. *J. Mol. Cell. Cardiol.* 49, 639–646. <https://doi.org/10.1016/j.yjmcc.2010.06.005>.
- Gaborit, N., Sakuma, R., Wylie, J.N., Kim, K.-H., Zhang, S.-S., Hui, C.-C., and Bruneau, B.G. (2012). Cooperative and antagonistic roles for *Irx3* and *Irx5* in cardiac morphogenesis and postnatal physiology. *Dev. Camb. Engl.* 139, 4007–4019. <https://doi.org/10.1242/dev.081703>.
- Galang, G., Mandla, R., Ruan, H., Jung, C., Sinha, T., Stone, N.R., Wu, R.S., Mannion, B.J., Allu, P.K.R., Chang, K., et al. (2020). ATAC-Seq Reveals an *Isl1* Enhancer That Regulates Sinoatrial Node Development and Function. *Circ. Res.* 127, 1502–1518. <https://doi.org/10.1161/CIRCRESAHA.120.317145>.
- Galouzis, C.C., and Furlong, E.E.M. (2022). Regulating specificity in enhancer–promoter communication. *Curr. Opin. Cell Biol.* 75, 102065. <https://doi.org/10.1016/j.ceb.2022.01.010>.
- Galupa, R., Nora, E.P., Worsley-Hunt, R., Picard, C., Gard, C., van Bommel, J.G., Servant, N., Zhan, Y., El Marjou, F., Johanneau, C., et al. (2020). A Conserved Noncoding Locus Regulates Random Monoallelic *Xist* Expression across a Topological Boundary. *Mol. Cell* 77, 352–367.e8. <https://doi.org/10.1016/j.molcel.2019.10.030>.
- Garfinkel, A.M., and Khokha, M.K. (2017). An interspecies heart-to-heart: Using *Xenopus* to uncover the genetic basis of congenital heart disease. *Curr. Pathobiol. Rep.* 5, 187–196. <https://doi.org/10.1007/s40139-017-0142-x>.
- Garriock, R.J., Vokes, S.A., Small, E.M., Larson, R., and Krieg, P.A. (2001). Developmental expression of the *Xenopus* Iroquois-family homeobox genes, *Irx4* and *Irx5*. *Dev. Genes Evol.* 211, 257–260. <https://doi.org/10.1007/s004270100147>.
- Gates, L.A., Foulds, C.E., and O’Malley, B.W. (2017). Histone Marks in the “Driver’s Seat”: Functional Roles in Steering the Transcription Cycle. *Trends Biochem. Sci.* 42, 977–989. <https://doi.org/10.1016/j.tibs.2017.10.004>.
- George, R.M., and Firulli, A.B. (2019). Hand Factors in Cardiac Development. *Anat. Rec. Hoboken NJ* 2007 302, 101–107. <https://doi.org/10.1002/ar.23910>.
- Ghavi-Helm, Y., Jankowski, A., Meiers, S., Viales, R.R., Korb, J.O., and Furlong, E.E.M. (2019). Highly rearranged chromosomes reveal uncoupling between genome topology and gene expression. *Nat. Genet.* 51, 1272–1282. <https://doi.org/10.1038/s41588-019-0462-3>.

Ghimire, S., Mantziou, V., Moris, N., and Martinez Arias, A. (2021). Human gastrulation: The embryo and its models. *Dev. Biol.* *474*, 100–108. <https://doi.org/10.1016/j.ydbio.2021.01.006>.

Gibson, B.A., Doolittle, L.K., Schneider, M.W.G., Jensen, L.E., Gamarra, N., Henry, L., Gerlich, D.W., Redding, S., and Rosen, M.K. (2019). Organization of Chromatin by Intrinsic and Regulated Phase Separation. *Cell* *179*, 470–484.e21. <https://doi.org/10.1016/j.cell.2019.08.037>.

Girbig, M., Misiaszek, A.D., and Müller, C.W. (2022). Structural insights into nuclear transcription by eukaryotic DNA-dependent RNA polymerases. *Nat. Rev. Mol. Cell Biol.* <https://doi.org/10.1038/s41580-022-00476-9>.

Golbabapour, S., Majid, N.A., Hassandarvish, P., Hajrezaie, M., Abdulla, M.A., and Hadi, A.H.A. (2013). Gene silencing and Polycomb group proteins: an overview of their structure, mechanisms and phylogenetics. *Omic J. Integr. Biol.* *17*, 283–296. <https://doi.org/10.1089/omi.2012.0105>.

Goldberg, A.D., Allis, C.D., and Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell* *128*, 635–638. <https://doi.org/10.1016/j.cell.2007.02.006>.

Golson, M.L., and Kaestner, K.H. (2016). Fox transcription factors: from development to disease. *Dev. Camb. Engl.* *143*, 4558–4570. <https://doi.org/10.1242/dev.112672>.

Gomez-Skarmeta, J.L., Diez del Corral, R., de la Calle-Mustienes, E., Ferré-Marcó, D., and Modolell, J. (1996). Araucan and caupolican, two members of the novel iroquois complex, encode homeoproteins that control proneural and vein-forming genes. *Cell* *85*, 95–105. [https://doi.org/10.1016/s0092-8674\(00\)81085-5](https://doi.org/10.1016/s0092-8674(00)81085-5).

Gonzalez-Sandoval, A., and Gasser, S.M. (2016). On TADs and LADs: Spatial Control Over Gene Expression. *Trends Genet. TIG* *32*, 485–495. <https://doi.org/10.1016/j.tig.2016.05.004>.

Gonzalez-Teran, B., Pittman, M., Felix, F., Thomas, R., Richmond-Buccola, D., Hüttenhain, R., Choudhary, K., Moroni, E., Costa, M.W., Huang, Y., et al. (2022). Transcription factor protein interactomes reveal genetic determinants in heart disease. *Cell* *185*, 794–814.e30. <https://doi.org/10.1016/j.cell.2022.01.021>.

Groff, A.F., Barutcu, A.R., Lewandowski, J.P., and Rinn, J.L. (2018). Enhancers in the Peril lincRNA locus regulate distant but not local genes. *Genome Biol.* *19*, 219. <https://doi.org/10.1186/s13059-018-1589-8>.

Grosselin, K., Durand, A., Marsolier, J., Poitou, A., Marangoni, E., Nemati, F., Dahmani, A., Lameiras, S., Reyat, F., Frenoy, O., et al. (2019). High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nat. Genet.* *51*, 1060–1066. <https://doi.org/10.1038/s41588-019-0424-9>.

GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* *45*, 580–585. <https://doi.org/10.1038/ng.2653>.

- Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M.B., Talhout, W., Eussen, B.H., de Klein, A., Wessels, L., de Laat, W., et al. (2008). Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* 453, 948–951. <https://doi.org/10.1038/nature06947>.
- Günthel, M., Barnett, P., and Christoffels, V.M. (2018). Development, Proliferation, and Growth of the Mammalian Heart. *Mol. Ther. J. Am. Soc. Gene Ther.* 26, 1599–1609. <https://doi.org/10.1016/j.ymthe.2018.05.022>.
- Guo, Y., and Pu, W.T. (2020). Cardiomyocyte Maturation. *Circ. Res.* 126, 1086–1106. <https://doi.org/10.1161/CIRCRESAHA.119.315862>.
- Guo, C., Wang, Q., Wang, Y., Yang, L., Luo, H., Cao, X.F., An, L., Qiu, Y., Du, M., Ma, X., et al. (2017). Exome sequencing reveals novel IRXI mutation in congenital heart disease. *Mol. Med. Rep.* 15, 3193–3197. <https://doi.org/10.3892/mmr.2017.6410>.
- Gurdon, J.B., Javed, K., Vodnala, M., and Garrett, N. (2020). Long-term association of a transcription factor with its chromatin binding site can stabilize gene expression and cell fate commitment. *Proc. Natl. Acad. Sci.* 117, 15075–15084. <https://doi.org/10.1073/pnas.2000467117>.
- Hall, C., Gehmlich, K., Denning, C., and Pavlovic, D. (2021). Complex Relationship Between Cardiac Fibroblasts and Cardiomyocytes in Health and Disease. *J. Am. Heart Assoc.* 10, e019338. <https://doi.org/10.1161/JAHA.120.019338>.
- Hamamoto, K., and Fukaya, T. (2022). Molecular architecture of enhancer-promoter interaction. *Curr. Opin. Cell Biol.* 74, 62–70. <https://doi.org/10.1016/j.ceb.2022.01.003>.
- Hamamy, H.A., Teebi, A.S., Oudjhane, K., Shegem, N.N., and Ajlouni, K.M. (2007). Severe hypertelorism, midface prominence, prominent/simple ears, severe myopia, borderline intelligence, and bone fragility in two brothers: new syndrome? *Am. J. Med. Genet. A.* 143A, 229–234. <https://doi.org/10.1002/ajmg.a.31594>.
- Hanemaaijer, J., Gregorovicova, M., Nielsen, J.M., Moorman, A.F.M., Wang, T., Planken, R.N., Christoffels, V.M., Sedmera, D., and Jensen, B. (2019). Identification of the building blocks of ventricular septation in monitor lizards (Varanidae). *Dev. Camb. Engl.* 146, dev177121. <https://doi.org/10.1242/dev.177121>.
- Harvey, R.D., and Grant, A.O. (2018). Agents Used in Cardiac Arrhythmias. *Basic Clin. Pharmacol. Chapter 14*, 228. .
- He, A., Gu, F., Hu, Y., Ma, Q., Yi Ye, L., Akiyama, J.A., Visel, A., Pennacchio, L.A., and Pu, W.T. (2014). Dynamic GATA4 enhancers shape the chromatin landscape central to heart development and disease. *Nat. Commun.* 5, 4907. <https://doi.org/10.1038/ncomms5907>.
- He, W., Jia, Y., and Takimoto, K. (2009). Interaction between transcription factors Iroquois proteins 4 and 5 controls cardiac potassium channel Kv4.2 gene transcription. *Cardiovasc. Res.* 81, 64–71. <https://doi.org/10.1093/cvr/cvn259>.

Heicklen-Klein, A., McReynolds, L.J., and Evans, T. (2005). Using the zebrafish model to study GATA transcription factors. *Semin. Cell Dev. Biol.* 16, 95–106. <https://doi.org/10.1016/j.semcdb.2004.10.004>.

