HAL
open science

Deciphering the roles of Klf2a, Klf2b and Egr1 transcription factors in heart valve development using zebrafish as model organism

Nathalie Faggianelli-Conrozier

To cite this version:

Nathalie Faggianelli-Conrozier. Deciphering the roles of Klf2a, Klf2b and Egr1 transcription factors in heart valve development using zebrafish as model organism. Cardiology and cardiovascular system. Université de Strasbourg, 2018. English. NNT : 2018STRAJ128 . tel-02918076

HAL Id: tel-02918076

https://theses.hal.science/tel-02918076v1

Submitted on 20 Aug 2020

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.

Université de Strasbourg
Ecole doctorale ED 414
CNRS UMR 7104 – Inserm U 964

Thèse présentée par :

Nathalie FAGGIANELLI-CONROZIER

Soutenue le : 14 décembre 2018

Pour obtenir le grade de : **Docteur de l'Université de Strasbourg**
Discipline / Spécialité : Biologie du Développement

Deciphering the roles of *Klf2a*, *Klf2b* and *Egr1* transcription factors in heart valve development using zebrafish as model organism

Thèse dirigée par :

Dr Julien Vermot	DR, IGBMC, Strasbourg

Rapporteurs :

Prof. Christian Mosimann	Assistant professor, Institute of Molecular Life Sciences, Zurich
Dr Patrick Blader	DR, Centre de Biologie du Développement, Toulouse

Autres membres du jury :

Dr Gérard Gradwohl	DR, IGBMC, Strasbourg
Dr Sigolène Meilhac	CR, Institut Pasteur, Paris
Dr Eirini Trompouki	Assistant professor, Max Planck Institute, Freiburg

Table of contents

Remerciements 6

Avant-propos 12

Table of abbreviations 13

general vocabulary

Proteins, genes and molecular pathways

Table of figures 15

Introduction

I. About the importance of studying cardiovascular development 17

I.1 congenital heart diseases and valve defects 18

I.2 origin of the defects : genetic and epigenetic factors

II. Zebrafish, a powerful animal model for cardiac development and for human cardiovascular diseases modelling 20

II.1 Presentation of the zebrafish

II.2 Genetic considerations about zebrafish 22

a) Some considerations about zebrafish mutant models

b) Zebrafish transgenic lines, tol2 system and Gal4 gene trap 24

II.3 Advantages of using the zebrafish model to study heart development 28

III. Cardiovascular system organization and formation 29

III.1 General consideration about the vertebrate cardiovascular system

III.2 Presentation of the zebrafish cardiovascular system 30

a) Anatomy of the zebrafish cardiovascular system

b) General structural key steps in the zebrafish cardiovascular development 34

c) Key structural and molecular steps in zebrafish heart development 35

d) Further information about genetic regulation of heart formation 36

IV. Cardiac Valves 37

IV.1 State-of-art on the cellular processes involved in ZF valve formation 38

IV.2 State-of-art on the genetic control of ZF valve formation 41

a) Notch and Bmp signalings: bidirectional signaling in zebrafish AVC

b) VEGF signaling: regulation of endothelial division and ECM biosynthesis 44

c) Signaling pathways in cardiac jelly, role of the ECM 45

d) Wnt/beta-catenin pathway 47

IV.3 The role of biomechanical forces in cardiac valve development 48

V. Klf2a, an important gene for valvulogenesis 50

V.1 Klf2a, a member of Krüppel-like factor family of transcription factors

V.2 Functions of *Klf2/klf2a* 51

V.3 *Klf2/klf2a* is a flow-responsive gene 52

V.4 Genetic pathway regulated by *Klf2/klf2a* 53

V.5 Role of *klf2b*, paralog of *klf2a* 55

VI. Egr1, another important gene for valvulogenesis

VI.1 Egr1, a member of the 'Early Growth Factor' Family

VI.2 *Egr1* expression and functions 57

Objectives of this thesis 59

Results 60

Chapter 1: Generation of zebrafish polyclonal and monoclonal antibodies to validate Knock-Out mutant lines

I. Preamble about the optimization of zebrafish protein lysis and W.Blot assay 63

I.1 Experimental procedure to get zebrafish protein lysates

a) Choice of the extraction lysis buffer

b) Sample preparation: embryos dechorionating and deyolking steps 65

I.2 Western Blot optimization

a) The quantity of proteins loaded on gels

b) The blocking solution 66

c) The concentration of primary antibodies 67

d) The time of exposure 69

I.3. Results and discussion

a) The choice of the lysis buffer

b) The quantity of loading material 70

c) Finding a good control protein 71

d) The time of exposure

II. The need of home-made anti-Klf2a/Klf2b antibodies 72

II.1 Generation of polyclonal antibodies against Klf2a/Klf2b 73

a) Presentation of the chosen peptides

b) Purification and test of the polyclonal antibodies

II.2 Monoclonal antibodies against Klf2a, Klf2b, Egr1, Klf4 and Klf17 74

a) Monoclonal antibodies production

b) Results and discussion

Chapter 2: Use of the home-made monoclonal antibodies to study Klf2b and Egr1 protein localization by Immunofluorescence 82

I. Problematic and optimization steps 83

II. Preliminary results

III. Perspectives 85

Chapter 3: Elucidating the roles of *egr1*, *klf2a*/*klf2b* and *flt1* in valve formation 86

Manuscript of the paper which will be submitted 87

I. Cardiac expression patterns of *egr1* in zebrafish 120

I.1 *Egr1* cardiac expression 121

I.2 Additional data about *egr2* gene 122

II. Some klf members cardiac expression`

II.1 *klf2b* AVC-expression would be *klf2a*-dependent

II.2 *Klf4* and *klf17* are not expressed in AVC/OFT between 30 and 72hpf

III. Additional information about *flt1* and *flt4* gene expression 127

III.1 *flt1* cardiac expression is *klf2a*- and *egr1*-dependent

III.2 No *flt4* cardiac expression at 48hpf

IV. Additional information about *wnt9b* spati-temporal expression in ZF heart 129

Chapter 4: deciphering the cardiac downstream transcriptional network of klf2a and egr1 at 48hpf 130
I. Preamble: introduction to integrative genomic studies used to understand regulatory networks in development 131
I.1 chromatin and gene expression
I.2 genome-wide analysis techniques 133
I.3 ChIP followed by high-throughput sequencing 135
I.4 ATAC-seq 136
I.5 RNA-seq 138
II. Optimization of a ChIP method 139
II.1 Optimization steps
a) quantity of material
b) sonication step
c) antibody efficiency 140
d) choice of controls
II.2 results of the optimized steps 141
a) sonication time
b) antibodies test
II.3 Preliminary results 144
a) ChIP-qPCR
b) Results of the ChIP-seq assays 149
II.4 Discussion and perspectives 151
a) the quality of the starting material
b) the cross-linking step
c) antibody specificity and efficacy
III. Improvement: establishment of in vivo biotinylation tagging of Egr1 and Klf2a in zebrafish endocardial cells 152
III.1 Presentation of the designed constructs
III.2 Construction protocol 154
III.3 Fish mutant generation 155
III.4 Preliminary results and perspectives
IV. Additional information about the ATAC-seq/mRNA seq combined approach 159
IV.1 Discussion about the quantity of material required, choice of the kit and protocols used for ATAC-seq and mRNAseq
a) Optimisation of a FACS sorting protocol
b) ATAC-seq protocol
c) Collecting RNA from 48hpf-sorted cells
IV.2 Complementary information not included in the paper about the mRNA and ATACseq data sets 162
a) Data on the double mutant *klf2a;klf2b*
b) Some disparities with ISH data 164
Chapter 5: Design of CRISPR/cas9 system for generating KO-fn1b zebrafish mutant 166
V.1. Presentation of CRISPR-Cas9 system

V.2 Design of a CRISPR/cas9 approach to target *fn1b* zebrafish gene 170
a) Identification and customization of gRNA
b) gRNA production
V.3 Discussion and perspectives 173
General conclusion and perspectives 176
Materials and Methods 178
References 191
Résumé détaillé de these en français 208
Annex 1 216
paper: **Steed, E., N. Faggianelli, S. Roth, C. Ramspacher, J. P. Concordet and J. Vermot.** "klf2a couples mechanotransduction and zebrafish valve morphogenesis through fibronectin synthesis." Nat Commun 7: 11646.

Annex 2
manuscript: **Duchemin, A.-L., H. Vignes, N. Faggianelli and J. Vermot.** "Piezo channels control mechanosensitive outflow tract valve development through the Hippo pathway effector Yap1 and the Klf2-Notch signaling axis".

Abstract/résumé

Remerciements

Me voilà enfin arrivée au bout de ces trois (longues) années de thèse avec l'achèvement de ce manuscrit récapitulant le travail accompli. Un doctorat, comme chacun le sait et je le confirme, est un travail de longue haleine ! Mais c'est aussi, je crois, un défi que l'on se donne à soi-même. C'est aujourd'hui un chapitre de ma vie qui se termine, pendant lequel j'ai dû faire face à de gros questionnements intellectuels et à une remise en question personnelle. Je suis néanmoins heureuse de ce que cette thèse m'a apportée !

C'est trois années n'ont pas constitué un travail purement solitaire, bien au contraire. Et de nombreuses personnes se sont retrouvées impliquées, qui m'ont aidé à obtenir les résultats présentés dans ce mémoire. Ce sont ces personnes que j'aimerais mettre en avant dans ces remerciements, et souligner que ce chapitre de ma vie a aussi été l'occasion de belles rencontres et amitiés au-delà de l'aspect purement professionnel et scientifique.

Aux membres du jury

Je tiens tout d'abord à remercier Monsieur **Patrick Blader**, Monsieur **Christian Mosimann**, et Monsieur **Gérard Gradwohl** d'avoir accepté de lire cette thèse, et de juger mon travail en tant que rapporteurs. Je remercie tous les membres du jury, les rapporteurs précédemment nommés, ainsi que Madame **Sigolène Meilhac** et Madame **Eirini Trompouki**, d'avoir accepté d'assister à la présentation de ce travail, et de s'être déplacés sur Strasbourg.

A mon directeur de thèse

J'aimerais ensuite remercier mon directeur de thèse, **Julien Vermot**, pour m'avoir donné la possibilité d'effectuer ma thèse dans ton laboratoire et pour ta confiance accordée tout au long de ce travail de recherche. Merci de m'avoir laissé des libertés d'initiative tout au long de ces trois années de travail en commun et le temps que tu m'as consacré.

A nos collaborateurs / to our collaborators

Then, I would like to thank you in particular, Dr **Eirini Trompouki**. Thank you so much for your support, your great and valuable help during these three years! This work wouldn't have been possible without you. I have learnt a lot about chromatin assays and treatment of -omics data. I would like to thank you also for your kindness, your constant support and optimism with my results! It was a great pleasure to come and work from time to time in your lab in Freiburg with you and your team! Thank you for your help and corrections for some paragraphs of this thesis manuscript, for the scientific discussions we have had! A big thank also to Mrs **Katerina Polizou** for the great job you did for me to analyze the data and the huge time you have consumed to work on them nights and days!

Aux membres de l'IGBMC

Je voudrais également remercier les différentes plateformes de l'IGBMC avec qui j'ai eu l'opportunité de travailler. Merci à :

M. **Bernard Jost**, Mme **Christelle Thibault-Carpentier**, Mme **Cathy Herouard** de la plateforme de séquençage.

Mr **Bernardo Da Reina** et Mr **Vincent Heyer** pour votre aide dans le design et le clonage de mes constructions - qui nous ont fait nous arracher les cheveux pendant quelques semaines !

Mmes **Muriel Philipps** et **Claudine Ebel** de la plateforme de FACS pour les nombreuses heures passées à trier mes cellules et votre bonne humeur !

Mr **Mustapha Oulad-Adbelghani** et à la plateforme anticorps pour votre aide dans la génération des anticorps polyclonaux et monoclonaux

Mme **Elvire Guiot**, M. **Didier Hentsch** et M. **Erwan Gandgirard** (promis je pars avec ma poisse des microscopes !) de la plateforme d'imagerie.

Sandrine, Sylvie et Clarisse de l'animalerie poisson pour leur travail et leur aide pour prendre soin de nos petits protégés sans qui ces expériences ne pourraient se faire !

M. **Guy Joubert** du service info pour toujours répondre rapidement à nos demandes et résoudre efficacement les problèmes !

Mme **France Chivaille** pour ton aide dans l'organisation administrative de la fin de thèse !