Hesson, L.B., Sloane, M.A., Wong, J.W., Nunez, A.C., Srivastava, S., Ng, B., Hawkins, N.J., Bourke, M.J., and Ward, R.L. (2014). Altered promoter nucleosome positioning is an early event in gene silencing. *Epigenetics* 9, 1422–1430. <https://doi.org/10.4161/15592294.2014.970077>.

Hirschenberger, M., Hayn, M., Laliberté, A., Koepke, L., Kirchhoff, F., and Sparrer, K.M.J. (2021). Luciferase reporter assays to monitor interferon signaling modulation by SARS-CoV-2 proteins. *STAR Protoc.* 2, 100781. <https://doi.org/10.1016/j.xpro.2021.100781>.

Hofbauer, P., Jahnel, S.M., Papai, N., Giesshammer, M., Deyett, A., Schmidt, C., Penc, M., Tavernini, K., Grdseloff, N., Meledeth, C., et al. (2021). Cardioids reveal self-organizing principles of human cardiogenesis. *Cell* <https://doi.org/10.1016/j.cell.2021.04.034>.

Holland, P.W., Booth, H.A.F., and Bruford, E.A. (2007). Classification and nomenclature of all human homeobox genes. *BMC Biol.* 5, 47. <https://doi.org/10.1186/1741-7007-5-47>.

Holliday, R. (1994). Epigenetics: an overview. *Dev. Genet.* 15, 453–457. <https://doi.org/10.1002/dvg.1020150602>.

Hsieh, T.-H.S., Cattoglio, C., Slobodyanyuk, E., Hansen, A.S., Rando, O.J., Tjian, R., and Darzacq, X. (2020). Resolving the 3D Landscape of Transcription-Linked Mammalian Chromatin Folding. *Mol. Cell* 78, 539–553.e8. <https://doi.org/10.1016/j.molcel.2020.03.002>.

Hu, W., Xin, Y., Zhang, L., Hu, J., Sun, Y., and Zhao, Y. (2018). Iroquois Homeodomain transcription factors in ventricular conduction system and arrhythmia. *Int. J. Med. Sci.* 15, 808–815. <https://doi.org/10.7150/ijms.25140>.

Huang, K., Maruyama, T., and Fan, G. (2014). The Naive State of Human Pluripotent Stem Cells: A Synthesis of Stem Cell and Preimplantation Embryo Transcriptome Analyses. *Cell Stem Cell* 15, 410–415. <https://doi.org/10.1016/j.stem.2014.09.014>.

Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* 49, e324. <https://doi.org/10.1038/emm.2017.11>.

Ibrahim, D.M., and Mundlos, S. (2020). Three-dimensional chromatin in disease: What holds us together and what drives us apart? *Curr. Opin. Cell Biol.* 64, 1–9. <https://doi.org/10.1016/j.ceb.2020.01.003>.

Igolkina, A.A., Zinkevich, A., Karandasheva, K.O., Popov, A.A., Selifanova, M.V., Nikolaeva, D., Tkachev, V., Penzar, D., Nikitin, D.M., and Buzdin, A. (2019). H3K4me3, H3K9ac, H3K27ac, H3K27me3 and H3K9me3 Histone Tags Suggest Distinct Regulatory Evolution of Open and Condensed Chromatin Landmarks. *Cells* 8, 1034. <https://doi.org/10.3390/cells8091034>.

Jablonka, E., and Lamb, M.J. (2002). The changing concept of epigenetics. *Ann. N. Y. Acad. Sci.* 981, 82–96. <https://doi.org/10.1111/j.1749-6632.2002.tb04913.x>.

- Jaworska-Wilczynska, M., Trzaskoma, P., Szczepankiewicz, A.A., and Hryniewiecki, T. (2016). Pericardium: structure and function in health and disease. *Folia Histochem. Cytobiol.* 54, 121–125. <https://doi.org/10.5603/FHC.a2016.0014>.
- Jensen, E. (2014). Technical Review: In Situ Hybridization. *Anat. Rec.* 297, 1349–1353. <https://doi.org/10.1002/ar.22944>.
- Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316, 1497–1502. <https://doi.org/10.1126/science.1141319>.
- Jorgensen, J.S., and Gao, L. (2005). Irx3 is differentially up-regulated in female gonads during sex determination. *Gene Expr. Patterns* 5, 756–762. <https://doi.org/10.1016/j.modgep.2005.04.011>.
- Joseph, E.M. (2004). Zebrafish IRX1b in the embryonic cardiac ventricle. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* 231, 720–726. <https://doi.org/10.1002/dvdy.20172>.
- Jouni, M., Si-Tayeb, K., Es-Salah-Lamoureux, Z., Latypova, X., Champon, B., Caillaud, A., Rungoat, A., Charpentier, F., Loussouarn, G., Baró, I., et al. (2015). Toward Personalized Medicine: Using Cardiomyocytes Differentiated From Urine-Derived Pluripotent Stem Cells to Recapitulate Electrophysiological Characteristics of Type 2 Long QT Syndrome. *J. Am. Heart Assoc. Cardiovasc. Cerebrovasc. Dis.* 4. <https://doi.org/10.1161/JAHA.115.002159>.
- Jung, G.T., Kim, K.-P., and Kim, K. (2020). How to interpret and integrate multi-omics data at systems level. *Anim. Cells Syst.* 24, 1–7. <https://doi.org/10.1080/19768354.2020.1721321>.
- Junion, G., Spivakov, M., Girardot, C., Braun, M., Gustafson, E.H., Birney, E., and Furlong, E.E.M. (2012). A transcription factor collective defines cardiac cell fate and reflects lineage history. *Cell* 148, 473–486. <https://doi.org/10.1016/j.cell.2012.01.030>.
- Kang, Y., Kim, Y.W., Kang, J., and Kim, A. (2021). Histone H3K4me1 and H3K27ac play roles in nucleosome eviction and eRNA transcription, respectively, at enhancers. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* 35, e21781. <https://doi.org/10.1096/fj.202100488R>.
- Karbassi, E., Fenix, A., Marchiano, S., Muraoka, N., Nakamura, K., Yang, X., and Murry, C.E. (2020). Cardiomyocyte maturation: advances in knowledge and implications for regenerative medicine. *Nat. Rev. Cardiol.* 17, 341–359. <https://doi.org/10.1038/s41569-019-0331-x>.
- Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13, 424. <https://doi.org/10.1186/1471-2164-13-424>.
- Karr, J.P., Ferrie, J.J., Tjian, R., and Darzacq, X. (2022). The transcription factor activity gradient (TAG) model: contemplating a contact-independent mechanism for enhancer-promoter communication. *Genes Dev.* 36, 7–16. <https://doi.org/10.1101/gad.349160.121>.

Kathiriya, I.S., Nora, E.P., and Bruneau, B.G. (2015). Investigating the transcriptional control of cardiovascular development. *Circ. Res.* *116*, 700–714. <https://doi.org/10.1161/CIRCRESAHA.116.302832>.

Kathiriya, I.S., Rao, K.S., Iacono, G., Devine, W.P., Blair, A.P., Hota, S.K., Lai, M.H., Garay, B.I., Thomas, R., Gong, H.Z., et al. (2021). Modeling Human TBX5 Haploinsufficiency Predicts Regulatory Networks for Congenital Heart Disease. *Dev. Cell* *56*, 292-309.e9. <https://doi.org/10.1016/j.devcel.2020.11.020>.

Kehat, I., Kenyagin-Karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-Eldor, J., and Gepstein, L. (2001). Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J. Clin. Invest.* *108*, 407–414. <https://doi.org/10.1172/JCI12131>.

Kelly, R.G., Buckingham, M.E., and Moorman, A.F. (2014). Heart Fields and Cardiac Morphogenesis. *Cold Spring Harb. Perspect. Med.* *4*, a015750. <https://doi.org/10.1101/cshperspect.a015750>.

Kerner, P., Ikmi, A., Coen, D., and Vervoort, M. (2009). Evolutionary history of the iroquois/Irx genes in metazoans. *BMC Evol. Biol.* *9*, 74. <https://doi.org/10.1186/1471-2148-9-74>.

Kester, L., and van Oudenaarden, A. (2018). Single-Cell Transcriptomics Meets Lineage Tracing. *Cell Stem Cell* *23*, 166–179. <https://doi.org/10.1016/j.stem.2018.04.014>.

Kim, K.H. (2012). The Role Of Homeodomain Transcription Factor Irx5 In Cardiac Contractility and Hypertrophic Response. Thesis.

Kim, J., and Kim, H. (2012). Recruitment and Biological Consequences of Histone Modification of H3K27me3 and H3K9me3. *ILAR J.* *53*, 232–239. <https://doi.org/10.1093/ilar.53.3-4.232>.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., Kim, J., Aryee, M., Ji, H., Ehrlich, L., et al. (2010). Epigenetic memory in induced pluripotent stem cells. *Nature* *467*, 285–290. <https://doi.org/10.1038/nature09342>.

Kim, K.-H., Rosen, A., Bruneau, B.G., Hui, C., and Backx, P.H. (2012). Iroquois homeodomain transcription factors in heart development and function. *Circ. Res.* *110*, 1513–1524. <https://doi.org/10.1161/CIRCRESAHA.112.265041>.

Kim, K.-H., Rosen, A., Hussein, S.M.I., Puviindran, V., Korogyi, A.S., Chiarello, C., Nagy, A., Hui, C.-C., and Backx, P.H. (2016). Irx3 is required for postnatal maturation of the mouse ventricular conduction system. *Sci. Rep.* *6*, 19197. <https://doi.org/10.1038/srep19197>.

Kim, K.-H., Oh, Y., Liu, J., Dababneh, S., Xia, Y., Kim, R.Y., Kim, D.-K., Ban, K., Husain, M., Hui, C.-C., et al. (2022). Irx5 and transient outward K⁺ currents contribute to transmural contractile heterogeneities in the mouse ventricle. *Am. J. Physiol. Heart Circ. Physiol.* *322*, H725–H741. <https://doi.org/10.1152/ajpheart.00572.2021>.

- Kim, S., Yu, N.-K., and Kaang, B.-K. (2015). CTCF as a multifunctional protein in genome regulation and gene expression. *Exp. Mol. Med.* 47, e166. <https://doi.org/10.1038/emm.2015.33>.
- Kimura, Y., Aiba, T., Sasano, T., Furukawa, T., Kusano, K., and Shimizu, W. (2016). IRX3 variant as a modifier of Brugada syndrome with frequent ventricular fibrillation. *Hear. Case Rep.* 2, 465–468. <https://doi.org/10.1016/j.hrcr.2016.06.002>.
- Kleber, A.G., and Saffitz, J.E. (2014). Role of the intercalated disc in cardiac propagation and arrhythmogenesis. *Front. Physiol.* 5. .
- Klemm, S.L., Shipony, Z., and Greenleaf, W.J. (2019). Chromatin accessibility and the regulatory epigenome. *Nat. Rev. Genet.* 20, 207–220. <https://doi.org/10.1038/s41576-018-0089-8>.
- Koizumi, A., Sasano, T., Kimura, W., Miyamoto, Y., Aiba, T., Ishikawa, T., Nogami, A., Fukamizu, S., Sakurada, H., Takahashi, Y., et al. (2016). Genetic defects in a His-Purkinje system transcription factor, IRX3, cause lethal cardiac arrhythmias. *Eur. Heart J.* 37, 1469–1475. <https://doi.org/10.1093/eurheartj/ehv449>.
- Krenning, G., Zeisberg, E.M., and Kalluri, R. (2010). The Origin of Fibroblasts and Mechanism of Cardiac Fibrosis. *J. Cell. Physiol.* 225, 631–637. <https://doi.org/10.1002/jcp.22322>.
- Kribelbauer, J.F., Rastogi, C., Bussemaker, H.J., and Mann, R.S. (2019). Low-Affinity Binding Sites and the Transcription Factor Specificity Paradox in Eukaryotes. *Annu. Rev. Cell Dev. Biol.* 35, 357–379. <https://doi.org/10.1146/annurev-cellbio-100617-062719>.
- Krishnan, A., Samtani, R., Dhanantwari, P., Lee, E., Yamada, S., Shiota, K., Donofrio, M.T., Leatherbury, L., and Lo, C.W. (2014). A detailed comparison of mouse and human cardiac development. *Pediatr. Res.* 76, 500–507. <https://doi.org/10.1038/pr.2014.128>.
- Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D.J., Hicks, S.C., Robinson, M.D., Vallejos, C.A., Campbell, K.R., Beerenwinkel, N., Mahfouz, A., et al. (2020). Eleven grand challenges in single-cell data science. *Genome Biol.* 21, 31. <https://doi.org/10.1186/s13059-020-1926-6>.
- Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.-P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: Using Gene-Expression Signatures to Connect Small Molecules, Genes, and Disease. *Science* 313, 1929–1935. <https://doi.org/10.1126/science.1132939>.
- Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. *Cell* 172, 650–665. <https://doi.org/10.1016/j.cell.2018.01.029>.
- Lamberto, F., Peral-Sanchez, I., Muenthaisong, S., Zana, M., Willaime-Morawek, S., and Dinnyés, A. (2021). Environmental Alterations during Embryonic Development: Studying the Impact of Stressors on Pluripotent Stem Cell-Derived Cardiomyocytes. *Genes* 12, 1564. <https://doi.org/10.3390/genes12101564>.

- Lane, J.D., and Tinker, A. (2017). Have the Findings from Clinical Risk Prediction and Trials Any Key Messages for Safety Pharmacology? *Front. Physiol.* *8*, 890. <https://doi.org/10.3389/fphys.2017.00890>.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* *9*, 357–359. <https://doi.org/10.1038/nmeth.1923>.
- Latchman, D.S. (1997). Transcription factors: An overview. *Int. J. Biochem. Cell Biol.* *29*, 1305–1312. [https://doi.org/10.1016/S1357-2725\(97\)00085-X](https://doi.org/10.1016/S1357-2725(97)00085-X).
- Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* *25*, 1–18. .
- Lebel, M., Agarwal, P., Cheng, C.W., Kabir, M.G., Chan, T.Y., Thanabalasingham, V., Zhang, X., Cohen, D.R., Husain, M., Cheng, S.H., et al. (2003). The Iroquois Homeobox Gene *Ir2* Is Not Essential for Normal Development of the Heart and Midbrain-Hindbrain Boundary in Mice. *Mol. Cell. Biol.* *23*, 8216–8225. <https://doi.org/10.1128/MCB.23.22.8216-8225.2003>.
- Lecaudey, V., Anselme, I., Dildrop, R., Rütger, U., and Schneider-Maunoury, S. (2005). Expression of the zebrafish Iroquois genes during early nervous system formation and patterning. *J. Comp. Neurol.* *492*, 289–302. <https://doi.org/10.1002/cne.20765>.
- Lemoine, M.D., and Christ, T. (2021). Human Induced Pluripotent Stem Cell-Derived Cardiomyocytes: The New Working Horse in Cardiovascular Pharmacology? *J. Cardiovasc. Pharmacol.* *77*, 265–266. <https://doi.org/10.1097/FJC.0000000000000981>.
- Leon-Mimila, P., Wang, J., and Huertas-Vazquez, A. (2019). Relevance of Multi-Omics Studies in Cardiovascular Diseases. *Front. Cardiovasc. Med.* *6*. .
- Lescroart, F., and Meilhac, S.M. (2012). Cell lineages, growth and repair of the mouse heart. *Results Probl. Cell Differ.* *55*, 263–289. https://doi.org/10.1007/978-3-642-30406-4_15.
- Lescroart, F., and Zaffran, S. (2018). Hox and Tale transcription factors in heart development and disease. *Int. J. Dev. Biol.* *62*, 837–846. <https://doi.org/10.1387/ijdb.180192sz>.
- Lescroart, F., Chabab, S., Lin, X., Rulands, S., Paulissen, C., Rodolosse, A., Auer, H., Achouri, Y., Dubois, C., Bondue, A., et al. (2014). Early lineage restriction in temporally distinct populations of *Mesp1* progenitors during mammalian heart development. *Nat. Cell Biol.* *16*, 829–840. <https://doi.org/10.1038/ncb3024>.
- Lescroart, F., Wang, X., Lin, X., Swedlund, B., Gargouri, S., Sánchez-Dànes, A., Moignard, V., Dubois, C., Paulissen, C., Kinston, S., et al. (2018). Defining the earliest step of cardiovascular lineage segregation by single-cell RNA-seq. *Science* *359*, 1177–1181. <https://doi.org/10.1126/science.aao4174>.
- Leyns, L., Gómez-Skarmeta, J.L., and Dambly-Chaudière, C. (1996). *iroquois*: a prepatterning gene that controls the formation of bristles on the thorax of *Drosophila*. *Mech. Dev.* *59*, 63–72. [https://doi.org/10.1016/0925-4773\(96\)00577-1](https://doi.org/10.1016/0925-4773(96)00577-1).

- Li, M., and Belmonte, J.C.I. (2018). Deconstructing the pluripotency gene regulatory network. *Nat. Cell Biol.* 20, 382–392. <https://doi.org/10.1038/s41556-018-0067-6>.
- Li, D., Sakuma, R., Vakili, N.A., Mo, R., Puvindran, V., Deimling, S., Zhang, X., Hopyan, S., and Hui, C. (2014a). Formation of Proximal and Anterior Limb Skeleton Requires Early Function of *Irx3* and *Irx5* and Is Negatively Regulated by *Shh* Signaling. *Dev. Cell* 29, 233–240. <https://doi.org/10.1016/j.devcel.2014.03.001>.
- Li, X., Martinez-Fernandez, A., Hartjes, K.A., Kocher, J.-P.A., Olson, T.M., Terzic, A., and Nelson, T.J. (2014b). Transcriptional atlas of cardiogenesis maps congenital heart disease interactome. *Physiol. Genomics* 46, 482–495. <https://doi.org/10.1152/physiolgenomics.00015.2014>.
- Li, Y., Nguyen, H.V., and Tsang, S.H. (2016). Skin Biopsy and Patient-Specific Stem Cell Lines. *Methods Mol. Biol. Clifton NJ* 1353, 77–88. https://doi.org/10.1007/7651_2015_225.
- Liao, Y., Wang, J., Jaehnig, E.J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47, W199–W205. <https://doi.org/10.1093/nar/gkz401>.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293. <https://doi.org/10.1126/science.1181369>.
- Lind, M.I., and Spagopoulou, F. (2018). Evolutionary consequences of epigenetic inheritance. *Heredity* 121, 205–209. <https://doi.org/10.1038/s41437-018-0113-y>.
- Litviňuková, M., Talavera-López, C., Maatz, H., Reichart, D., Worth, C.L., Lindberg, E.L., Kanda, M., Polanski, K., Heinig, M., Lee, M., et al. (2020). Cells of the adult human heart. *Nature* 588, 466–472. <https://doi.org/10.1038/s41586-020-2797-4>.
- Liu, Z., and Tjian, R. (2018). Visualizing transcription factor dynamics in living cells. *J. Cell Biol.* 217, 1181–1191. <https://doi.org/10.1083/jcb.201710038>.
- Liu, C., Feng, X., Li, G., Gokulnath, P., and Xiao, J. (2022). Generating 3D human cardiac constructs from pluripotent stem cells. *EBioMedicine* 76, 103813. <https://doi.org/10.1016/j.ebiom.2022.103813>.
- Liu, H., Ma, W., Xie, J., Li, H., Luo, K., Luo, D., Liu, L., and Sun, X. (2018a). Nucleosome Positioning and Its Role in Gene Regulation in Yeast. *Yeast Role Med. Appl.* 115. .
- Liu, L., Jin, G., and Zhou, X. (2015). Modeling the relationship of epigenetic modifications to transcription factor binding. *Nucleic Acids Res.* 43, 3873–3885. <https://doi.org/10.1093/nar/gkv255>.
- Liu, Y.-W., Chen, B., Yang, X., Fugate, J.A., Kalucki, F.A., Futakuchi-Tsuchida, A., Couture, L., Vogel, K.W., Astley, C.A., Baldessari, A., et al. (2018b). Human embryonic stem cell-derived cardiomyocytes restore function in infarcted hearts of non-human primates. *Nat. Biotechnol.* 36, 597–605. <https://doi.org/10.1038/nbt.4162>.