A mes collègues

Un grand merci à tout le **labo Vermot** bien sûr – ceux qui sont déjà partis et ceux qui sont encore là (Marina, Pedro, Laia, Hajime, Francesco) ! Et plus particulièrement à :

- Thank you first to **Emily Steed**: you launched with me my PhD project and taught me all your techniques! thanks for your constant cheerfulness, optimism and support! Was such a pleasure to work this you!
- Merci à une autre **Emilie** ☺ ma première année de thèse en ta compagnie a été un plaisir et je suis heureuse de t'avoir rencontrée ! Promis le Canada est sur ma liste de voyages ! Et un grand merci pour avoir relu avec attention mon introduction !
- Merci à Lhéanna pour ton sourire, ta bonne humeur ! tu étais toujours là pour donner un coup de main même depuis Freiburg ;)
- Un MERCI tout particulier à **Hélène et Anne-Laure** pour tous les bons moments passés ensemble dans et hors labo. Certains paragraphes de ce mémoire ont bénéficié de votre lecture attentive et de vos remarques et je vous suis vraiment reconnaissantes de m'avoir aidé sur cette phase de rédaction ! J'aimerais vous dire ô combien j'ai apprécié votre immense soutien pendant les périodes moins drôles de cette thèse et pour m'avoir supportée ! Vous allez me manquer *« et Etienne Daho aussi ;) »*
- Merci à toi **Stéphane** aussi, mon binôme de paillasse, pour nos fou-rires, les bons moments passés, et pour ton aide dans le labo et notamment sur mes dernières qPCR ! « *Remercier Stéphane, voilà ça, c'est fait ! ;)* »
- Thanks Renee for your great help with the photoconversion experiment!
- Thank you, Rita, for your great support the nice time we had with you in the lab!
- Merci à Mélanie, notre ex-voisine de labo, pour ton soutien et les bons moments passés. « *A nos tupps presque parfaits avec Hélène et Anne-Laure* ;) »

A mes amis

Et enfin, last but not least, merci à mes amis 'hors labo'. J'aurais tellement de choses à vous dire et ce serait toujours en deçà de la vérité. Sans vous tous et votre soutien, cette thèse n'aurait peut-être pas vu le jour ! Merci d'être entrés dans ma vie et pour tout ce que vous m'apportez au quotidien !

Tout d'abord à toi ma **Clémence** ! merci pour ce que tu es et ce que tu apportes dans ma vie, merci d'avoir toujours été là et surtout au moment où j'en avais le plus besoin ! le début de cette thèse a aussi été le début d'une merveilleuse rencontre et Amitié, je pense que tu ne conçois pas à quel point tu as été précieuse pour moi ces dernières années ! Petite dédicace à **Guillaume** aussi et pleine de belles choses pour Bookinou !! Sans aucun degré d'importance, merci tout autant à toi ma **Sophie**, pour cette belle Amitié partagée avec toi depuis l'école d'ingénieur, pour ton soutien et ton positivisme constant ! Et bien sûr à Aurélie et Marie pour tous les bons moments que l'on a déjà vécu ensemble ;)

Merci à mon petit groupe d'amis de la danse et de l'escalade avec qui j'ai passé de très bons moments à la danse et en dehors !
A **Joseph, Prudence** (merci pour ton inébranlable optimisme, tous les soirées passées qui remontaient mon curseur de bonne humeur au sommet), **Seb** et **Alice** (pour ton soutien tout au long de ma thèse et cette belle amitié !), **Pauline** (pour nos atelier couture, nos sorties et ta gentillesse !), **Julien E., Julie** (une très belle rencontre et j'ai hâte de nos prochaines pauses couture !), **Rachel** (pour nos petits moments à tous les deux, les danses et notre semaine mémorable à Avoriaz il y a deux ans !), **Caro et Seb** (notamment pour les nombreux gouters et dîners chez vous avec thermomix ;))**, Célia** (en souvenir notamment de nos super macarons de nouvel an !)**, Pierre** (pour ton coaching escalade et ton écoute!)**, Marion, Claire** (pour nos trop rares moments magiques de cuisine ensemble)**, Victor** (pour toutes les aprems jeux) mais aussi à **Julien M**., **Arnaud**, **Paul, Paula**, **Luc, Noël et Kim** (pour nos ateliers bricolage, couture et cuisine et nos riches discussions !), **Célia J.** (pour nos petites pauses midi qui ont égayées notamment ma dernière année!). A **Jeanne** et **Daniel** que je suis ravie d'avoir rencontrés grâce au WCS ! A **Anne** et **Jérome T.** pour me transmettre patiemment et avec plaisir votre passion **!** Et à tous les autres membres de la petite troupe de danseurs de West Coast Swing strasbourgeois. La danse a toujours été mon échappatoire aux longues journées de thèse et aux baisses de moral et la bouffée d'air qui me permettait de prendre du recul sur ce travail !

Merci à la fine équipe « parisienne et co. » !! Que de bons moments déjà passés, merci de toujours répondre présents pour tout ! **Thibaut**, **Louise, Jérôme, Clément, Corentin, Sophie, Edouard, Nico**, et **Lucie** (pour nos deux années de coloc aux lentilles, tous les bons moments passés et nos vacances au Brésil !). Vivement l'édition 2 de Tourtour !!

Merci à **Laure** (pour ta bonne humeur et tes mains de kiné magiques) et **Bruno** (notamment pour cette belle année et demie de coloc et les kilomètres de trails partagés) ! Vous êtes merveilleux tous les deux !

A Maïwenn pour les bons moments passés ! A **Amalia**, pour ton soutien depuis Lyon. A **Isabelle Pfeil** pour ton écoute, ton constant optimisme et bonne humeur !

A Glenn, Raphaëlle, Oriane, Hélène, Mélissa, Thibaut, Armand et Alex pour les bons moments passés, que de souvenirs ;) Et en mémoire de Pierre, parti beaucoup trop tôt.

A l'équipe Openlab et au Jardin des Sciences

Merci à toute l'équipe d'**Openlab** (dont Téo, Elodie, Sarah et Madgalène). Ce fut une formidable aventure humaine, je suis très heureuse de ces journées hors labo passées sur les routes alsaciennes, lors de ma dernière année, pour faire découvrir aux lycéens notre travail et les métiers de la recherche ! Merci à **Lucile et à Amandine** du Jardin des Sciences pour votre gentillesse, votre aide et votre soutien dans ma recherche de carrière future ! Ce fut un plaisir de travailler avec vous sur différentes missions du JDS.

A ma famille

Un merci tout particulier à ma famille, à mes **parents et à toi Isa**, qui m'ont soutenue tout au long de ces années, et qui ont suivi les difficultés, les moments de doute tout au long de l'avancement chaotique de mes travaux jusqu'à aujourd'hui. Malgré mon éloignement, merci pour votre compréhension, votre patience et votre confiance dans mes choix ! Je suis fière que vous soyez présents aujourd'hui à la fin de cette thèse ! Dédicace spéciale à mes grands-parents : vous allez pouvoir découvrir enfin ce que je faisais avec ces poissons ;)

Et à toi, Téo

Mes derniers remerciements te sont adressés **Téo**, pour toi qui est entré dans ma vie pendant ma troisième année de thèse, et qui a fait entrer le soleil dans chacune de mes journées depuis lors ! merci pour ta grande aide dans la réalisation de mes figures et pour m'avoir supportée et soutenue dans ces dernières semaines de rédaction !

J'en oublie certainement encore et je m'en excuse ! Encore un grand MERCI à tous pour m'avoir aidé à arriver au bout !

Avant-Propos

My PhD work aimed at connecting different aspects of biology, trying to make links between molecular and cellular observations and genome-wide data to study the role of two transcription factors (*klf2a* and *egr1*) in a developmental context: the formation of cardiac atrio-ventricular valves.

This manuscript starts by describing, within a state-of-art introduction, the vertebrate cardiovascular system focusing on valvulogenesis and the advantages of using *Danio rerio*, better known as zebrafish, as an animal model to study it. The manuscript is then divided into five main chapters.

The first section describes the generation of monoclonal antibodies in order to validate the knock-out of *egr1, klf2a* and *klf2b* genes in the zebrafish mutant lines used in the lab. Related preliminary data of their use to study protein localization *in vivo* are presented in chapter two.

The third and fourth parts correspond to the main project of my thesis: the study of *klf2a* and *egr1* transcription factors. The characterization of their roles in valvulogenesis, their interaction and the discovery of their downstream transcriptional signalling pathway was carefully explored. Part four presents the paper which will be sent to editors gathering the results. Specific and more detailed analysis will complete it.

Finally, chapter five will present the preliminary work aiming at the generation of zebrafish mutants using the CRISPR/Cas9 system, with *fn1b* as a target gene. This gene encodes for fibronectin protein and was shown to be a Klf2a downstream target in valvulogenesis.

Table of abbreviations

General vocabulary

A: Atrium
Ab: antibody
ALPM: anterior lateral plate mesoderm
AVSD: atrioventricular and septal defects
BDM: 2,3-butanedione-2-monoxime
CCM: Cerebral Cavernous Malformation pathway
DA: dorsal aorta
DV: dorsal vein
DNA: Deoxyribonucleic Acid
ATAC-seq: Assay for Transposase Activity sequencing
AVC / AV canal: Atrio-Ventricular Canal
bp: base pair
CHD: Congenital Heart Disease
ChIP: Chromatin Immuno-Precipitation
dpf: days post fertilization
ds: double stranded
DTT: dithiothreitol
EC: endothelial cell
ECM: Extra-Cellular Matrix
EMT: Endocardial to Mesenchymal Transformation
ENU: N-ethyl-nitroso-urea
FACS: Fluorescence-Activated Cell Sorting
FHF: First Heart Field
hpf: hours post-fertilization
IFT: Inflow Tract
ISH: In Situ Hybridization
ISV: intersegmental vessels
KD: Knock-down
KI: knock-in
KO: knock-out
mRNAseq: messenger RNA sequencing
MO: morpholino
NMD: Non-sense Mediated Decay
NP-40: Tergitol-type NP-40, nonyl phenoxypolyethoxylethanol
nt: nucleotide
PCR: Polymerase Chain Reaction
PFA: Para-Formaldehyde
PTU: 1-phenyl 2-thiourea
PCV: Posterior Caudal Vein
OFT: Outflow Tract
qPCR: quantitative Polymerase Chain Reaction

RNA: Ribonucleic Acid
RT: Room Temperature
SDS: Sodium Dodecyl-Sulfate
SHF: Second Heart Field
UAS: Upstream Activating Sequence
seq: sequencing
V: Ventricle
WES: Whole Genome Sequencing
ZIRC: Zebrafish International Resource Center
ZF: zebrafish

Proteins, genes and molecular pathways

Bmp: bone morphogenic protein
Cmlc2: cardiac myosin light chain 2
CRISPR/Cas9: Clustered Regularly Interspaced Short Palindromic Repeats / Cluster Associated 9
Gata: GATA binding protein or Erythroid transcription factor
EGR: Early Growth Factor
erbB: family of receptor tyrosine kinase
Flk*:* kinase insert domain receptor like or *kdrl* promoter
Fli/fli1a*:* Friend Leukemia Integration 1a promoter
Flt (flt1): Fms-related tyrosine kinase or VEGF receptor
Has (has2): hyaluronan synthase
HDAC (HDAC5): class II histone deacetylase
Hh: Hedgehog
HSC: Hematopoietic Stem Cell
KLF: Krüppel-Like Factor
Isl (Isl2): Islet family, ISL LIM homeobox gene
Mef2cb: myocyte enhancer factor 2cb
Myl: Myosin regulatory light chain polypeptide
n-fact1: calcineurin- dependent factor 1
nkx: homeodomain transcription factors
nos: nitrogen-oxyde synthase (NO: nitrogen oxydase)
PrKD2: protein kinase D2
R: receptor
RA: retionoic acid
TALEN: Transcription activator-like effector nuclease
TBX: T-box transcription factor
TGFβ: Transforming Growth Factor Beta
tnnt2/tnnt2a : cardiac troponin T
VEGF: Vascular endothelial growth factor
WNT (Wnt9b): signaling gene family Wingless-type MMTV integration site family (factor 9b)

Table of figures

Figure 1: Zebrafish embryo at different developmental stages. 21
Figure 2: Genetic tools in zebrafish. Use of the Gal4FF/UAS system to express endogenous sequence in the tissue of interest. 27
Figure 3: comparison between the circulatory systems of fish and mammals. 31
Figure 4: Simplified diagram of the zebrafish cardiovascular system overall organization. 32
Figure 5: Comparative anatomical schemes of a zebrafish and human hearts. 33
Figure 6: Models of the early stages of AV valve development 40
Figure 7: Cardiac jelly protein versican as a downstream target of CCM pathway. 46
Figure 8: Role of biochemical forces in AV valve development through KLF-WNT signalling. 50
Figure 9: General genetic and molecular pathways leading to the formation of AVC valve leaflets. Schematic representation of the mechanisms of activation of KLF2/klf2a and of the signaling pathways activated by this transcription factor. 56
Figure 10: zebrafish Kf2a protein domains. 58
Figure 11: Phylogram of KLF zebrafish proteins. 58
Figure 12: zebrafish Egr1 protein domains. 58
Figure 13: Presentation of the zebrafish KO-mutant lines for *egr1, klf2a* and *klf2b* genes used in this thesis. 64
Figure 14: Optimization of a Western Blot protocol for protein lysates of 48hpf-embryos. 70
Figure 15: *egr1* WISH on the transgenic lines over-expressing *klf2a* in endocardial cells. 79
Figure 16: test of zebrafish monoclonal anti-Klf2a, anti-Klf2b, anti-Klf17 and anti-Egr1 antibodies, generated and purified at IGBMC. 80
Figure 17: CLUSTAL O(1.2.4) multiple sequence alignment of Klf2a, Klf2b, Klf4 and Klf17 zebrafish proteins. 82
Figure 18: Immuno-Fluorescence assay with anti-Klf2b and anti-Egr1 antibodies 86
Figure 19: *egr2b/krox20* expression pattern between 30 and 72hpf in WT zebrafish embryo. 121
Figure 20: *klf2b* spatio-temporal expression (WISH assay) on 30-48 and 72hpf zebrafish embryos ..124
Figure 21: *klf2b* WISH on blood-flow, *klf* and *egr1*^{sa64} mutants at 48hpf 125
Figure 22: *klf4* and *klf17* WISH expression in 30-48 and 72hpf WT zebrafish embryos. 126
Figure 23: *flt4* WISH on 48hpf-zebrafish embryos 128
Figure 24: *wnt9b* WISH in 30 to 72hpf zebrafish embryos 128
Figure 25: Hierarchical levels of DNA organization in the eukaryotic nucleus. 132
Figure 26: Integration of transcriptomic and epigenetic studies. 134
Figure 27: Comparative scheme presenting ATAC-seq and ChIP-seq principles and the method used in this thesis on 48hpf-zebrafish endocardial cells. 137
Figure 28: Optimization of the sonication step of the ChIP protocol. 142
Figure 29: Test of the antibody efficiency to immune-precipitate Klf2a and H3K4 proteins in ChIP assay. 143
Figure 30: ChIP-qPCR 146
Figure 31: Positions of negative/positive couples of primers on *nos1* and *fn1b* zebrafish genes used in qPCR. 147
Figure 32: comparison of ChIP-seq data on *erg1* promoter 150
Figure 33: Immunoprecipitation of Klf2a protein 150
Figure 34: schematic representation of the transgenic constructions designed to over-express a biotin-tag version of Klf2a and Egr1 in endocardial cells of living fish. 153
Figure 35: Presentation of birA biotinylation system used to create tagged proteins 153
Figure 36: Megawhoop Golden Gate cloning 156

Figure 37: cDNA quality profiles. Analysis performed with Bioanalyser. 161
Figure 38: Genome-wide data obtained for the double mutant *klf2a-klf2b*.....163
Figure 39: *has2* WISH on 48hpf-zebrafish embryos 165

Figure 40: Overview of the CRISPR/Cas system…….170
Figure 41: T7 endonuclease I assay to detect Cas9 induced mutations……178

Figures of the scientific paper

Manuscript, figure 1: *egr1* WISH in 48hpf-ZF embryos.
Manuscript, figure 2: Description of the valve defects observed in *egr1*^{sa64}; *klf2a;klf2b and klf2b* mutants
Manuscript, figure 3: mRNA sequencing results.
Manuscript, figure 4: ATAC-seq results and over-lapping between ATAC and mRNAseq data
Manuscript, figure 5: footprinting analysis
Manuscript, figure 6: Description of *flt1*^{sa1504} mutant and *flt1* WISH in 48hpf-ZF embryos.
Manuscript, figure 7: *wnt9b* WISH in 48hpf-ZF embryos.
suppl. figure 1: *egr1* WISH in 30-48 and 72hpf-ZF embryos
suppl. figure 2: *klf2a* WISH in 48hpf-ZF embryos

Introduction

The development of a well-organized and functional organ is a multistep process referred as morphogenesis. The classical view of organogenesis suggest that organs tri dimensional shape is the result of genetically programmed events. Tissues are patterned into distinct area by gene expression domain. At the cellular scale, tissue shape results from cell remodelling, migration and differentiation. Most of the cellular behaviours are governed by spatiotemporal changes in gene expression. Morphogenesis also relies on the mechanical forces defined by the physical properties of the system. In developing systems, mechanical forces can be active because cells can sense and react to mechanical stimuli through the process of mechanotransduction. To date, it is still unclear how genetic programs and mechanical forces interact to control morphogenetic programs.

During my phD thesis, I studied the roles of three mechanosensitive transcription factors: Klf2a, Klf2b *and* Egr1 *during the process of cardiac valve morphogenesis.*

I. About the importance of studying cardiovascular development

Organogenesis is a well-organized and controlled process. Nevertheless, it happens one of the steps may be impaired and therefore impacts the correct formation of organs, e.g. the heart and/or the valves. It is obvious that valve function has to be mechanically flawless, as the heart beats continuously during lifespan. Either the obstruction of forward flow or a backwards regurgitation of blood due to defective valves could cause heart failure or could lead to congenital heart diseases.

I.1 Congenital heart diseases and valve defects

Congenital Heart Defects (CHDs) represent the most common form of birth defects occurring in 1% of live birth, with similar prevalence throughout the world (CDC statistics). The defects are diverse and too severe alteration can even result in death at birth. Therefore, CHDs constitute a leading cause of infant mortality and morbidity worldwide (Kuo et al., 1997)

Among the possible defects, atrioventricular and septal defects (AVSDs) account for the majority of CHDs (Hoffmann and Kaplan 2002) and, to date, are still a challenging medical issue. In humans, cardiac valve defects account for 20-30% of all congenital cardiovascular malformations, with an incidence of ≈2% in all live births (Hoffman and Kaplan, 2002). They can affect the mature heart structure and its beating function. The most common are mitral insufficiency (i.e., regurgitation), mitral stenosis, myxomatous degeneration (formation of soft tumor) as well as mitral valve prolapse and can eventually lead to early death (Pagnozzi and Butcher 2017)

I.2 Origin of the defects: genetic and epigenetic factors

Majority of the CH defects comes from abnormal embryogenic atrio-ventricular (AV) valve development: for example, delays or alterations in the signalling pathway regulating heart septation (Lin et al., 2012). Although the precise origin of each defect and related-disease is elusive, multifactorial causes like genetic, environmental or a combination of both seems to be link. In main cases, whitout clear origini, epigenetic factors are thought to play an important role in the pathogenesis of congenital valve defects (Goddard et al., 2017). Among the environmental causes, alcohol or tobacco, the use of certain drugs and medications (barbituric, anti-inflammatory drugs) (van Gelder et al., 2011), poor nutritional status (Hassan et al., 2015) , obesity of the mother (Brite et al., 2014) or viral infections (like rubella, (Stuckley 1956) contracted during pregnancy can lead to heart malformations.

Most of the diseases are usually linked to specific genetic mutations in genes related to heart valve development, such as the signalling factors: Notch1 and Transforming Growth Factor Beta (TGFβ) for the aortic valves (Kerstjens-Frederikse et al., 2016) (Arthur and Bamforth 2011) (Theodoris et al., 2015), T-box transcription factor 5 & 20 (TBX5,TBX20) (Chiplunkar et al., 2013), GATA4 (Guo et al., 2017), the transcription coactivator LIM and cysteine-rich domains protein 1 (LMCD1), (Dina et al., 2015), and also proteins of the

extracellular matrix : Tensin-1 (TNS1), the member of cadherin and cell adhesion proteins protein dachsous homolog 1 (DCHS1) and Filamin A (FLNA), an actin-interacting protein (Sauls et al., 2012) involved in mitral valvulopathies. Affected valves in patients carrying FLNA mutations often display defects in extracellular matrix (ECM) deposition. thus leading to affected mitral valve architecture and degeneration. The gold standard treatment for advanced heart valve disease is surgical replacement. But none of the currently available mechanical and biological heart valve substitutes are ideal solutions, though living autografts (from the same body) are clearly have a better performance than non-living homografts (mechanical valve) (Walter et al., 2012). Understanding how mechanical forces can be used to grow better living valves is an ongoing challenge for biomedical engineers of valve defects (MacGrogan et al., 2014). Another important research field investigates how these leaflets develop during embryogenesis, in particular regarding the spatio-temporal signalling regulations and the biochemical signals which are involved. In addition, it may help to create engineered viable tissue for heart valves and developing novel treatment strategies for valve diseases.

In order to study cardiac development and some diseases linked to it, a powerful animal model has been used since 1990s: zebrafish.

II. Zebrafish, a powerful animal model for cardiac development and for human cardiovascular diseases modelling

Among in vivo model systems, unique opportunities for understanding morphogenesis, and in particular for studying heart development and valvulogenesis, are provided by zebrafish embryos. Zebrafish has become a popular animal model for genetics and developmental biology since the 1990s.

II.1 Presentation of the zebrafish

Danio Rerio or *Brachydanio Rerio* is a vertebrate from the Telosteis intraclass of fish. It is better known as the zebrafish, appropriately named due to its black-and-white stripes (Figure 1). It is a small (three centimetres in length for an adult) and tropical freshwater fish, originally found in slow streams, rice paddies and in the Ganges River in East India and Burma. It is a common and robust aquarium species, where it breeds all year round.

George Streisinger was the first scientist to use zebrafish as an animal model for embryogenesis and developmental studies because of its "desirable attributes". Amongst these, include its small size, easy maintenance in an aquarium, the external fertilization and extrauterine development allow access to the embryos easily from one-cell stage. 100 to 300 embryos can be obtained per couple per week, considerably enhancing the experimental throughput in comparison to other vertebrate models such as chicks or mice. Generation time is short (for a vertebrate), typically 3 to 4 months, making it suitable for selection experiments. To make a comparison with human development, the first two days of zebrafish development correspond to the 35 first days of human embryo and 12 first days of mouse embryos (reviewed by (Stainier 2002). Organogenesis is rapid, with precursors of all major organs developing within 36 hours, larvae displaying food-seeking and active avoidance behaviours within five days after fertilization.

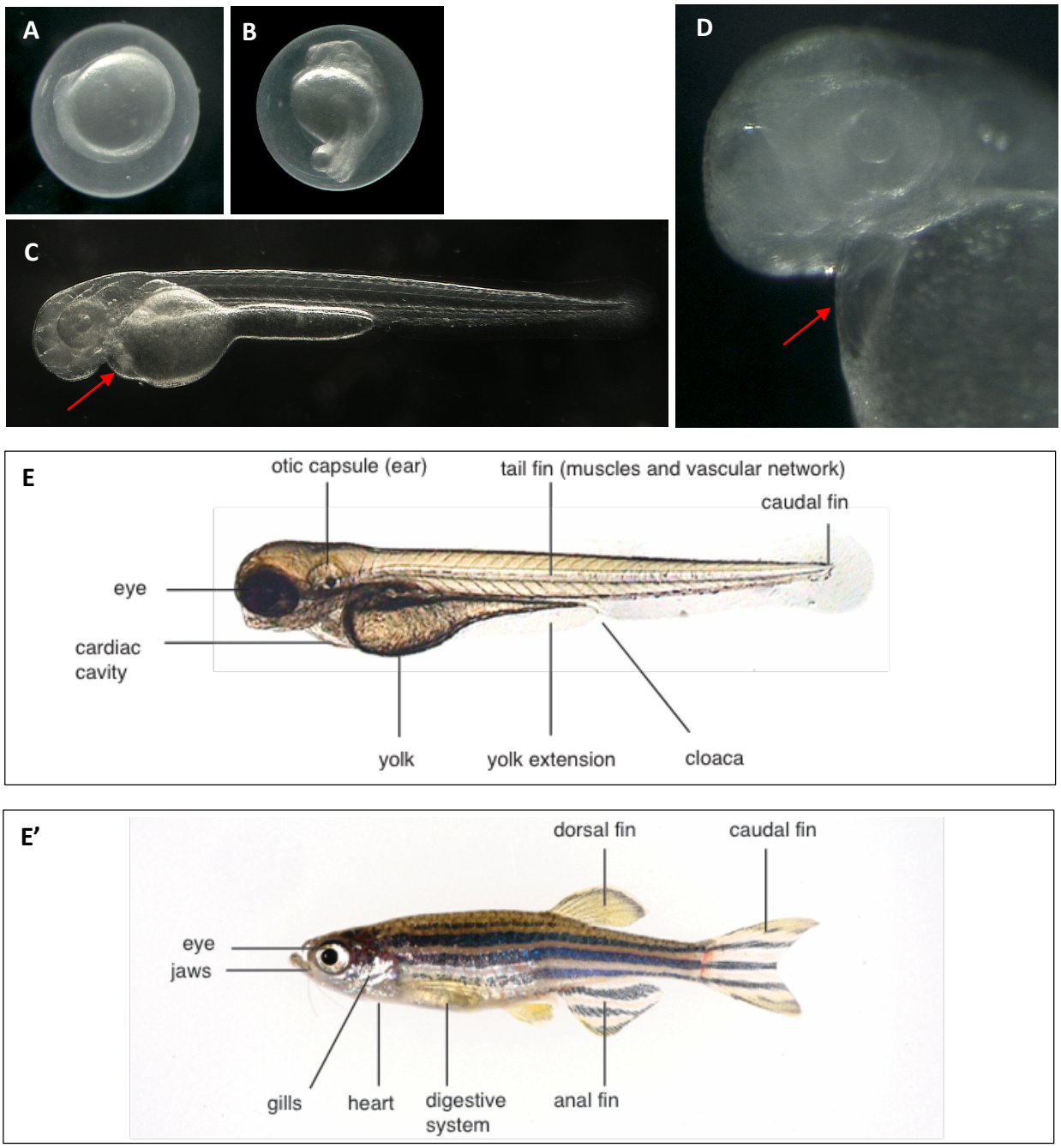


Figure 1: Zebrafish embryo at different developmental stages.