- Lou, Q., Fedorov, V.V., Glukhov, A.V., Moazami, N., Fast, V.G., and Efimov, I.R. (2011). Transmural Heterogeneity and Remodeling of Ventricular Excitation-Contraction Coupling in Human Heart Failure. *Circulation* *123*, 1881–1890. <https://doi.org/10.1161/CIRCULATIONAHA.110.989707>.
- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* *15*, 550. <https://doi.org/10.1186/s13059-014-0550-8>.
- Lowe, R., Shirley, N., Bleackley, M., Dolan, S., and Shafee, T. (2017). Transcriptomics technologies. *PLOS Comput. Biol.* *13*, e1005457. <https://doi.org/10.1371/journal.pcbi.1005457>.
- Luger, K. (2006). Dynamic nucleosomes. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* *14*, 5–16. <https://doi.org/10.1007/s10577-005-1026-1>.
- Luna-Zurita, L., and Bruneau, B.G. (2013). Chromatin modulators as facilitating factors in cellular reprogramming. *Curr. Opin. Genet. Dev.* *23*, 556–561. <https://doi.org/10.1016/j.gde.2013.07.002>.
- Luna-Zurita, L., Stirnimann, C.U., Glatt, S., Kaynak, B.L., Thomas, S., Baudin, F., Samee, M.A.H., He, D., Small, E.M., Mileikovsky, M., et al. (2016). Complex Interdependence Regulates Heterotypic Transcription Factor Distribution and Coordinates Cardiogenesis. *Cell* *164*, 999–1014. <https://doi.org/10.1016/j.cell.2016.01.004>.
- Luo, M., Bai, J., Liu, B., Yan, P., Zuo, F., Sun, H., Sun, Y., Xu, X., Song, Z., Yang, Y., et al. (2019). H3K18ac Primes Mesendodermal Differentiation upon Nodal Signaling. *Stem Cell Rep.* *13*, 642–656. <https://doi.org/10.1016/j.stemcr.2019.08.016>.
- Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* *161*, 1012–1025. <https://doi.org/10.1016/j.cell.2015.04.004>.
- Ma, J. (2011). Transcriptional activators and activation mechanisms. *Protein Cell* *2*, 879–888. <https://doi.org/10.1007/s13238-011-1101-7>.
- Magdy, T., Schuldt, A.J.T., Wu, J.C., Bernstein, D., and BurrIDGE, P.W. (2018). Human Induced Pluripotent Stem Cell (hiPSC)-Derived Cells to Assess Drug Cardiotoxicity: Opportunities and Problems. *Annu. Rev. Pharmacol. Toxicol.* *58*, 83–103. <https://doi.org/10.1146/annurev-pharmtox-010617-053110>.
- Maston, G.A., Evans, S.K., and Green, M.R. (2006). Transcriptional regulatory elements in the human genome. *Annu. Rev. Genomics Hum. Genet.* *7*, 29–59. <https://doi.org/10.1146/annurev.genom.7.080505.115623>.
- Mattout, A., Biran, A., and Meshorer, E. (2011). Global epigenetic changes during somatic cell reprogramming to iPS cells. *J. Mol. Cell Biol.* *3*, 341–350. <https://doi.org/10.1093/jmcb/mjr028>.

- McCulley, D.J., and Black, B.L. (2012). Transcription factor pathways and congenital heart disease. *Curr. Top. Dev. Biol.* *100*, 253–277. <https://doi.org/10.1016/B978-0-12-387786-4.00008-7>.
- McGinty, R.K., and Tan, S. (2015). Nucleosome Structure and Function. *Chem. Rev.* *115*, 2255–2273. <https://doi.org/10.1021/cr500373h>.
- McKinnon, D., and Rosati, B. (2016). Transmural Gradients in Ion Channel and Auxiliary Subunit Expression. *Prog. Biophys. Mol. Biol.* *122*, 165–186. <https://doi.org/10.1016/j.pbiomolbio.2016.09.012>.
- McKinsey, T.A. (2012). Therapeutic potential for HDAC inhibitors in the heart. *Annu. Rev. Pharmacol. Toxicol.* *52*, 303–319. <https://doi.org/10.1146/annurev-pharmtox-010611-134712>.
- McLean, C.Y., Bristor, D., Hiller, M., Clarke, S.L., Schaar, B.T., Lowe, C.B., Wenger, A.M., and Bejano, G. (2010). GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* *28*, 495–501. <https://doi.org/10.1038/nbt.1630>.
- McNeill, H., Yang, C.H., Brodsky, M., Ungos, J., and Simon, M.A. (1997). mirror encodes a novel PBX-class homeoprotein that functions in the definition of the dorsal-ventral border in the *Drosophila* eye. *Genes Dev.* *11*, 1073–1082. <https://doi.org/10.1101/gad.11.8.1073>.
- Meilhac, S.M. (2018). Cardiac growth I: Cardiomyocyte proliferation. *ESC Textb. Cardiovasc. Dev.* *75*. .
- Meilhac, S.M., and Buckingham, M.E. (2018). The deployment of cell lineages that form the mammalian heart. *Nat. Rev. Cardiol.* *15*, 705–724. <https://doi.org/10.1038/s41569-018-0086-9>.
- Meilhac, S.M., Lescroart, F., Blanpain, C., and Buckingham, M.E. (2014). Cardiac Cell Lineages that Form the Heart. *Cold Spring Harb. Perspect. Med.* *4*, a013888. <https://doi.org/10.1101/cshperspect.a013888>.
- Miquerol, L. (2018). Origin and development of the cardiac conduction system. *ESC Textb. Cardiovasc. Dev.* *136*. .
- Miquerol, L., Bellon, A., Moreno, N., Beyer, S., Meilhac, S.M., Buckingham, M., Franco, D., and Kelly, R.G. (2013). Resolving cell lineage contributions to the ventricular conduction system with a Cx40-GFP allele: a dual contribution of the first and second heart fields. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* *242*, 665–677. <https://doi.org/10.1002/dvdy.23964>.
- Mirzoyan, Z., and Pandur, P. (2013). The Iroquois Complex Is Required in the Dorsal Mesoderm to Ensure Normal Heart Development in *Drosophila*. *PLOS ONE* *8*, e76498. <https://doi.org/10.1371/journal.pone.0076498>.
- Misteli, T. (2020). The Self-Organizing Genome: Principles of Genome Architecture and Function. *Cell* *183*, 28–45. <https://doi.org/10.1016/j.cell.2020.09.014>.

Mitsis, T., Efthimiadou, A., Bacopoulou, F., Vlachakis, D., Chrousos, G.P., and Eliopoulos, E. (2020). Transcription factors and evolution: An integral part of gene expression (Review). *World Acad. Sci. J.* 2, 3–8. <https://doi.org/10.3892/wasj.2020.32>.

Mohan, R.A., Boukens, B.J., and Christoffels, V.M. (2018). Developmental Origin of the Cardiac Conduction System: Insight from Lineage Tracing. *Pediatr. Cardiol.* 39, 1107–1114. <https://doi.org/10.1007/s00246-018-1906-8>.

Mommersteeg, M.T.M., Domínguez, J.N., Wiese, C., Norden, J., de Gier-de Vries, C., Burch, J.B.E., Kispert, A., Brown, N.A., Moorman, A.F.M., and Christoffels, V.M. (2010). The sinus venosus progenitors separate and diversify from the first and second heart fields early in development. *Cardiovasc. Res.* 87, 92–101. <https://doi.org/10.1093/cvr/cvq033>.

Moore-Morris, T., Cattaneo, P., Pucoat, M., and Evans, S.M. (2016). Origins of cardiac fibroblasts. *J. Mol. Cell. Cardiol.* 91, 1–5. <https://doi.org/10.1016/j.yjmcc.2015.12.031>.

Moore-Morris, T., van Vliet, P.P., Andelfinger, G., and Pucoat, M. (2018). Role of Epigenetics in Cardiac Development and Congenital Diseases. *Physiol. Rev.* 98, 2453–2475. <https://doi.org/10.1152/physrev.00048.2017>.

Moorman, A.F.M., and Christoffels, V.M. (2003). Cardiac chamber formation: development, genes, and evolution. *Physiol. Rev.* 83, 1223–1267. <https://doi.org/10.1152/physrev.00006.2003>.

Moorman, A.F., van den Berg, G., Anderson, R.H., and Christoffels, V.M. (2010). Early cardiac growth and the ballooning model of cardiac chamber formation. In *Heart Development and Regeneration*, (Elsevier), pp. 219–236.

Morgan, M.A.J., and Shilatifard, A. (2020). Reevaluating the roles of histone-modifying enzymes and their associated chromatin modifications in transcriptional regulation. *Nat. Genet.* 52, 1271–1281. <https://doi.org/10.1038/s41588-020-00736-4>.

Morrison, O., and Thakur, J. (2021). Molecular Complexes at Euchromatin, Heterochromatin and Centromeric Chromatin. *Int. J. Mol. Sci.* 22, 6922. <https://doi.org/10.3390/ijms22136922>.

Mummenhoff, J., Houweling, A.C., Peters, T., Christoffels, V.M., and Rütger, U. (2001). Expression of *Irx6* during mouse morphogenesis. *Mech. Dev.* 103, 193–195. [https://doi.org/10.1016/s0925-4773\(01\)00353-7](https://doi.org/10.1016/s0925-4773(01)00353-7).

Mummery, C., Ward-van Oostwaard, D., Doevendans, P., Spijker, R., van den Brink, S., Hassink, R., van der Heyden, M., Opthof, T., Pera, M., de la Riviere, A.B., et al. (2003). Differentiation of human embryonic stem cells to cardiomyocytes: role of coculture with visceral endoderm-like cells. *Circulation* 107, 2733–2740. <https://doi.org/10.1161/01.CIR.0000068356.38592.68>.

Muniz, L., Nicolas, E., and Trouche, D. (2021). RNA polymerase II speed: a key player in controlling and adapting transcriptome composition. *EMBO J.* 40, e105740. <https://doi.org/10.15252/embj.2020105740>.

- Muralidhar, S.A., and Sadek, H.A. (2016). Meis1 Regulates Postnatal Cardiomyocyte Cell Cycle Arrest. In *Etiology and Morphogenesis of Congenital Heart Disease: From Gene Function and Cellular Interaction to Morphology*, T. Nakanishi, R.R. Markwald, H.S. Baldwin, B.B. Keller, D. Srivastava, and H. Yamagishi, eds. (Tokyo: Springer), p.
- Nanney, D.L. (1958). EPIGENETIC CONTROL SYSTEMS. *Proc. Natl. Acad. Sci. U. S. A.* *44*, 712–717. <https://doi.org/10.1073/pnas.44.7.712>.
- Nelson, D.O., Jin, D.X., Downs, K.M., Kamp, T.J., and Lyons, G.E. (2014). Irx4 identifies a chamber-specific cell population that contributes to ventricular myocardium development. *Dev. Dyn. Off. Publ. Am. Assoc. Anat.* *243*, 381–392. <https://doi.org/10.1002/dvdy.24078>.
- Nelson, D.O., Lalit, P.A., Biermann, M., Markandeya, Y.S., Capes, D.L., Adesso, L., Patel, G., Han, T., John, M.C., Powers, P.A., et al. (2016). Irx4 Marks a Multipotent, Ventricular-Specific Progenitor Cell. *Stem Cells Dayt. Ohio* *34*, 2875–2888. <https://doi.org/10.1002/stem.2486>.
- Neph, S., Stergachis, A.B., Reynolds, A., Sandstrom, R., Borenstein, E., and Stamatoyannopoulos, J.A. (2012). Circuitry and dynamics of human transcription factor regulatory networks. *Cell* *150*, 1274–1286. <https://doi.org/10.1016/j.cell.2012.04.040>.
- Nerbonne, J.M., and Kass, R.S. (2005). Molecular physiology of cardiac repolarization. *Physiol. Rev.* *85*, 1205–1253. <https://doi.org/10.1152/physrev.00002.2005>.
- Nguyen, P.D., de Bakker, D.E.M., and Bakkers, J. (2021). Cardiac regenerative capacity: an evolutionary afterthought? *Cell. Mol. Life Sci.* *78*, 5107–5122. <https://doi.org/10.1007/s00018-021-03831-9>.
- Nica, A.C., and Dermitzakis, E.T. (2013). Expression quantitative trait loci: present and future. *Philos. Trans. R. Soc. B Biol. Sci.* *368*, 20120362. <https://doi.org/10.1098/rstb.2012.0362>.
- Nicin, L., Wagner, J.U.G., Luxán, G., and Dimmeler, S. (2022). Fibroblast-mediated intercellular crosstalk in the healthy and diseased heart. *FEBS Lett.* *596*, 638–654. <https://doi.org/10.1002/1873-3468.14234>.
- Nora, E.P., Goloborodko, A., Valton, A.-L., Gibcus, J.H., Uebersohn, A., Abdennur, N., Dekker, J., Mirny, L.A., and Bruneau, B.G. (2017). Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* *169*, 930–944.e22. <https://doi.org/10.1016/j.cell.2017.05.004>.
- Noyes, M.B., Christensen, R.G., Wakabayashi, A., Stormo, G.D., Brodsky, M.H., and Wolfe, S.A. (2008). Analysis of homeodomain specificities allows the family-wide prediction of preferred recognition sites. *Cell* *133*, 1277–1289. <https://doi.org/10.1016/j.cell.2008.05.023>.
- O'Donnell, A., and Yutzey, K.E. (2020). Mechanisms of heart valve development and disease. *Development* *147*, dev183020. <https://doi.org/10.1242/dev.183020>.
- Ogura, K., Matsumoto, K., Kuroiwa, A., Isobe, T., Otaguro, T., Jurecic, V., Baldini, A., Matsuda, Y., and Ogura, T. (2001). Cloning and chromosome mapping of human and chicken Iroquois (IRX) genes. *Cytogenet. Cell Genet.* *92*, 320–325. <https://doi.org/10.1159/000056921>.

- Okita, K., Ichisaka, T., and Yamanaka, S. (2007). Generation of germline-competent induced pluripotent stem cells. *Nature* 448, 313–317. <https://doi.org/10.1038/nature05934>.
- Olson, E.N. (2006). Gene Regulatory Networks in the Evolution and Development of the Heart. *Science* 313, 1922–1927. <https://doi.org/10.1126/science.1132292>.
- Oudelaar, A.M., and Higgs, D.R. (2021). The relationship between genome structure and function. *Nat. Rev. Genet.* 22, 154–168. <https://doi.org/10.1038/s41576-020-00303-x>.
- Paige, S.L., Plonowska, K., Xu, A., and Wu, S.M. (2015). Molecular Regulation of Cardiomyocyte Differentiation. *Circ. Res.* 116, 341–353. <https://doi.org/10.1161/CIRCRESAHA.116.302752>.
- Pang, B., and Snyder, M.P. (2020). Systematic identification of silencers in human cells. *Nat. Genet.* 52, 254–263. <https://doi.org/10.1038/s41588-020-0578-5>.
- Panigrahi, A., and O'Malley, B.W. (2021). Mechanisms of enhancer action: the known and the unknown. *Genome Biol.* 22, 108. <https://doi.org/10.1186/s13059-021-02322-1>.
- Park, P.J. (2009). ChIP-Seq: advantages and challenges of a maturing technology. *Nat. Rev. Genet.* 10, 669–680. <https://doi.org/10.1038/nrg2641>.
- Park, D.S., and Fishman, G.I. (2017). Development and Function of the Cardiac Conduction System in Health and Disease. *J. Cardiovasc. Dev. Dis.* 4, 7. <https://doi.org/10.3390/jcdd4020007>.
- Parrotta, E., De Angelis, M.T., Scalise, S., Candeloro, P., Santamaria, G., Paonessa, M., Coluccio, M.L., Perozziello, G., De Vitis, S., Sgura, A., et al. (2017). Two sides of the same coin? Unraveling subtle differences between human embryonic and induced pluripotent stem cells by Raman spectroscopy. *Stem Cell Res. Ther.* 8, 271. <https://doi.org/10.1186/s13287-017-0720-1>.
- Parrotta, E.I., Scalise, S., Taverna, D., De Angelis, M.T., Sarro, G., Gaspari, M., Santamaria, G., and Cuda, G. (2019). Comprehensive proteogenomic analysis of human embryonic and induced pluripotent stem cells. *J. Cell. Mol. Med.* 23, 5440–5453. <https://doi.org/10.1111/jcmm.14426>.
- Parrotta, E.I., Lucchino, V., Scaramuzzino, L., Scalise, S., and Cuda, G. (2020). Modeling Cardiac Disease Mechanisms Using Induced Pluripotent Stem Cell-Derived Cardiomyocytes: Progress, Promises and Challenges. *Int. J. Mol. Sci.* 21, 4354. <https://doi.org/10.3390/ijms21124354>.
- Paul, S. (2021). Histone “acidic patch”: a hotspot in chromatin biology. *The Nucleus* 64, 271–275. .
- Piazza, N., and Wessells, R.J. (2011). Drosophila Models of Cardiac Disease. *Prog. Mol. Biol. Transl. Sci.* 100, 155–210. <https://doi.org/10.1016/B978-0-12-384878-9.00005-4>.
- Piquereau, J., and Ventura-Clapier, R. (2018). Maturation of Cardiac Energy Metabolism During Perinatal Development. *Front. Physiol.* 9, 959. <https://doi.org/10.3389/fphys.2018.00959>.

- Potier, D., Davie, K., Hulselmans, G., Naval Sanchez, M., Haagen, L., Huynh-Thu, V.A., Koldere, D., Celik, A., Geurts, P., Christiaens, V., et al. (2014a). Mapping gene regulatory networks in *Drosophila* eye development by large-scale transcriptome perturbations and motif inference. *Cell Rep.* *9*, 2290–2303. <https://doi.org/10.1016/j.celrep.2014.11.038>.
- Potier, D., Seyres, D., Guichard, C., Iche-Torres, M., Aerts, S., Herrmann, C., and Perrin, L. (2014b). Identification of cis-regulatory modules encoding temporal dynamics during development. *BMC Genomics* *15*, 534. <https://doi.org/10.1186/1471-2164-15-534>.
- Pozo, M.R., Meredith, G.W., and Entcheva, E. (2022). Human iPSC-Cardiomyocytes as an Experimental Model to Study Epigenetic Modifiers of Electrophysiology. *Cells* *11*, 200. <https://doi.org/10.3390/cells11020200>.
- Pradeepa, M.M., Grimes, G.R., Kumar, Y., Olley, G., Taylor, G.C.A., Schneider, R., and Bickmore, W.A. (2016). Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat. Genet.* *48*, 681–686. <https://doi.org/10.1038/ng.3550>.
- Probst, S., and Arnold, S.J. (2017). Eomesodermin-At Dawn of Cell Fate Decisions During Early Embryogenesis. *Curr. Top. Dev. Biol.* *122*, 93–115. <https://doi.org/10.1016/bs.ctdb.2016.09.001>.
- Prosdocimo, D.A., Sabeh, M.K., and Jain, M.K. (2015). Kruppel-Like Factors in Muscle Health and Disease. *Trends Cardiovasc. Med.* *25*, 278–287. <https://doi.org/10.1016/j.tcm.2014.11.006>.
- Protze, S.I., Lee, J.H., and Keller, G.M. (2019). Human Pluripotent Stem Cell-Derived Cardiovascular Cells: From Developmental Biology to Therapeutic Applications. *Cell Stem Cell* *25*, 311–327. <https://doi.org/10.1016/j.stem.2019.07.010>.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* *26*, 841–842. <https://doi.org/10.1093/bioinformatics/btq033>.
- Radman-Livaja, M., and Rando, O.J. (2010). Nucleosome positioning: how is it established, and why does it matter? *Dev. Biol.* *339*, 258–266. <https://doi.org/10.1016/j.ydbio.2009.06.012>.
- Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* *44*, W160–W165. <https://doi.org/10.1093/nar/gkw257>.
- Rassek, K., and Izykowska, K. (2020). Single-Cell Heterogeneity of Cutaneous T-Cell Lymphomas Revealed Using RNA-Seq Technologies. *Cancers* *12*, 2129. <https://doi.org/10.3390/cancers12082129>.
- Razin, S.V., and Ulianov, S.V. (2017). Gene functioning and storage within a folded genome. *Cell. Mol. Biol. Lett.* *22*, 18. <https://doi.org/10.1186/s11658-017-0050-4>.