A: 12hpf, B: 24hfp, C and D': 48hpf, E: 96hpf, E': adult

The principal anatomical structures in a 48hpf-embryo and in an adult fish are presented in E and E'. Due to the optical transparency, enhanced by the addition of PTU in the medium which prevents the formation of pigments, cardiac vessels and organs can be observed easily under visible light on a microscope (A, B, C, E x50 ; D: x250, E': 1x). E: visualization of the heart cavity.

II.2 Genetic considerations about zebrafish

Another important characteristic that makes the zebrafish an appealing model for research is its use as a genetic tool.

Its diploid genome has been sequenced and is completely annotated. It contains 25 chromosomes (without the heteromorphic sexual ones). It has 1.4 billion bp size, corresponding to a little less than half of the human genome. Interestingly, 69% of the 26.000 protein-coding genes have a human ortholog, making zebrafish to be a pertinentn animal model to study human diseases (Howe et al., 2013). A whole genome duplication event has been postulated to explain the high redundancy of zebrafish genes and the fact that approximately 15% of the human genes have more than one ortholog in *Danio rerio (Postlethwait et al., 2000)*. Paralogs can have either redundant functionalities or retain one of the functions of their human ortholog.

II.2.a) Some considerations about zebrafish mutant models

The ease of depleting a candidate gene in zebrafish was extensively for studying loss-of-function effects of mutations and for identifying new genes implicated in development and pathologies.

Rapid screening for candidate gene validation, genome annotation and disease modelling was first realized using morpholinos (MO)-based antisense Knock-Down (KD) (Summerton 1999). MOs are 25 morpholine (organic chemical compound) base-oligomers, that are complementary in sequence to their target region. Their backbone confers to them a high affinity with RNA. MOs have three ways of action: first they can modify the splicing of the pre-mNA causing a specific exon to be spliced out. Second, they can inhibit the maturation and activity of miRNA. Finally, they can block the initiation of translation due to steric hindrance on the start site, leading in most cases to the production of non-functional protein. Despite special care brought to ensure specificity of the targeting and to avoid off-target effects and toxicity has been developed (Bill et al., 2009), some conflicting observations between morphants (MO-injected embryos) and mutants raise serious concern about the specificity of the morpholino anti-sense technology. (Kok et al., 2015) conclude that "*after looking at more than 80 genes, that approximately 80% of morphant phenotypes were not observed in the corresponding mutants*" (Stainier et al., 2015). Because of these drawbacks, this method tends

to be neglected compared to Knock-Out (KO) models; or used as a complementary study.

Forward genetic screens has generated thousands of mutants by introducing random point mutations in the genome using either gamma rays (Chakrabarti et al., 1983) or the more commonly used point mutagen N-ethyl-nitroso-urea (ENU) (Driever et al., 1996) ; (Haffter et al., 1996) The mutated gene was then identified using gene mapping techniques and mutants showing recessive inheritance of the phenotype were selected. The largest random mutagenesis project is carried out by the *Welcome Trust Sanger institute* (Kettleborough et al., 2013). These screens have produced mutants with a variety of cardiac phenotypes (Stainier et al., 1996), many of which resemble human cardiac malformations.

The repertoire of zebrafish mutagenesis tools, breeding strategies and mutant selection approaches has no match in any other vertebrate. Nowadays, new techniques have emerged to generate directed-mutants using nuclease-based techniques including Transcription Activator-Like Effector Nucleases or TALENs (Bedell et al., 2012) and Clustered Regularly Interspaced Short Palindromic Repeats / Cluster Associated 9 or CRISPR-cas9 (Hwang et al., 2013). These techniques are based on the insertion of a double-stranded cut on genomic DNA by a nuclease domain which can specially recognized at a locus of interest a DNA binding motif. This motif is peptidic in the case of TALEN and RNA-based in the case of CRISPR/Cas9. the induced cut will be repaired by the cell error-prone repair mechanism, which likely lead to insertion and/or deletion. This can lead in modifications of the reading frame and therefore to a shortened and un-functional protein.

The *Zebrafish Mutation Project* (ZMP) now aims to create a knockout allele in every protein-coding gene in the zebrafish genome, using a combination of whole exome enrichment, nuclease-based techniques and Illumina next generation sequencing.

II.2.b) Zebrafish transgenic lines, tol2 system and Gal4 gene trap

In addition, a number of transgenic strains have been generated, expressing fluorescent proteins in tissue specific cell types (Long et al., 1997), (Huang et al., 2003) and (Motoike et al., 2000) making zebrafish a powerful model for imaging as well. Aside from its high genetic homology with humans (Howe et al., 2013), a wide range of transgenic reporter lines with tissue-specific expression of fluorescent proteins are available for the scientific community. The Zebrafish International Resource Centre (ZIRC) gathers mutagenized and wild-type zebrafish strains library and for materials and information about zebrafish research, public available on the online platform Zfin.

Tol2 transposon system

Introducing transgenes into the zebrafish germline to generate stable transgenic lines is facilitated by the use of Tol2 transposable element.

Tol2 system consists of a transposon-donor plasmid carrying a Tol2 construct with gene of interest and the transposase activity supplied in the form of mRNA synthetized in vitro. The tol2 construct contains DNA sequences recognized by the transposase. These minimal cis-sequences essential and sufficient for transposition are 200-bp from the left end (L200) and 150-bp from the right end (R150) of the Tol2 element ((Balciunas et al., 2006); (Urasaki et al., 2006). Any DNA fragment can be cloned between these cis-sequences. Tol2 system is active in all vertebrate cells tested so far (Kawakami 2007).

In zebrafish, plasmid carrying *Tol2* elements can be micro-injected into fertilized eggs with the Tol2 transposase mRNA. Tol2 - excised from the donor plasmid - is integrated into the genome of the germ cell lineage during embryonic development, and the transposon insertions are transmitted to the next generation through germ cells (Kawakami and Shima 1999) ; (Kawakami et al., 2004). Because of the high transposition efficiency in the germ line and the capacity to carry a large DNA fragment, Tol2-mediated transgenesis has become a popular method to create zebrafish transenic lines. Moreover, it allows the integration of transgenes as single copies, thereby eliminating the problems associated with insertions containing complex concatemeric arrays (Kawakami et al., 2000).

Gal4/UAS regulatory system

Tol 2 system was successfully applied to gene and enhancer trap methods[1]. It was in particular associated with the Gal4 gene trap as well as enhancer trap methods (Scott et al., 2007) ; (Asakawa and Kawakami 2008).

The yeast transcriptional activator Gal4[2] was used to develop a powerful system to activate the transcription of gene of interest. Gal4 and its variant Gal4FF (Asakawa and Kawakami 2008) binds to a specific recognition sequence called UAS (for Upstream Activating Sequence) and activates transcription of the target gene placed downstream the UAS. Gal4/Gal4FF can be expressed in particular tissues and stimulates expression of a gene linked to UAS in a tissue-specific manner. The Gal4/UAS system was adapted to zebrafish by Scheer and Campos-Ortega (Sheer and Campos-Ortega 1999) who assayed reporter expression under the control of 5 UAS copies (5X UAS). It is widely employed to analyse zebrafish gene functions *in vivo* by crossing specific Gal4FF-expressing lines with UAS-reporter / UAS-effector lines. An example is presented on Figure 2, with the over-expression of Klf2a in endocardial cells.

[1] Gene trapping method is used to disrupt genes by inserting a trapping cassette consisting of a promoter-less reporter gene and selectable genetic marker. When inserted into an intron of an expressed gene, the trap cassette is transcribed from the endogenous promoter of that gene in the form of a fusion transcript in which the exon(s) upstream of the insertion site is spliced in frame to the reporter/selectable marker gene. Since transcription is terminated prematurely at the inserted polyadenylation site, the processed fusion transcript encodes a truncated and nonfunctional version of the cellular protein and the reporter/selectable marker. Thus, gene traps simultaneously inactivate and report the expression of the trapped gene at the insertion site. The same method can be applied to enhancer trapping to investigate how and when enhancer DNA sequences can affect gene regulation, and also aid in the determination of their possible location

[2] In yeast, the Gal4 transcription factor binds to upstream activating sequences (UAS) to direct transcription of genes necessary for metabolism of galactose. UAS is found between the GAL1 and GAL10 genes from the yeast galactose metabolic system. **Giniger, E., S. Varnum and M. Ptashne,** (1895). "Specific DNA Binding of GAL4, A Positive Regulatory Protein of Yeast." Cell 40: 767-774.

Notes about UAS sequences

UAS are repetitive sequences which exhibited high affinity for the Gal4 DNA binding domain in vitro (Kang et al., 1993). Synthetic promoters consisting of basal transcription initiating elements combined with tandem repeats of the UAS sequence were designed. Each UAS is 17 base pairs long, roughly palindromic, and in the form of CGG-N11-CCG. The CpG dinuleotides are essential for Gal4 binding (Marmorstein et al., 1992) and serve as a target for methylation (Goll et al., 2009). UAS are assembled into multiple tandem copies and separated by 10bp spacer sequence. Four or five distinct upstream activator copies (4x – 5X UAS) were shown to be far less susceptible to methylation than insertions containing fourteen repetitions used in the past (Akitake et al., 2011). It was difficult to obtain high levels of expression from these constructs, most likely because they were integrated as large concatemers of multiple transgenes, which made them susceptible to silencing. The second and fourth repetitions are placed in reverse orientation to minimize further the repetitive nature of the multicopy UAS.

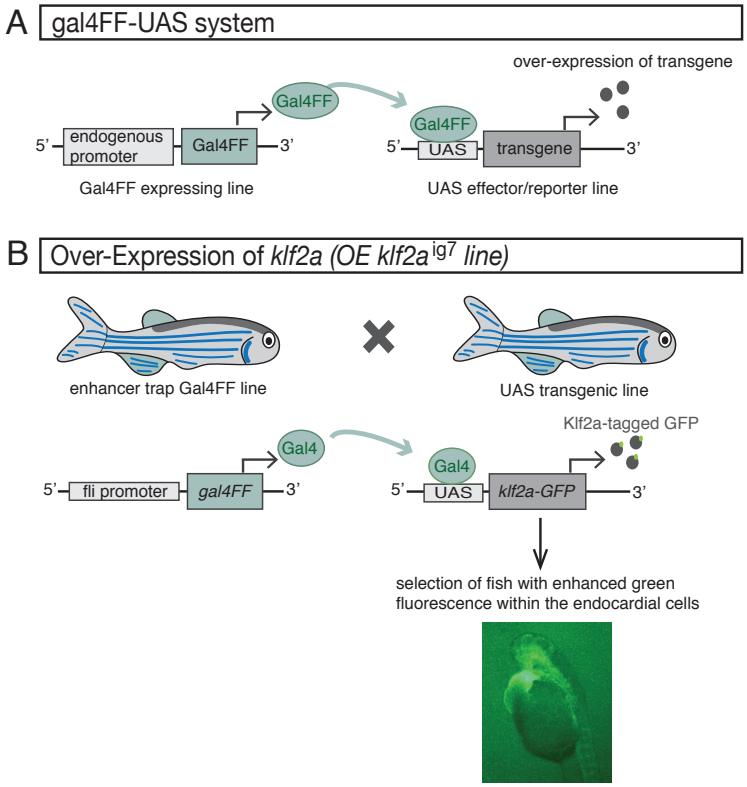


Figure 2: Genetic tools in ZF. Use of the Gal4FF/UAS system to express endogenous sequence in the tissue of interest.