Razin, S.V., Gavrilov, A.A., Vassetzky, Y.S., and Ulianov, S.V. (2016). Topologically-associating domains: gene warehouses adapted to serve transcriptional regulation. *Transcription* 7, 84–90. <https://doi.org/10.1080/21541264.2016.1181489>.

Reddy, K.L., Zullo, J.M., Bertolino, E., and Singh, H. (2008). Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* 452, 243–247. <https://doi.org/10.1038/nature06727>.

Regadas, I., Dahlberg, O., Vaid, R., Ho, O., Belikov, S., Dixit, G., Deindl, S., Wen, J., and Mannervik, M. (2021). A unique histone 3 lysine 14 chromatin signature underlies tissue-specific gene regulation. *Mol. Cell* 81, 1766-1780.e10. <https://doi.org/10.1016/j.molcel.2021.01.041>.

Ren, J., Miao, D., Li, Y., and Gao, R. (2021). Spotlight on Isl1: A Key Player in Cardiovascular Development and Diseases. *Front. Cell Dev. Biol.* 9, 793605. <https://doi.org/10.3389/fcell.2021.793605>.

Reubinoff, B.E., Pera, M.F., Fong, C.Y., Trounson, A., and Bongso, A. (2000). Embryonic stem cell lines from human blastocysts: somatic differentiation in vitro. *Nat. Biotechnol.* 18, 399–404. <https://doi.org/10.1038/74447>.

Richter, F., Morton, S.U., Won Kim, S., Kitaygorodsky, A., Wasson, L.K., Chen, K.M., Zhou, J., Qi, H., Patel, N., DePalma, S.R., et al. (2020). Genomic analyses implicate noncoding de novo variants in congenital heart disease. *Nat. Genet.* 52, 769–777. <https://doi.org/10.1038/s41588-020-0652-z>.

Robson, M.I., Ringel, A.R., and Mundlos, S. (2019). Regulatory Landscaping: How Enhancer-Promoter Communication Is Sculpted in 3D. *Mol. Cell* 74, 1110–1122. <https://doi.org/10.1016/j.molcel.2019.05.032>.

Rodriguez, J., and Larson, D.R. (2020). Transcription in Living Cells: Molecular Mechanisms of Bursting. *Annu. Rev. Biochem.* 89, 189–212. <https://doi.org/10.1146/annurev-biochem-011520-105250>.

Ronaldson-Bouchard, K., Ma, S.P., Yeager, K., Chen, T., Song, L., Sirabella, D., Morikawa, K., Teles, D., Yazawa, M., and Vunjak-Novakovic, G. (2018). Advanced maturation of human cardiac tissue grown from pluripotent stem cells. *Nature* 556, 239–243. <https://doi.org/10.1038/s41586-018-0016-3>.

Rosati, B., Pan, Z., Lypen, S., Wang, H.S., Cohen, I., Dixon, J.E., and McKinnon, D. (2001). Regulation of KChIP2 potassium channel beta subunit gene expression underlies the gradient of transient outward current in canine and human ventricle. *J. Physiol.* 533, 119–125. <https://doi.org/10.1111/j.1469-7793.2001.0119b.x>.

Rossant, J., and Tam, P.P.L. (2017). New Insights into Early Human Development: Lessons for Stem Cell Derivation and Differentiation. *Cell Stem Cell* 20, 18–28. <https://doi.org/10.1016/j.stem.2016.12.004>.

- Rossant, J., and Tam, P.P.L. (2022). Early human embryonic development: Blastocyst formation to gastrulation. *Dev. Cell* 57, 152–165. <https://doi.org/10.1016/j.devcel.2021.12.022>.
- Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R.G., Kwon, C., and Lutolf, M.P. (2021). Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell* 28, 230–240.e6. <https://doi.org/10.1016/j.stem.2020.10.013>.
- Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L., et al. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 409, 928–933. <https://doi.org/10.1038/35057149>.
- Sadler, T.W., and Langman, J. (2012). *Langman’s medical embryology* (Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins).
- Sanulli, S., Trnka, M.J., Dharmarajan, V., Tibble, R.W., Pascal, B.D., Burlingame, A.L., Griffin, P.R., Gross, J.D., and Narlikar, G.J. (2019). HP1 reshapes nucleosome core to promote phase separation of heterochromatin. *Nature* 575, 390–394. <https://doi.org/10.1038/s41586-019-1669-2>.
- Schmitt, A.D., Hu, M., Jung, I., Xu, Z., Qiu, Y., Tan, C.L., Li, Y., Lin, S., Lin, Y., Barr, C.L., et al. (2016). A Compendium of Chromatin Contact Maps Reveal Spatially Active Regions in the Human Genome. *Cell Rep.* 17, 2042–2059. <https://doi.org/10.1016/j.celrep.2016.10.061>.
- Schoenwolf, G.C., Bleyl, S.B., Brauer, P.R., and Francis-West, P.H. (2014). *Larsen’s Human Embryology* (Elsevier Health Sciences).
- Schones, D.E., Cui, K., Cuddapah, S., Roh, T.-Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic Regulation of Nucleosome Positioning in the Human Genome. *Cell* 132, 887–898. <https://doi.org/10.1016/j.cell.2008.02.022>.
- Schuettengruber, B., Martinez, A.-M., Iovino, N., and Cavalli, G. (2011). Trithorax group proteins: switching genes on and keeping them active. *Nat. Rev. Mol. Cell Biol.* 12, 799–814. <https://doi.org/10.1038/nrm3230>.
- Schuettengruber, B., Bourbon, H.-M., Di Croce, L., and Cavalli, G. (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* 171, 34–57. <https://doi.org/10.1016/j.cell.2017.08.002>.
- Schwarzer, W., Abdennur, N., Goloborodko, A., Pekowska, A., Fudenberg, G., Loe-Mie, Y., Fonseca, N.A., Huber, W., Haering, C.H., Mirny, L., et al. (2017). Two independent modes of chromatin organization revealed by cohesin removal. *Nature* 551, 51–56. <https://doi.org/10.1038/nature24281>.
- Sheng, J.-J., and Jin, J.-P. (2014). Gene regulation, alternative splicing, and posttranslational modification of troponin subunits in cardiac development and adaptation: a focused review. *Front. Physiol.* 5. .

- Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* 357, eaaf4382. <https://doi.org/10.1126/science.aaf4382>.
- Si-Tayeb, K., Noto, F.K., Nagaoka, M., Li, J., Battle, M.A., Duris, C., North, P.E., Dalton, S., and Duncan, S.A. (2010). Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* 51, 297–305. <https://doi.org/10.1002/hep.23354>.
- Son, J.E., Dou, Z., Kim, K.-H., Wanggou, S., Cha, V.S.B., Mo, R., Zhang, X., Chen, X., Ketela, T., Li, X., et al. (2021). *Irx3* and *Irx5* in *Ins2-Cre+* cells regulate hypothalamic postnatal neurogenesis and leptin response. *Nat. Metab.* 3, 701–713. <https://doi.org/10.1038/s42255-021-00382-y>.
- Spitz, F., and Furlong, E.E.M. (2012). Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* 13, 613–626. <https://doi.org/10.1038/nrg3207>.
- Spivakov, M. (2014). Spurious transcription factor binding: Non-functional or genetically redundant? *Bioessays* 36, 798–806. <https://doi.org/10.1002/bies.201400036>.
- Srivastava, D. (2006). Making or breaking the heart: from lineage determination to morphogenesis. *Cell* 126, 1037–1048. <https://doi.org/10.1016/j.cell.2006.09.003>.
- van Steensel, B., and Furlong, E.E.M. (2019). The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* 20, 327–337. <https://doi.org/10.1038/s41580-019-0114-6>.
- Stefanovic, S., and Zaffran, S. (2017). Mechanisms of retinoic acid signaling during cardiogenesis. *Mech. Dev.* 143, 9–19. <https://doi.org/10.1016/j.mod.2016.12.002>.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45. <https://doi.org/10.1038/47412>.
- Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. *Nature* 547, 241–245. <https://doi.org/10.1038/nature22989>.
- Sun, Z., Zhang, Y., Jia, J., Fang, Y., Tang, Y., Wu, H., and Fang, D. (2020). H3K36me3, message from chromatin to DNA damage repair. *Cell Biosci.* 10, 9. <https://doi.org/10.1186/s13578-020-0374-z>.
- Supek, F., Bošnjak, M., Škunca, N., and Šmuc, T. (2011). REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms. *PLOS ONE* 6, e21800. <https://doi.org/10.1371/journal.pone.0021800>.
- Swedlund, B., and Lescroart, F. (2020). Cardiopharyngeal Progenitor Specification: Multiple Roads to the Heart and Head Muscles. *Cold Spring Harb. Perspect. Biol.* 12, a036731. <https://doi.org/10.1101/cshperspect.a036731>.
- Sylva, M., Hoff, M.J.B. van den, and Moorman, A.F.M. (2014). Development of the human heart. *Am. J. Med. Genet. A.* 164, 1347–1371. <https://doi.org/10.1002/ajmg.a.35896>.

- Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* *47*, D607–D613. <https://doi.org/10.1093/nar/gky1131>.
- Tai, Y.C. (2022). *timecourse: Statistical Analysis for Developmental Microarray Time Course Data* (Bioconductor version: Release (3.14)).
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* *126*, 663–676. <https://doi.org/10.1016/j.cell.2006.07.024>.
- Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell* *131*, 861–872. <https://doi.org/10.1016/j.cell.2007.11.019>.
- Takeuchi, J.K., and Bruneau, B.G. (2009). Directed transdifferentiation of mouse mesoderm to heart tissue by defined factors. *Nature* *459*, 708–711. <https://doi.org/10.1038/nature08039>.
- Tan, C.M.J., and Lewandowski, A.J. (2020). The Transitional Heart: From Early Embryonic and Fetal Development to Neonatal Life. *Fetal Diagn. Ther.* *47*, 373–386. <https://doi.org/10.1159/000501906>.
- Tan, J.T., Korzh, V., and Gong, Z. (1999). Expression of a zebrafish iroquois homeobox gene, *Ziro3*, in the midline axial structures and central nervous system. *Mech. Dev.* *87*, 165–168. [https://doi.org/10.1016/s0925-4773\(99\)00121-5](https://doi.org/10.1016/s0925-4773(99)00121-5).
- Taylor, D., Linnarsson, S., Aronow, B.J., Bader, G.D., Barker, R.A., Camara, P.G., Camp, J.G., Chédotal, A., Copp, A., Etchevers, H.C., et al. (2021). A roadmap for the Human Developmental Cell Atlas. *Nature* *597*, 196–205. <https://doi.org/10.1038/s41586-021-03620-1>.
- Tena, J.J., and Santos-Pereira, J.M. (2021). Topologically Associating Domains and Regulatory Landscapes in Development, Evolution and Disease. *Front. Cell Dev. Biol.* *9*.
- Tena, J.J., Alonso, M.E., de la Calle-Mustienes, E., Splinter, E., de Laat, W., Manzanares, M., and Gómez-Skarmeta, J.L. (2011). An evolutionarily conserved three-dimensional structure in the vertebrate *lrx* clusters facilitates enhancer sharing and coregulation. *Nat. Commun.* *2*, 310. <https://doi.org/10.1038/ncomms1301>.
- Thesleff, I. (1995). Homeobox genes and growth factors in regulation of craniofacial and tooth morphogenesis. *Acta Odontol. Scand.* *53*, 129–134. <https://doi.org/10.3109/00016359509005962>.
- Thiery, J.P., Acloque, H., Huang, R.Y.J., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions in development and disease. *Cell* *139*, 871–890. <https://doi.org/10.1016/j.cell.2009.11.007>.

Thompson, D., Regev, A., and Roy, S. (2015). Comparative Analysis of Gene Regulatory Networks: From Network Reconstruction to Evolution. *Annu. Rev. Cell Dev. Biol.* *31*, 399–428. <https://doi.org/10.1146/annurev-cellbio-100913-012908>.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* *282*, 1145–1147. <https://doi.org/10.1126/science.282.5391.1145>.

Tian, P., Wang, J., Shen, X., Rey, J.F., Yuan, Q., and Yan, Y. (2017). Fundamental CRISPR-Cas9 tools and current applications in microbial systems. *Synth. Syst. Biotechnol.* *2*, 219–225. <https://doi.org/10.1016/j.synbio.2017.08.006>.

Tucker, N.R., Chaffin, M., Fleming, S.J., Hall, A.W., Parsons, V.A., Bedi, K.C., Akkad, A.-D., Herndon, C.N., Arduini, A., Papangelis, I., et al. (2020). Transcriptional and Cellular Diversity of the Human Heart. *Circulation* *142*, 466–482. <https://doi.org/10.1161/CIRCULATIONAHA.119.045401>.

Ulianov, S.V., Khrameeva, E.E., Gavrillov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M.V., Chertovich, A., et al. (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.* *26*, 70–84. <https://doi.org/10.1101/gr.196006.115>.

Välimäki, M.J., and Ruskoaho, H.J. (2020). Targeting GATA4 for cardiac repair. *iubmb Life* *72*, 68–79. <https://doi.org/10.1002/iub.2150>.

Van Campenhout, M.J.H., Yaksh, A., Kik, C., de Jaegere, P.P., Ho, S.Y., Allesie, M.A., and de Groot, N.M.S. (2013). Bachmann's Bundle. *Circ. Arrhythm. Electrophysiol.* *6*, 1041–1046. <https://doi.org/10.1161/CIRCEP.113.000758>.

Van den Aemele, J., Tiberi, L., Bondue, A., Paulissen, C., Herpoel, A., Iacovino, M., Kyba, M., Blanpain, C., and Vanderhaeghen, P. (2012). Eomesodermin induces *Mesp1* expression and cardiac differentiation from embryonic stem cells in the absence of Activin. *EMBO Rep.* *13*, 355–362. <https://doi.org/10.1038/embor.2012.23>.

Van Tuyl, M., Liu, J., Groenman, F., Ridsdale, R., Han, R.N.N., Venkatesh, V., Tibboel, D., and Post, M. (2006). Iroquois genes influence proximo-distal morphogenesis during rat lung development. *Am. J. Physiol. Lung Cell. Mol. Physiol.* *290*, L777–L789. <https://doi.org/10.1152/ajplung.00293.2005>.

VanOudenhove, J., Yankee, T.N., Wilderman, A., and Cotney, J. (2020). Epigenomic and Transcriptomic Dynamics During Human Heart Organogenesis. *Circ. Res.* *127*, e184–e209. <https://doi.org/10.1161/CIRCRESAHA.120.316704>.

Vazin, T., and Freed, W.J. (2010). Human embryonic stem cells: Derivation, culture, and differentiation: A review. *Restor. Neurol. Neurosci.* *28*, 589–603. <https://doi.org/10.3233/RNN-2010-0543>.

Veerman, C.C., Podliesna, S., Tadros, R., Lodder, E.M., Mengarelli, I., de Jonge, B., Beekman, L., Barc, J., Wilders, R., Wilde, A.A.M., et al. (2017). The Brugada Syndrome Susceptibility Gene

- HEY2 Modulates Cardiac Transmural Ion Channel Patterning and Electrical Heterogeneity. *Circ. Res.* *121*, 537–548. <https://doi.org/10.1161/CIRCRESAHA.117.310959>.
- Voss, T.C., and Hager, G.L. (2014). Dynamic regulation of transcriptional states by chromatin and transcription factors. *Nat. Rev. Genet.* *15*, 69–81. <https://doi.org/10.1038/nrg3623>.
- Waddington, C.H. (2012). The epigenotype. 1942. *Int. J. Epidemiol.* *41*, 10–13. <https://doi.org/10.1093/ije/dyr184>.
- Wagner, E.J., and Carpenter, P.B. (2012). Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.* *13*, 115–126. <https://doi.org/10.1038/nrm3274>.
- Wamaitha, S.E., and Niakan, K.K. (2018). Human Pre-gastrulation Development. *Curr. Top. Dev. Biol.* *128*, 295–338. <https://doi.org/10.1016/bs.ctdb.2017.11.004>.
- Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., et al. (2012). Dynamic and Coordinated Epigenetic Regulation of Developmental Transitions in the Cardiac Lineage. *Cell* *151*, 206–220. <https://doi.org/10.1016/j.cell.2012.07.035>.
- Wang, G.F., Nikovits, W., Bao, Z.Z., and Stockdale, F.E. (2001). Irx4 forms an inhibitory complex with the vitamin D and retinoic X receptors to regulate cardiac chamber-specific slow MyHC3 expression. *J. Biol. Chem.* *276*, 28835–28841. <https://doi.org/10.1074/jbc.M103716200>.
- Wang, Q., Jia, Y., Wang, Y., Jiang, Z., Zhou, X., Zhang, Z., Nie, C., Li, J., Yang, N., and Qu, L. (2019). Evolution of cis- and trans-regulatory divergence in the chicken genome between two contrasting breeds analyzed using three tissue types at one-day-old. *BMC Genomics* *20*, 933. <https://doi.org/10.1186/s12864-019-6342-5>.
- Wang, Z., Cui, M., Shah, A.M., Tan, W., Liu, N., Bassel-Duby, R., and Olson, E.N. (2020). Cell-Type-Specific Gene Regulatory Networks Underlying Murine Neonatal Heart Regeneration at Single-Cell Resolution. *Cell Rep.* *33*, 108472. <https://doi.org/10.1016/j.celrep.2020.108472>.
- Warrier, S., Van der Jeught, M., Duggal, G., Tilleman, L., Sutherland, E., Taelman, J., Popovic, M., Lierman, S., Chuva De Sousa Lopes, S., Van Soom, A., et al. (2017). Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nat. Commun.* *8*, 15055. <https://doi.org/10.1038/ncomms15055>.
- Weber, C.M., Ramachandran, S., and Henikoff, S. (2014). Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. *Mol. Cell* *53*, 819–830. <https://doi.org/10.1016/j.molcel.2014.02.014>.
- van Weerd, J.H., and Christoffels, V.M. (2016). The formation and function of the cardiac conduction system. *Development* *143*, 197–210. <https://doi.org/10.1242/dev.124883>.
- Wen, B., Jung, H.J., Chen, L., Saeed, F., and Knepper, M.A. (2020). NGS-Integrator: An efficient tool for combining multiple NGS data tracks using minimum Bayes' factors. *BMC Genomics* *21*, 806. <https://doi.org/10.1186/s12864-020-07220-7>.

Whitfield, T.W., Wang, J., Collins, P.J., Partridge, E.C., Aldred, S.F., Trinklein, N.D., Myers, R.M., and Weng, Z. (2012). Functional analysis of transcription factor binding sites in human promoters. *Genome Biol.* *13*, R50. <https://doi.org/10.1186/gb-2012-13-9-r50>.

van Wijk, B., Barnett, P., and van den Hoff, M.J. (2018). The developmental origin of myocardium at the venous pole of the heart. *ESC Textb. Cardiovasc. Dev.* *64*. .

Williams, L.A., Davis-Dusenbery, B.N., and Eggan, K.C. (2012). SnapShot: Directed Differentiation of Pluripotent Stem Cells. *Cell* *149*, 1174-1174.e1. <https://doi.org/10.1016/j.cell.2012.05.015>.

Williams, T.L., Colzani, M.T., Macrae, R.G.C., Robinson, E.L., Bloor, S., Greenwood, E.J.D., Zhan, J.R., Strachan, G., Kuc, R.E., Nyimanu, D., et al. (2021). Human embryonic stem cell-derived cardiomyocyte platform screens inhibitors of SARS-CoV-2 infection. *Commun. Biol.* *4*, 926. <https://doi.org/10.1038/s42003-021-02453-y>.

Wittig, J.G., and Münsterberg, A. (2020). The Chicken as a Model Organism to Study Heart Development. *Cold Spring Harb. Perspect. Biol.* *12*, a037218. <https://doi.org/10.1101/cshperspect.a037218>.

Wittkopp, P.J., and Kalay, G. (2012). Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* *13*, 59–69. <https://doi.org/10.1038/nrg3095>.

Wu, H., Lee, J., Vincent, L.G., Wang, Q., Gu, M., Lan, F., Churko, J.M., Sallam, K.I., Matsa, E., Sharma, A., et al. (2015). Epigenetic Regulation of Phosphodiesterases 2A and 3A Underlies Compromised β -Adrenergic Signaling in an iPSC Model of Dilated Cardiomyopathy. *Cell Stem Cell* *17*, 89–100. <https://doi.org/10.1016/j.stem.2015.04.020>.