UAS: upstream activation sequence, specific recognition sequence of Gal4FF protein.

fli/fli1a promoter: Friend Leukemia Integration 1a (fli) promoter, specific of endothelial and hematopoietic cells. It widely used to drive the expression of enhanced green fluorescent protein (EGFP) in all blood vessels throughout embryogenesis (Lawson and Weinstein 2002).

Panel A: general presentation of the Gal4FF/UAS system.

Panel B: Example of the use of this system to over-express *klf2a* in endocardial cells.

II.3 Advantages of using the zebrafish model to study heart development

Although the chick and amphibian have provided most of the current knowledge about cardiac development (reviewed by Litvin, 1992), the zebrafish offers non-negligible advantages.

Organ formation in zebrafish is easily accessible at all stages of development due to their optical transparency. This allows simple non-invasive whole-mount observation/imaging of developmental processes, in particular of heart development, as shown on Figure 1. The optical clarity is due either to the absence of pigment ("casper" line, (White et al., 2008)) or the possibility to inhibit pigmentation by adding to the growing medium a non-invasive and non-toxic drug called 1-phenyl 2-thiourea or PTU, a couple of hours after fertilization (*see materials & methods section*).

The oxygenation of the early zebrafish embryo does not rely on heart beat and blood flow during the first days of development (Stainier et al., 1996). Indeed, oxygen can diffuse into the embryo during the early stages of development due to their small size, avoiding the need for an intact circulatory system (Peal et al., 2011). This small vertebrate is therefore particularly well suited to study development and in particular cardiac formation. Heartbeat can be stopped temporarily and reversibly using 2,3-butanedione monoxime (BDM, a myosin ATPase inhibitor for few hours without affecting the development during the first days (Banjo et al., 2013). This would not be possible in the mouse, for example, as heartbeat is necessary at any time. Therefore, the zebrafish embryo gives the opportunity to address the role of blood flow-mediated effects in cardiac development.

Moreover, and contrary to zebrafish, higher vertebrates present some inherent difficulty to study their AVC development due to the early lethality that results from essential gene knockouts (Peal et al., 2011). These unique zebrafish characteristics strongly facilitate the study of heart malformations and this for a considerable period of development. Furthermore, these features have allowed the characterization of a large number of cardiovascular mutants and the zebrafish is now a widely recognized model organism to study heart diseases (Lombardo et al., 2015).

III. Cardiovascular system organisation and formation

The cardiovascular system allows blood to circulate through the organism to transport oxygen and nutrients, and to take waste away from all tissues. It is a complex network of vessels: arteries, veins and capillaries. The arteries carry the blood away from the heart, the veins pump it back to it. In addition, the capillaries enable the exchange of oxygen, water and chemicals between the blood and the cells.

III.1 General considerations about the vertebrate cardiovascular system

Cardiovascular system of vertebrates is a closed system. Blood should never leave the vessel network. Vasculature anatomy and processes leading to the complete and functional network are highly similar to those described in higher vertebrates, in particular humans (simplified representation of the zebrafish cardiovascular system is presented in Figure 4).

In all organisms, blood flow is set in motion by rhythmic contractions of a muscular pump, the heart. The adult vertebrate heart is a complex organ, divided into chambers – called atria and ventricles - separated by different sets of valves, which help maintain unidirectional blood flow within the heart and towards the vascular system. The atria act as the receiving chambers pumping the blood through the ventricles, the discharging chambers.

In mammals, the cardiovascular system is organized into a double circulation system (Figure 3). The heart is composed of four chambers: two atria and two ventricles. The oxygenated blood comes from the lungs and enters the left ventricle via the left atrium, and is pumped out into the body. The deoxygenated blood returns to the heart through its right side, entering the right atrium and ventricle to be pumped into the lungs.

III.2 Presentation of the zebrafish cardiovascular system

III.2.a) Anatomy of the ZF cardiovascular system

Danio rerio has a simpler cardiovascular system: a single pattern, wherein the blood passes through the heart only once during each complete circuit. Their heart is homologous to the human one in terms of constitution, except it has only two chambers: one atrium and one ventricle (Figure 5). However, recent studies show a possible laterality in the zebrafish atrium (Guerra et al., 2018) mimicking both atria in mammals.

A simplified anatomy of the zebrafish cardiovascular network is presented on Figure 4 and a complete anatomical description of it with a precise nomenclature of vessels was established and reviewed by (Isogai et al., 2001). Oxygen-deprived blood from the body tissues comes to the heart, from where it is pumped to the gills. Gaseous exchange happens within the gills, and the oxygenated blood from the gills is circulated throughout the body.

As in mammals, the cardiac pump is composed of three layers of cells: the first one, the muscle tissue layer myocardium, is able to contract and forms a thick middle layer between the outer mesothelial envelope, the epicardium, and the inner and non-contractile layer, the endocardium, in direct contact with the blood. Myocardium and endocardium are separated by an elastic, cell-free layer known as the cardiac jelly.

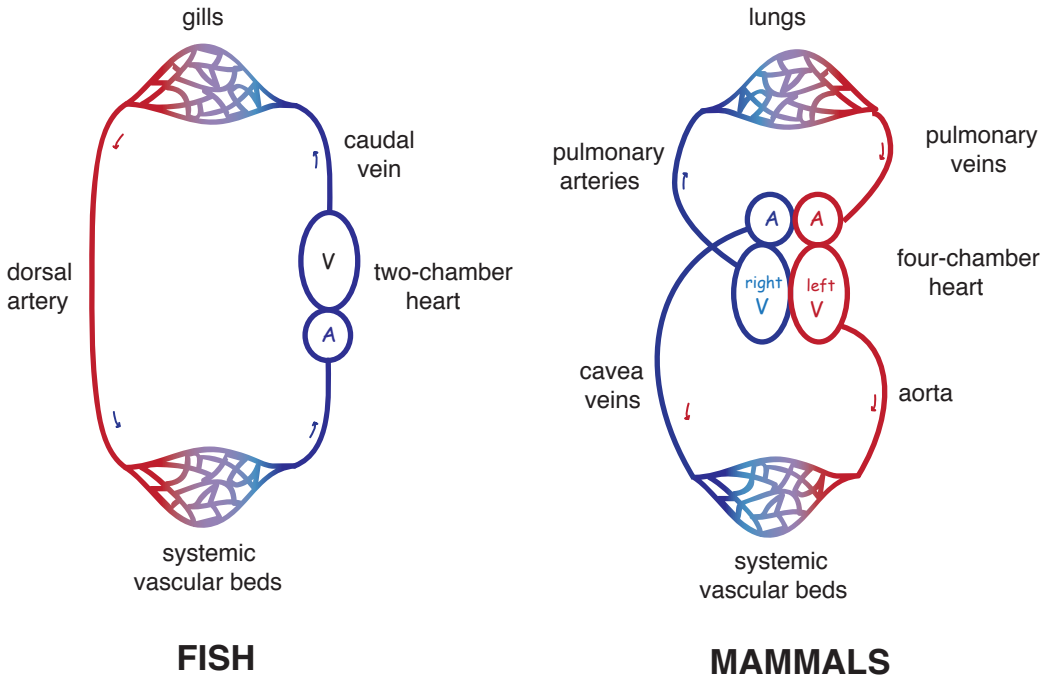


Figure 3: comparison between the circulatory systems of fish and mammals.

A: atrium, V: ventricle

In mammals, the cardiovascular system is organized into a double circulation system. The heart is composed of four chambers: two atria and two ventricles. The oxygenated blood comes from the lungs and enters the left ventricle via the left atrium, and is pumped out into the body. The deoxygenated blood returns to the heart through its right side, entering the right atrium and ventricle to be pumped into the lungs.

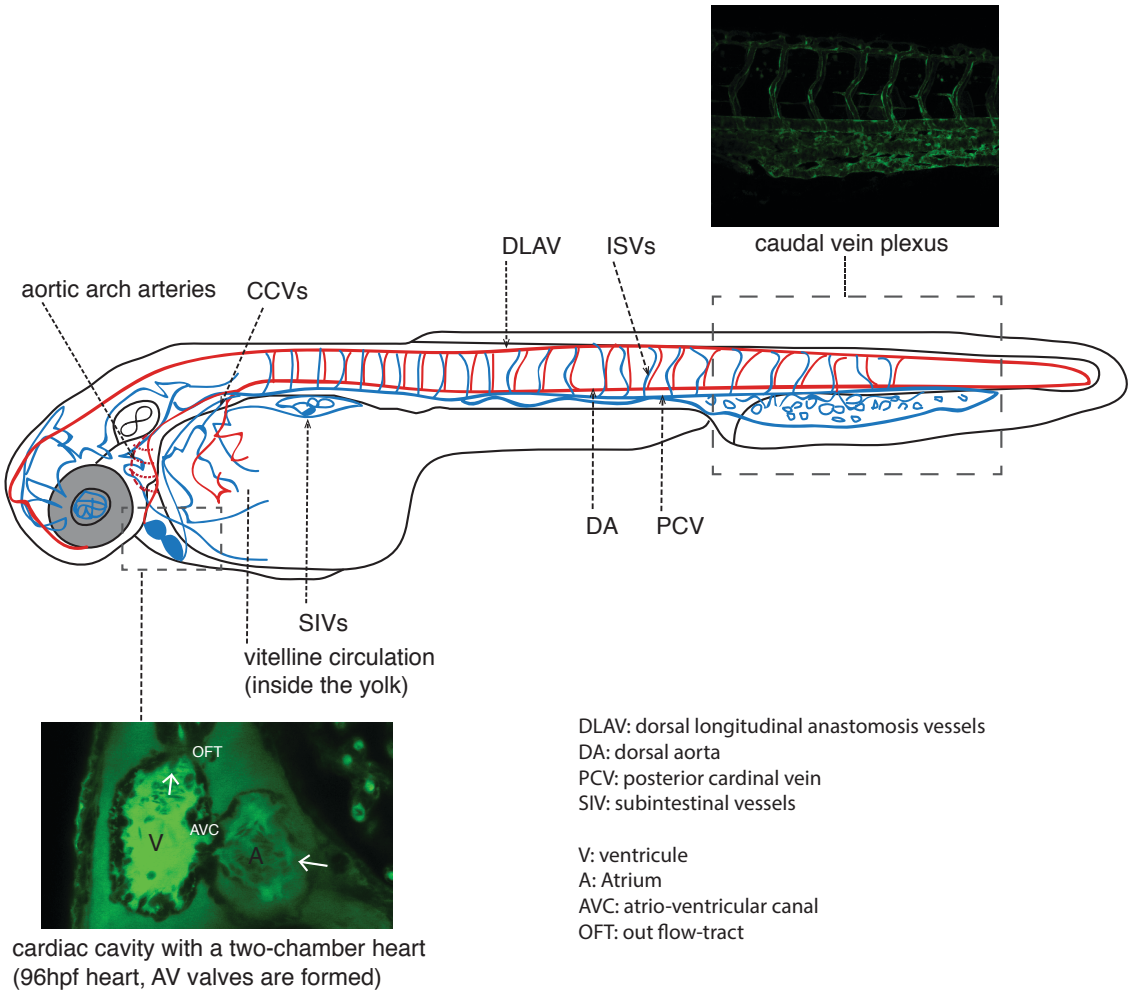


Figure 4: Simplified diagram of the ZF cardiovascular system overall organization.
48hpf-zebrafish embryo, embryo length: approx. 3.1mm.

Early zebrafish embryonic vascular development begins approximately at 12hpf, when a precursor pool of cells from originates in the lateral plate mesoderm and will become later, at approximately 24hpf, the dorsal aorta (DA) and dorsal vein (DV) forming the first circulation loop. Blood exits the heart through the OFT into the bulbus arteriosus and the ventral aorta, passes through the aortic arches and continues into the DA. This single media vessel runs into the tail and turns 180 degrees at its caudal most end to empty into the posterior caudal vein (PCV), a single medial unpaired tube. Between both vessels, blood flow circulates through the intersegmental vessels (ISVs). Post-ventral vasculature forms an interconnecting network of venous tubes, called caudal vein plexus, formed by angiogenesis. Blood is then collected in the common cardinal veins (CCVs) and transported back to the heart. At 48hpf stage, the embryo still depends on lipid nutriments contained inside the yolk. The heart tube has looped and beats vigorously, the two chambers – atrium and ventricle – become separated from the constriction inside the Atrio-ventricular canal (AVC) region. Endothelial cushions are beginning to form in this area.
Adapted from (Garcia-Caballero et al., 2018) and (Lawson and Weinstein 2002).

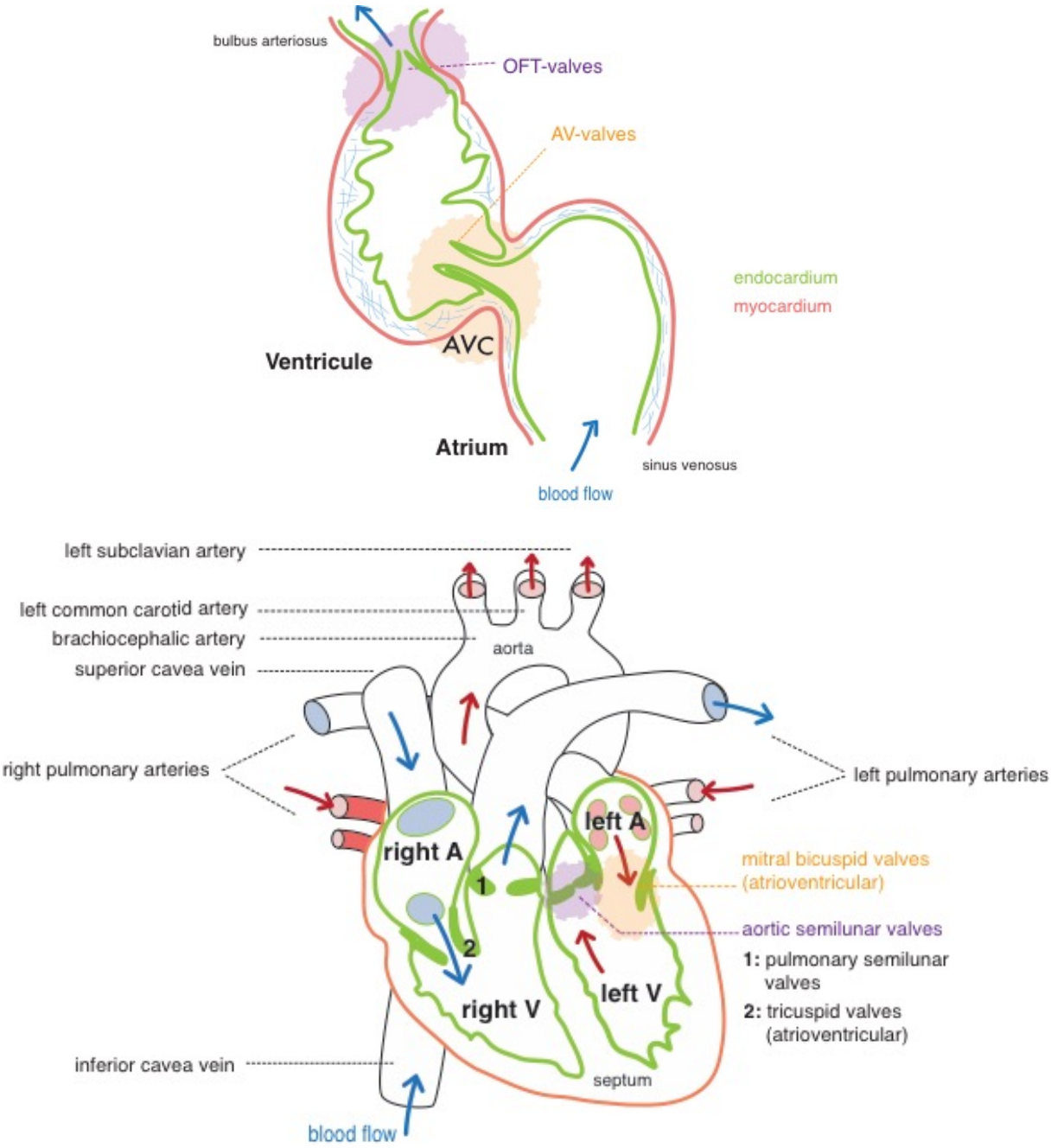


Figure 5: Comparative anatomical schemes of a zebrafish and human hearts.
Simplified zebrafish heart on the upper part at 96hpf and adult human heart on the lower part

In zebrafish, the Atrio-Ventricular Canal (AVC) separates the two cardiac chambers, the atrium (A) from the ventricle (V). Valve leaflets form from invagination and tissue remodeling from the endocardium (green layer) in this area.

III.2.b) General structural key steps in the zebrafish cardiovascular development

The developmental steps of the zebrafish heart can be extrapolated to the human's. As in higher vertebrates, the zebrafish heart is one of the first organs to be formed and functional during embryogenesis. It starts beating peristaltically at 24 hours post-fertilization (hpf) when the blood flow is very low, and reaches its mature configuration within 5 days post fertilization (dpf).

Cardiac development begins with cardiac cell specification at the blastula stage (5hpf) (Stainier et al., 1993). The heart first forms as a linear simple tube, incompletely lumenized, which lacks septation and chambers. Shortly after, the lumen starts opening and the stroke volume and frequency quickly increase (Boselli et al., 2015). The linear tube then undergoes a conformational change known as looping, with a displacement of the ventricle to the embryo's right-hand side, and by 36hpf the cardiac looping is done. Beatings are now 180beats/min and provide a strong circulation to the trunk and head (Stainier et al., 1993).

With the onset of blood flow, the two cardiac chambers greatly expand in size in a process called cardiac ballooning (between 30 and 54hfp), which occurs through a continuous accretion of cells at the arterial and venous poles of the heart (Dietrich et al., 2014). At the same time, the constricted region between the forming atrium and the ventricle specializes, forming the so-called Atrio-Ventricular Canal (AVC), where one set of valves will emerge. As the heart loops, it indeed becomes less efficient at preventing retrograde flow / reversing flow, requiring functional valves (Liebling et al., 2006).

Embryonic heart growth requires the coordinated expansion and patterning of two major cell types, endothelial cells that line the lumen of the cardiac chambers and contractile myocardial cells that pump blood (Zhou et al., 2015).

III.2.c) Key cellular and molecular steps in zebrafish heart development

The heart results from the accretion of distinct progenitor cell populations, which have specific regional contributions to its development.

In higher vertebrates, the earliest population of cardiac progenitors is called the first heart field (FHF) and originate from the anterior lateral plate mesoderm (ALPM). A midline fusion of these mesodermal cells leads by differentiation to the initial tube myocardium formation and expression of muscle-specific proteins. In mammals, these progenitors will form the left ventricle and parts of the atria. Then, the heart expands by the addition of Second Heart Field (SHF) cells, late-differentiating mesodermal progenitor cells, accrete new myocardium to the heart arterial and venous poles. They contribute to the smooth muscles and myocardium of the atria, right ventricle, and OFT. (Stainier et al., 1993) (Lescroart et al., 2010); (Felker et al., 2018). Several studies have shown the conservation of these mechanisms in zebrafish. A population of late-differentiating cardiac progenitors gathering at the arterial pole was shown to be analogous to the mammalian SHF (Knight and Yelon 2016).

At the early cardiac crescent stage, a bilateral expression of conserved cardiac transcription factors (Nkx2.5, Gata4/5, Hand2) can be detected defining the first and second HF (Serbedzija et al., 1996) (Lazic and Scott 2011) (Yelon 2001).

The fusion of the progenitors occurs by 14–18 hpf, the subsequently linear tube composed of endocardium and the surrounding early-differentiating FHF cardiomyocytes emerges at 16–18 hpf under expression of differentiation markers including *myosin regulatory light chain 7* (*myl7*) (Felker et al., 2018); (Stainier et al., 1993) (Bakkers 2011). Starting from 26 hpf, a late-differentiating wave of cardiomyocytes occurs. These progenitors from the SHF are patterned along the anterior-posterior axis of the embryos. Anterior cells, adjacent to the arterial cardiac pole, contribute to the formation of the right ventricle and the OFT myocardium. Whereas the posterior SHF, adjacent to the venous pole, contributes to the venous inflow tract (IFT) atrial myocardium and the sinoatrial node (Knight and Yelon 2016); (Felker et al., 2018).

III.2.d Further information about genetic regulation of heart formation

During the early somitogenesis stage and the first steps of cardiac specification, retinoic acid (RA) signalling has a potent repressive role, limiting the density of myocardial progenitor cells. It later influences cardiac chamber identity, terminal myocardial differentiation, cardiac looping, and ventricular maturation and growth (Keegan et al., 2004).

Some important factors regulate SHF. Among them, the Nkx. Nkx genes encode homeobox-containing transcription factors. *Nkx2.5 and nkx2.7* (two zebrafish homolog of mice *Nkx2.5)* are expressed in cells of the FHF and SHF (Stanley et al., 2002), and regulate cardiac development (Guner-Ataman et al., 2013), *Nkx2.5* is associated with a myriad of human CHDs (Bruneau 2008). They restrict the proliferation of anterior SHF progenitors in OFT, will delimit later the number of posterior SHF progenitors at the venous pole and pattern the sinoatrial node acting through *Isl1* repression (Kelly 2005); (Colombo et al., 2018). *Isl1* is a member of the Islet family. Isl1 mutant hearts are lacking outflow tract, right ventricle, and have a severe reduction in atrial tissue (Lin et al., 2006). By regulating *sonic hedgehog (shh)* expression, Isl1 participates to the morphogenesis at the anterior pole of the heart, in particular the formation of the aortic arch and outflow tract formation (Lin et al., 2006).

Hedgehog (Hh) signalling is required for endocardial differentiation (Wong et al., 2012). In the absence of Hh signalling, endocardial progenitors fail to migrate to the midline and do not initiate endocardial *nfatc1* expression, while endothelial differentiation of blood vessels is not affected. This argues that Hh is a specific signal required for endocardial differentiation. It was also shown to activate fibronectin expression (Wong et al., 2012).

FGF signalling also regulates heart formation by modulating the expression of *myocyte enhancer factor 2cb (mef2cb*), expressed in the late ventricular region, and necessary for late myocardial addition to the arterial pole (Lazic and Scott 2011).

Two members of GATA family, *gata4* and *gata5*, are expressed in endocardial cells[3]. *gata4* expression precedes *gata5* and is essential for initiation of endocardial-endothelial differentiation, *gata5* rather appears to be required for progression of the differentiation

[3] GATA5 expression was also detected in few myocytes, suggesting a potential cell autonomous role in the myocardium (Nemer and Nemer, 2002)

program. Inhibition of either *gata5* expression or *nuclear factor of activated T-cells cytoplasmic* (NF-ATc) activation blocks terminal differentiation at a pre- endocardial stage. GATA5 and NF-ATc seem to synergistically activate endocardial transcription (Nemer 2002). GATA5 also functions upstream of *nkx2.5* to initiate myocardial differentiation.

Bmp and Wnt signals are also similarly required for endocardial and myocardial differentiation *in vitro* (Misfeldt et al., 2009).

IV. Cardiac valves

Cardiac valves are the "doors of the heart". These fine membranes open and close according to the pressure on each side of the valve and act as a physical barrier to blood flow.

Zebrafish heart contains two types of valves: mitral or Atrioventricular Valves (AV valves) located in between atrium and ventricle and define the Atrioventricular Canal region (AVC). The second set of valves, the pulmonary and aortic valves, called Out Flow-Tract valves (OFT valves) in zebrafish. They separate the ventricle from the aorta and permit the blood to be expelled within the general circulation through aorta.

In the human heart, the AV or mitral valves can be distinguished between the tricuspid valve, with three leaflets separating the right atrium and ventricle; and a bicuspid mitral valve separating the left-sided chambers. A comparison between both cardiac systems is presented on Figure 5.

The three-dimensional structure of the valve leaflets is key for efficient gating. Cardiac valves are usually tricuspid (three leaflets) or bicuspid (two leaflets), and their function relies on the fact that one leaflet is longer than the other(s) for the closing to be efficient. Valve mis-function occurs when the valves fail to open properly (stenosis) or do not shut completely (regurgitation).

Due to their essential role, valve formation is an important and controlled process during heart development, modulated spatiotemporally by a precise gene expression pattern. To date, the developmental mechanisms controlling the number and the three-dimensional shape of the valve leaflets are still not completely elucidated. In particular, the precise genetic signalling pathways that dictate valve formation and the interactions with the mechanical forces at work in the developing heart are not well understood.

IV.1 State-of-art on the cellular processes involved in the zebrafish valve formation

Zebrafish valvulogenesis starts from 48hpf when the endothelial ring is forming in AVC. The reversing flow is still quite high at this stage (Scherz et al., 2008), estimated at 30% with a heart rate between 1.5 and 2 Hz (Vermot et al., 2009). Valvulogenesis will be complete around 96hpf when the leaflets can actually be seen extending into the AVC lumen (Stainier 2002) and (Beis et al., 2005).