Wu, P., Deng, G., Sai, X., Guo, H., Huang, H., and Zhu, P. (2021a). Maturation strategies and limitations of induced pluripotent stem cell-derived cardiomyocytes. *Biosci. Rep.* *41*, BSR20200833. <https://doi.org/10.1042/BSR20200833>.

Wu, S.-H. (Sam), Lee, J.-H., and Koo, B.-K. (2019). Lineage Tracing: Computational Reconstruction Goes Beyond the Limit of Imaging. *Mol. Cells* *42*, 104–112. <https://doi.org/10.14348/molcells.2019.0006>.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021b). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. N. Y. N* *2*, 100141. <https://doi.org/10.1016/j.xinn.2021.100141>.

Xia, J., Meng, Z., Ruan, H., Yin, W., Xu, Y., and Zhang, T. (2020). Heart Development and Regeneration in Non-mammalian Model Organisms. *Front. Cell Dev. Biol.* *8*. .

Yamagishi, H. (2021). Cardiac Neural Crest. *Cold Spring Harb. Perspect. Biol.* *13*, a036715. <https://doi.org/10.1101/cshperspect.a036715>.

Ye, W., Wang, J., Song, Y., Yu, D., Sun, C., Liu, C., Chen, F., Zhang, Y., Wang, F., Harvey, R.P., et al. (2015). A common Shox2-Nkx2-5 antagonistic mechanism primes the pacemaker cell fate

- in the pulmonary vein myocardium and sinoatrial node. *Dev. Camb. Engl.* *142*, 2521–2532. <https://doi.org/10.1242/dev.120220>.
- Yeo, J.-C., and Ng, H.-H. (2013). The transcriptional regulation of pluripotency. *Cell Res.* *23*, 20–32. <https://doi.org/10.1038/cr.2012.172>.
- Yi, T., Wong, J., Feller, E., Sink, S., Taghli-Lamalle, O., Wen, J., Kim, C., Fink, M., Giles, W., Soussou, W., et al. (2012). Electrophysiological mapping of embryonic mouse hearts: mechanisms for developmental pacemaker switch and internodal conduction pathway. *J. Cardiovasc. Electrophysiol.* *23*, 309–318. <https://doi.org/10.1111/j.1540-8167.2011.02191.x>.
- Yoshida, Y., and Yamanaka, S. (2017). Induced Pluripotent Stem Cells 10 Years Later: For Cardiac Applications. *Circ. Res.* *120*, 1958–1968. <https://doi.org/10.1161/CIRCRESAHA.117.311080>.
- Young, R.A. (2011). Control of the embryonic stem cell state. *Cell* *144*, 940–954. <https://doi.org/10.1016/j.cell.2011.01.032>.
- Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. *Science* *318*, 1917–1920. <https://doi.org/10.1126/science.1151526>.
- Yu, Y., Ma, J., Xiao, Y., Yang, Q., Kang, H., Zhen, J., Yu, L., and Chen, L. (2015). KLF15 is an essential negative regulatory factor for the cardiac remodeling response to pressure overload. *Cardiology* *130*, 143–152. <https://doi.org/10.1159/000369382>.
- Zabet, N.R. (2012). System size reduction in stochastic simulations of the facilitated diffusion mechanism. *BMC Syst. Biol.* *6*, 121. <https://doi.org/10.1186/1752-0509-6-121>.
- Zaffran, S., and Frasch, M. (2002). Early Signals in Cardiac Development. *Circ. Res.* *91*, 457–469. <https://doi.org/10.1161/01.RES.0000034152.74523.A8>.
- Zaffran, S., Odelin, G., Stefanovic, S., Lescroart, F., and Etchevers, H.C. (2018). Ectopic expression of *Hoxb1* induces cardiac and craniofacial malformations. *Genes*. N. Y. N *2000* *56*, e23221. <https://doi.org/10.1002/dvg.23221>.
- 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. <https://doi.org/10.1038/nature12141>.
- Zeng, L., Gu, N., Chen, J., Jin, G., and Zheng, Y. (2019). IRX1 hypermethylation promotes heart failure by inhibiting CXCL14 expression. *Cell Cycle Georget. Tex* *18*, 3251–3262. <https://doi.org/10.1080/15384101.2019.1673099>.
- Zhai, J., Xiao, Z., Wang, Y., and Wang, H. (2022). Human embryonic development: from peri-implantation to gastrulation. *Trends Cell Biol.* *32*, 18–29. <https://doi.org/10.1016/j.tcb.2021.07.008>.

Zhang, Q.-J., and Liu, Z.-P. (2015). Histone methylations in heart development, congenital and adult heart diseases. *Epigenomics* 7, 321–330. <https://doi.org/10.2217/epi.14.60>.

Zhang, H., Tian, L., Shen, M., Tu, C., Wu, H., Gu, M., Paik, D.T., and Wu, J.C. (2019). Generation of Quiescent Cardiac Fibroblasts From Human Induced Pluripotent Stem Cells for In Vitro Modeling of Cardiac Fibrosis. *Circ. Res.* 125, 552–566. <https://doi.org/10.1161/CIRCRESAHA.119.315491>.

Zhang, S., Postnikov, Y., Lobanov, A., Furusawa, T., Deng, T., and Bustin, M. (2022). H3K27ac nucleosomes facilitate HMGN localization at regulatory sites to modulate chromatin binding of transcription factors. *Commun. Biol.* 5, 1–14. <https://doi.org/10.1038/s42003-022-03099-0>.

Zhang, S.-S., Kim, K.-H., Rosen, A., Smyth, J.W., Sakuma, R., Delgado-Olguín, P., Davis, M., Chi, N.C., Puvionindran, V., Gaborit, N., et al. (2011). Iroquois homeobox gene 3 establishes fast conduction in the cardiac His-Purkinje network. *Proc. Natl. Acad. Sci. U. S. A.* 108, 13576–13581. <https://doi.org/10.1073/pnas.1106911108>.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 9, R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.

Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., Tang, Y., Fang, Y., and Fang, D. (2021). Overview of Histone Modification. *Adv. Exp. Med. Biol.* 1283, 1–16. https://doi.org/10.1007/978-981-15-8104-5_1.

Zheng, X., Hu, J., Yue, S., Kristiani, L., Kim, M., Sauria, M., Taylor, J., Kim, Y., and Zheng, Y. (2018). Lamins Organize the Global Three-Dimensional Genome from the Nuclear Periphery. *Mol. Cell* 71, 802-815.e7. <https://doi.org/10.1016/j.molcel.2018.05.017>.

Zhu, P., and Li, G. (2016). Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr. Opin. Struct. Biol.* 36, 106–115. <https://doi.org/10.1016/j.sbi.2016.01.013>.

Zhu, Y., Sun, L., Chen, Z., Whitaker, J.W., Wang, T., and Wang, W. (2013). Predicting enhancer transcription and activity from chromatin modifications. *Nucleic Acids Res.* 41, 10032–10043. <https://doi.org/10.1093/nar/gkt826>.

Titre : Etude du développement cardiaque et ses facteurs de transcription par l'utilisation des cellules souches pluripotentes induites

Mots clés : Cellules souches pluripotentes induites humaines (iPS), Développement cardiaque, Facteurs de transcription Iroquois (IRX), Réseaux de régulation des gènes, Complexes de facteurs de transcription

Les facteurs de transcription sont des éléments clés du développement cardiaque, gouvernant la mise en place d'un programme d'expression de gènes spécifiques. L'enchaînement chronologique des étapes d'établissement de ce programme est contrôlé par un réseau de régulations dynamiques et temporelles de facteurs de transcription cardiaques. Cependant, la vision globale de ce réseau n'a pour l'instant pas été obtenue. De plus, certains facteurs de transcription, comme ceux de la famille Iroquois, qui ont pourtant été montrés comme importants dans le cœur, n'ont pas ou peu été analysés dans le contexte de réseaux avec d'autres facteurs de transcription cardiaques. Par une analyse transcriptomique de la différenciation cardiaque de cellules souches pluripotentes induites humaines, utilisée comme modèle de développement cardiaque

humain, nous avons identifié d'une part, un réseau global de régulation entre les facteurs de transcription impliqués dans le processus. D'autre part, nous avons identifié un rôle jusqu'à présent inconnu des facteurs de transcription Iroquois, IRX3 et IRX5, dans la régulation de l'expression et de l'activité de facteurs de transcription cardiaques majeurs, GATA4, TBX5 et NKX2-5. Une étude focalisée sur IRX5 a ensuite permis de mieux comprendre le rôle de ses domaines protéiques, ses mécanismes d'interaction et de régulation de l'ADN, et les fonctions biologiques qu'il régule dans la différenciation cardiomyocytaire. L'ensemble de ce travail offre ainsi une meilleure compréhension des mécanismes moléculaires du développement cardiaque et peut également servir de base pour l'étude de pathologies ou pour le développement de nouvelles thérapies.

Title: Study of cardiac development and its transcription factors using induced pluripotent stem cells

Keywords: Human induced pluripotent stem cells (hiPSCs), Cardiac development, Iroquois transcription factors (IRX), Gene regulatory networks, Transcription factor complexes

Transcription factors are key elements in cardiac development, governing the establishment of a specific gene expression program. The chronological sequence of steps in the establishment of this program is controlled by a network of dynamic and temporal regulations of cardiac transcription factors. However, a global view of this network has not yet been obtained. Moreover, some transcription factors, such as those of the Iroquois family, which have been shown to be important in the heart, have not been analyzed or only to a limited extent in the context of networks with other cardiac transcription factors. By a transcriptomic analysis of the cardiac differentiation of human induced pluripotent stem cells, used as a model of human cardiac development, we identified

on one hand, a global regulatory network between the transcription factors involved in the process. On the other hand, we identified a previously unknown role of the Iroquois transcription factors, IRX3 and IRX5, in the regulation of the expression and activity of major cardiac transcription factors, GATA4, TBX5 and NKX2-5. Then, a focused study on IRX5 allowed a better understanding of the role of its protein domains, its mechanisms of interaction and regulation of DNA, and the biological functions it regulates in cardiomyocyte differentiation. This study offers a better understanding of the molecular mechanisms of cardiac development and can also serve as a basis for the study of pathologies or for the development of new therapies.