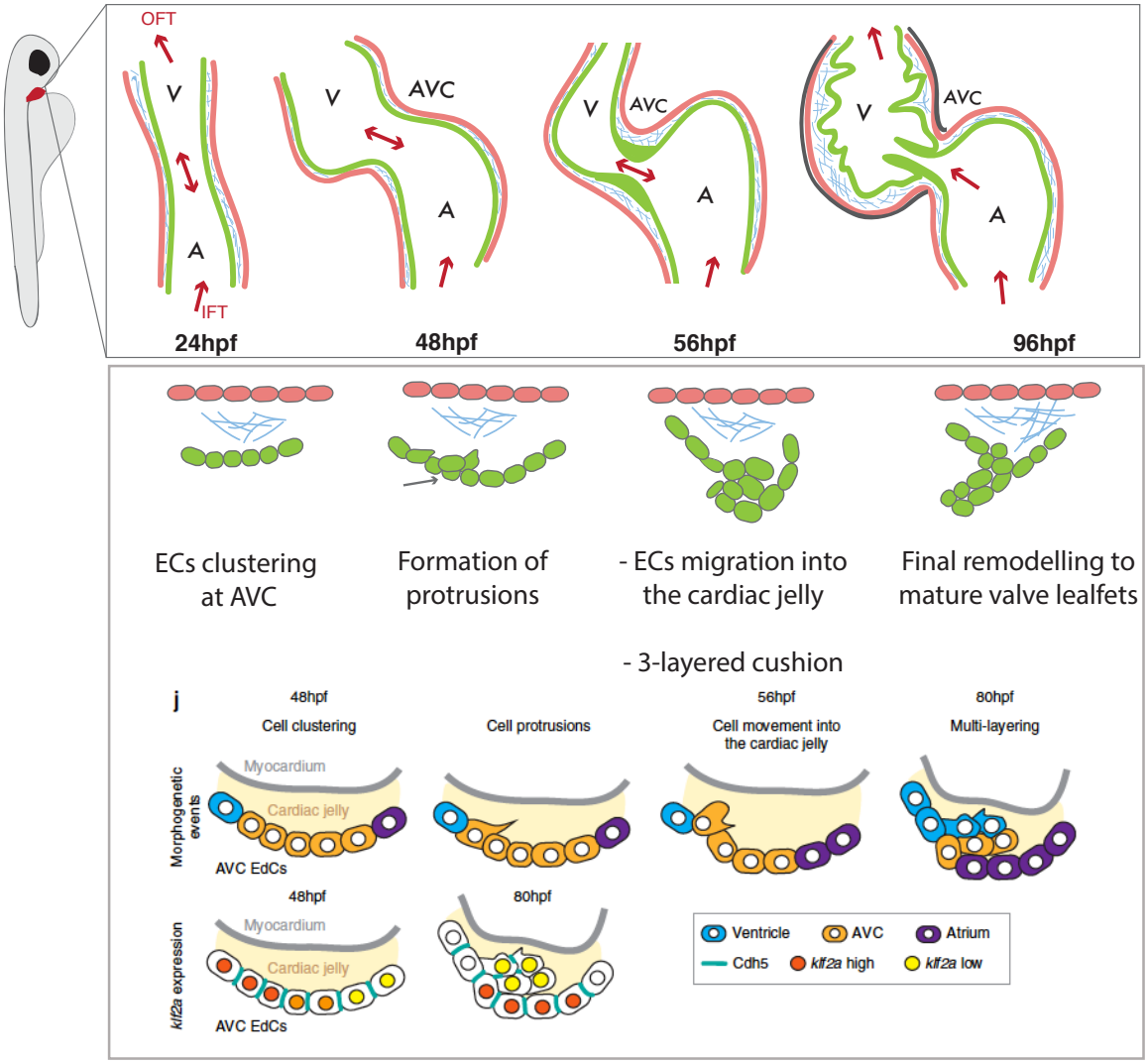
An initial and crucial step in their formation is the swelling of the adjacent ExtraCellular Matrix (ECM), also termed cardiac jelly, around the AVC. In mice and chicks, the composition of ECM changes as the heart develops (reviewed by (Armstrong and Bischoff 2004) and is mainly a hyaluronan (HA)-rich gelatinous matrix. This swelling is due to the delamination, subsequent migration and proliferation of endothelial cells (ECs) into the ECM (Beis et al., 2005) ; (Scherz et al., 2008) to form so-called endocardial pre-valvular cushions. These transient structures are progressively remodelled into mature valve leaflets. The ECs undergo an epithelial-to-mesenchymal transition (EMT) under a signal received from the overlying myocardial cells. This process was first thought to be similar in zebrafish (Beis et al., 2005) as several markers of EMT, such as *notch1b* (Vermot et al., 2009) and (Beis et al., 2005); (Timmerman et al., 2004), *nfat* (Chang et al., 2004), *wnt* (Hurlstone et al., 2003) and *erbB (Goishi et al., 2003)* are expressed in the zebrafish AVC as in mice. But some studies (Scherz et al., 2008) have shown zebrafish heart valves appear to emerge at the AVC boundary directly through an invagination of the AV endothelium rather than from mesenchymal cushion. Steed et al. (2016) [4] analysed the cellular processes leading to leaflet formation. The ECs first cluster at the AV boundary and form an endothelial ring lining the AV canal between 24 and 48hpf.

[4] Paper available in Annex 1

This initial clustering is followed by the appearance of cellular extensions (filopodia) towards the cardiac jelly. A subsequent tissue remodelling occurs to form multiple cell layer-leaflets composed of ventricular and AVC-derived ECs within the cardiac jelly, which overlie the atrial-derived ECs exposed to the lumen. Figure 6 presents a comparison of both processes.

Endocardial and myocardial cells were shown to differentiate morphologically before the onset of EMT/endocardial ring formation (Beis et al., 2005). From 36hpf, AV ECs undergoes a transition from squamous to cuboidal cell shape and start laterally expressing Dm-grasp, a cell-adhesion molecule Cuboidal AV. They also present a characteristic pattern of actin and catenin staining indicative of adherent junctions. By 55 hpf, the superior and inferior regions of the AVC are each lined by a sheet of cuboidal ECs interconnected by squamous cells. This spatial organisation could help preventing blood regurgitation prior to the formation of valves (Beis et al., 2005).

A


OFT
V
A
IFT
24hpf
V
AVC
A
48hpf
V
AVC
A
56hpf
V
AVC
A
96hpf
ECs clustering at AVC
Formation of protrusions
- ECs migration into the cardiac jelly
- 3-layered cushion
Final remodelling to mature valve lealfets
j
48hpf
Cell clustering
Cell protrusions
56hpf
Cell movement into the cardiac jelly
80hpf
Multi-layering
Morphogenetic events
Myocardium
Cardiac jelly
AVC EdCs
klf2a expression
48hpf
80hpf
Ventricle
AVC
Atrium
Cdh5
klf2a high
klf2a low

B

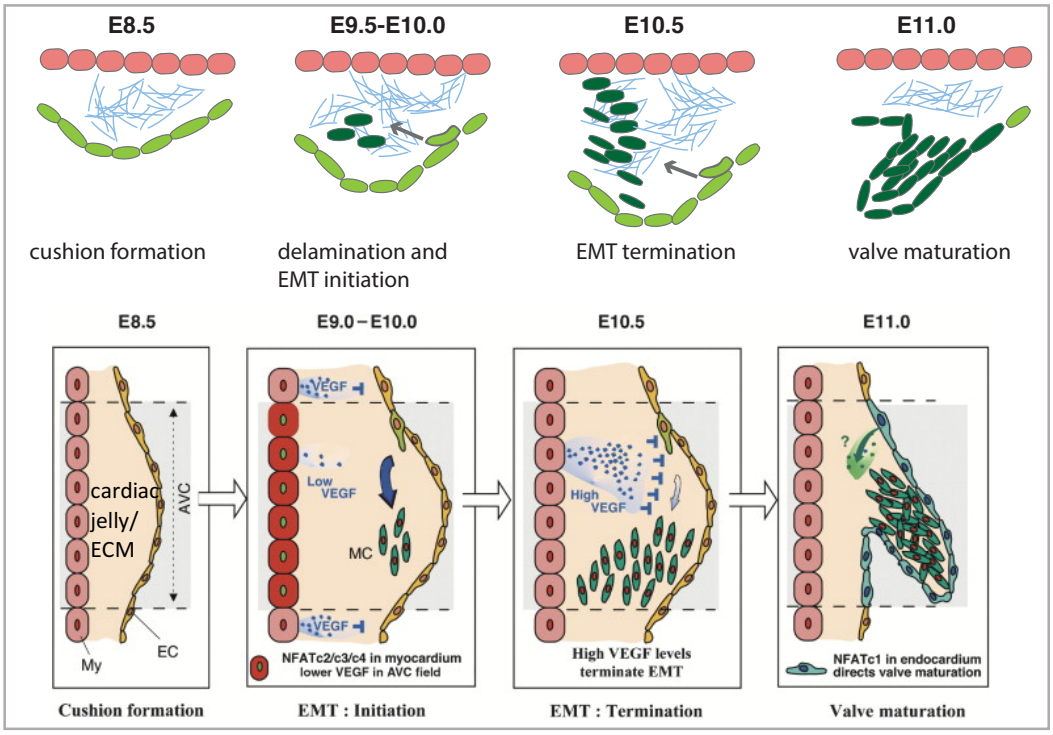

E8.5
E9.5-E10.0
E10.5
E11.0
cushion formation
delamination and EMT initiation
EMT termination
valve maturation
E8.5
E9.0 – E10.0
E10.5
E11.0
cardiac jelly/ ECM
AVC
My
EC
VEGF
Low VEGF
MC
NFATc2/c3/c4 in myocardium lower VEGF in AVC field
High VEGF
High VEGF levels terminate EMT
NFATc1 in endocardium directs valve maturation
Cushion formation
EMT : Initiation
EMT : Termination
Valve maturation

Figure 6: Models of the early stages of AV valve development

A: Zebrafish valvulogenesis. different steps leading to the formation of mature valve leaflets are summarized on the scheme. The primitive linear tube undergoes a change of conformation with the displacement of the ventricle of the right side of the embryos. Atrium and ventricle grow in size in a process called cardiac ballooning. EMT does not seem to occur in zebrafish. In contrast, ECs clustered at AVC and rearrange to form multi-layered cushions invading the cardiac jelly. Valve come from an invagination of endocardium. *(Figure from (Steed et al., 2016)– paper available at the end of chapter 1).*

B: Mice valvulogenesis. Endocardial cells (ECs) delaminate and migrate into the cardiac jelly/ Extra Cellular Matrix (ECM) where they undergo a transition to become mesenchymal cells (MC) under signals received from the myocardium (My) to populate and give birth to valve cushions. This process is called Endocardial-Mesenchymal Transition or EMT. *(Figures adapted from Lambrechts and Carmeliet, 2004)*

IV.2. State-of-art on the genetic control of zebrafish valve formation

The formation of functional valves, from the initiation of ECs migration to the global tissue remodelling, is regulated by a unique genetic program. The expression of genes specific to AVC depends on the communication between molecules secreted in the subjacent AV myocardium and the endocardium. An elaborate combination of signalling pathways between the two cell layers occurs.

Figure 9 summarizes then main genetic and molecular pathways occurring during AV valve formation.

IV.2.a) Notch and Bmp signalling: bidirectional signalling in zebrafish AVC

The myocardial and endocardial cells in the developing zebrafish AVC appear to communicate to each other across the cardiac jelly, specifically through the myocardium-derived Bone Morphogenetic Protein (BMP) and Notch signalling pathways (Vermot et al. 2009).

BMP already demonstrated as crucial for endocardial proliferation during the ballooning phase (Dietrich et al., 2014) is also required for EMT and for the formation of endocardial cushions and valves (Jiao et al., 2003); (Rivera-Feliciano and Tabin 2006) (Garside et al., 2013). *Bmp4* is expressed throughout the primitive myocardium, and becomes restricted from 37hpf to AVC myocardial cells (Walsh and Stainier 2001). It inhibits the expression of chamber-specific genes in the AVC region, and initiates the expression of another important markers of endocardial differentiation – nuclear factor of activated T-cells, cytoplasmic, calcineurin- dependent 1 - *nfact1* (Palencia-Desai et al., 2015). BMP also encodes for different ligand and membrane receptor pairs which leads to the activation of cytoplasmic SMAD proteins. These proteins regulate gene expression in the endocardium of the primitive heart tube by 24hpf, and later in myocardial cells of the AVC by 48hpf (Beets et al., 2016). Tbx2/*tbx2b* gene is one of their targets, it directly binds *Has2* and *Tgf-β2* promoters and increased their transcriptional activities required for EMT / cushions formation (for Has2 see next paragraph for further details ; for mice Tgf-β2 - (Shiraia et al., 2009). *Tbx2/tbx2b* expression pattern makes it as a marker of the AVC myocardium (Walsh and Stainier 2001).

Notch signalling is also involved in similar communication between cell layers and AVC formation. *Notch1b* expression is first detected throughout the endocardium and it becomes restricted to the endocardial AVC valve forming cells by 58 hpf (Vermot et al., 2009). Zebrafish *notch1b* was proved to be one of the potential endocardial signals for myocyte differentiation in the AV ring (Milan et al., 2006). Consequently, Notch signalling is required at early stages to prevent an AV cushion phenotype in the ventricle, and later on to modulate cushion formation in the AVC (Timmerman et al., 2004).

IV.2.b) VEGF signalling: Regulation of endothelial division and ECM biosynthesis

Vascular Endothelial Growth Factor (VEGF) signalling, pivotal regulator of vascular development, is essential for AVC formation at 17–19 hpf, it causes cell maker changes and division. Then calcineurin/NFAT represses its expression from 20–33 hpf (Lee et al., 2006). After myocardial cells migrate to surround endocardial progenitors, they initiate calcineurin-dependent transcription factor (nfatc1) expression in AVC (Wong et al., 2012), making of this gene a specific endocardial marker of the AVC. VEGF also cooperate with TGF-β1 (Lee et al., 2006), which was shown to induce EMT in clonal populations of mouse valve endothelial cells (Paranya et al., 2001). The upregulation of endoglin by TGFβ1 is cooperative with hypoxia, which is an upstream regulator of VEGF, suggesting cross-talk between VEGF and TGFβ1 in valvulogenesis (Lee et al., 2006). VEGF modulates extracellular matrix biosynthesis and in turn, the ECM influences cell behavior during development (Ortega et al., 1998). Cell-restricted expression of *notch1b* and *bmp4* is abolished when VEGF-R signalling was chemically disrupted (Lee et al., 2006). This blocking also mimics the genetic disruption of UDP-glucose dehydrogenase (UGDH) in zebrafish (Walsh and Stainier 2001) raising the hypothesis VEGF-R signaling and hyaluronic acid (HA) biosynthesis may act in an integrated manner during valve development.

Focus on one VEGF receptor, Fgefr1 or Flt1

Zebrafish VEGF receptor proteins are encoded by the *flt1 (Vegf receptor 1, vegfr1), kdr* (previously *kdrb, vegf receptor 2*), *kdr-like* (*kdrl*, previously *kdra*) and the *flt4* (vegfr receptor 3) (Bussmann et al., 2007) ortholog genes. Whereas the roles of *flk1 (also named kdr* and *kdrl*) and of *flt4* in mediating vascular development in zebrafish are well described

(Covassin et al., 2009; Covassin et al., 2006; Habeck et al., 2002; Hogan et al., 2009b; Lawson et al., 2002; Liao et al., less is known about *flt1. Flt1* was first well documented for its essential functions during blood vessels formation. Zebrafish embryos expressed soluble Flt1 protein (sFlt1) and membrane-bound protein (mFlt1). Both mFlt-1 and sFlt-1 are essential to establish and maintain the earliest vessels, a process termed vasculogenesis, as well as to direct later sprouting, termed angiogenesis. They modulate endothelial cell proliferation (Kappas et al., 2008) but only sFlt-1 regulates vessel branching by contributing to a local sprout guidance mechanism and by negatively regulating endothelial tip cell differentiation (Krueger et al., 2011) and (Chappell et al., 2013). *Flt1* is involved in a VEGF-Notch Feedback Loop, acting in a Notch-dependent manner as a negative regulator of and arterial branching formation. The use of MO targeting *flt1* showed reduced expression of Notch receptors and of the Notch downstream target *efnb2a*, and ectopic expression of *flt4* in fish arteries, consistent with the loss of Notch signaling. Conditional overexpression of the *notch1a* intracellular cleaved domain in *flt1* morphants restored segmental artery patterning (Chappell et al., 2013). *Flt1* is also expressed in the neural tube and it was shown that neuronal sFlt1 restricts Vegfaa-Kdrl mediated angiogenesis at the neurovascular interface (Wild et al., 2017).

Beside its involvement in angiogenesis, *flt1* is expressed in the endocardium at 26hpf, and is required during heart development. Although its roles in valvulogenesis have never been studied, *flt1* was shown to act upstream of the phospholipase C gamma 1 (PLCgamma1) to control cardiac ventricular contractility (Rottbauer et al., 2005). The VEGF–FLT1–PLC1 signalling pathway function in a reversible and rapid manner by modifying cardiomyocyte calcium cycling. In this thesis, I studied the *flt1*^{sa1504} mutant and found severe AV valve defects. I also uncovered a crosstalk with *egr1* and *klf2a* signaling pathways (*see chapter 3*).

IV.2.c) Signaling pathways in cardiac jelly, role of the ECM

An appropriate ECM environment in the AVC is necessary to provide a scaffold to support normal endocardial cushion formation (Camenisch et al., 2000) ; (Schroeder et al., 2003).
Some genes encoding for important components of ECM and cardiac jelly, like *ugdh*, *versican* and *has2* genes, undergo similar spatio-temporal changes as *notch1b* and *bmp4* in the AVC (Peal et al., 2011). *Versican-a* and *has2* genes becomes restricted to AVC in 48hpf-embryos (Hurlstone et al., 2003). Versican is a HA-binding proteoglycan, and Has2 synthetizes

hyaluronan or acid hyaluronic (HA). HA was shown to be important for cell migration into the cardiac jelly and for mice EMT in combination with Ras signalling and also involving CD44, and ErbB receptors/ligands (Camenisch et al., 2000). Both *versican a* and *has2* are transcriptional targets of the Wnt/β-catenin pathway (Hurlstone et al., 2003). Another gene, expressed in AVC endothelial cells, *crip2*, was discovered to have an antagonistic role (Kim et al., 2014) by downregulating their expression. A cross-talk between Crip2 and Wnt signaling allows the production of normal amount of Versican and HA. Indeed, either deficiency or over-expression of versican and has2 causes heart valve defects (Walsh and Stainier 2001) (Hurlstone et al., 2003) (Schroeder et al., 2003). Another signalling pathway, Wnt-independent, is also regulating Has2: the mediator complex subunit 10 (Med10) mediates Forkhead box protein N4 (Foxn4) signals and thereby activates the expression of *tbx2b* in AV myocardial cells; *tbx2b* which in turns regulates Has2 (Just et al., 2016).

Versican regulation is also downstream of Cerebral Cavernous Malformation (CCM) pathway. A role of CCM signaling in the developing heart was revealed by zebrafish embryos (Mably et al., 2003) (Mably et al., 2006) (Zheng et al., 2010). Its name is linked to its discovery through genetic studies of human patients with familial vascular malformations (Chan et al., 2010) ; (Riant et al., 2010). Loss of endocardial CCM signaling results in embryonic heart failure and reduced myocardial growth that is characterized by loss of cardiac jelly. Zhou et al. (2015) demonstrated that this phenotype was shown to be caused by an increased expression of *Klf2* and *Klf4* and of *Adamts4* and *Adamts5* genes endocardial secreted proteases that degrade versican and were already shown to be important in the regulation of cardiac jelly and heart valve formation (Dupuis et al., 2011) ; (Stankunas et al., 2008). CCM signaling plays a critical role in cardiac development through the regulation of the MEKK3/MAPK signaling pathway which then controls the expression of downstream KLF2/4 and ADAMTS. ADAMTS which finally regulates versican protein levels in the cardiac jelly modulating its degradation (Figure 7) (Zhou et al., 2015).

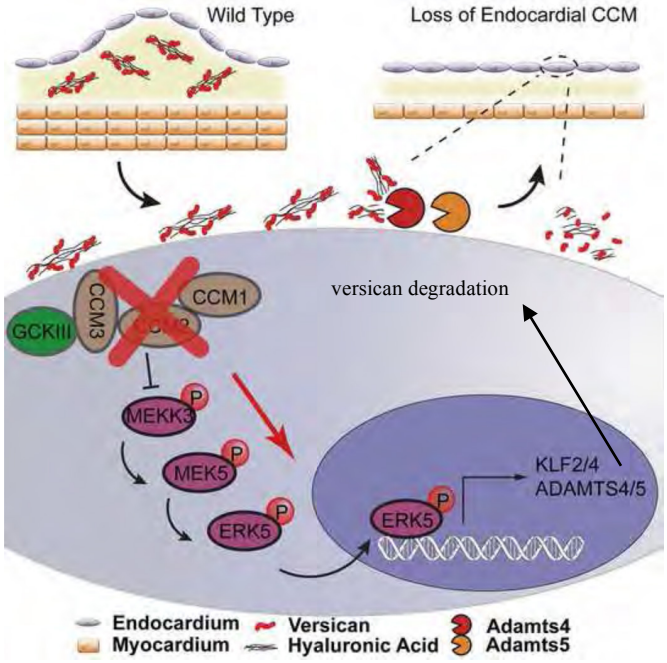


Figure 7: Cardiac jelly protein versican as a downstream target of CCM pathway.

The Cerebral Cavernous Malformation (CCM) pathway controls endocardial gene expression through regulation of MEKK3 signaling and KLF2/4 expression during valve cushion formation. CCM signaling controls the degradation of cardiac jelly (versican mainly) by negatively regulating endocardial expression of ADAMTS4/5 proteases.
(*From (Zhou et al., 2015)*

UDP-glucose dehydrogenase (UGDH)[5] is required to synthesize three important glycosaminoglycans (GAGs) of the cardiac jelly from the conversion of UDP-glucose: chondroitin sulfate (CS), heparan sulfate, and hyaluronic acid (HA) (Lander and Selleck 2000). This process is also regulated by retinoic acid pathway (Morton et al., 2008). GAGs synthetis are important to settle a boundary between the atrium and the ventricle and patterning the developing AVC. They help forming a hydrated gel and regulating ligand availability, while interacting with ECM components as versican (Armstrong and Bischoff 2004).

micro-RNAs (miRs) are also crucial molecular modulators during development[6]. Some of them were shown to be involved in cardiac jelly remodelling. For example, miR-138, this micro-RNA downregulates versican and retinoic acid dehydrogenase (raldh1a2) from VEGF cascade within the ventricle to keep them expressed only around the AVC (Morton et al., 2008). The response to flow patterns is also mediated by small RNAs, often referred to as mechano-miRs (Kumar et al., 2014). miR-21 is able to respond rapidly to blood flow, its expression in regions of high shear stress suppresses a number of target genes that would otherwise impair valve formation (Banjo et al., 2013). miR-143 is expressed within the OFT and ventricle in a flow-dependent manner too (Miyasaka et al., 2011), it targets retinoic acid (RA) signalling pathway components. Within endothelial cells in culture, Klf2 binds to the promoter region of the mir143/145 cluster to induce expression (Hergenreider et al., 2012). *(See paragraph IV.3 for further details about biochemical forces during cardiac development)*

Another gene seems to be important for the remodelling endocardial cushion matrix components: bHLH gene heart and neural crest derivatives expressed transcript 2 or hand2. Early in development, it regulates the epithelial polarity of myocardial precursors in zebrafish (Trinh 2004) and it was shown to be required for proper morphogenesis of the right ventricle and outflow tract, and to sustain mice cardiac development (Holler et al., 2010) (Laurent et al., 2017).

[5] "Jekyll mutation" – Walsh and Stainier, 2001
[6] Role of miR-126 acting downstream of klf2a to drive flow-stimulated angiogenesis. – Nicoli et al., 2010

IV.2.d) Wnt/ β-catenin pathway

Wnt signalling - though the binding with its ligands - controls the translocation of β-catenin into the nucleus to allow for expression of target genes. Among them, there are the T-Cell Factor (TCF) transcription factors (Hurlstone et al., 2003). Wnt/b-catenin signalling is operative in myocardial and endocardial cells only at AVC and regulates *bmp4*, *has2* and *versican*. Wnt/ ß-catenin regulates subsequent expression of valve markers as well as proliferation and transdifferentiation of endocardial cells to establish endocardial cushions (Hurlstone et al., 2003). However, high resolution imaging leads to the proposal that constitutively active Wnt/β-catenin signaling not only permits an increased cell proliferation but also increased endocardial cell movement towards the AVC cardiac jelly.

Wnt/β-catenin signalling was shown to be downstream of NOTCH signalling (Wang et al., 2013) and upstream of BMP (Verhoeven et al., 2011). For example, the complex Jagged1-Notch1 induces the mice endocardial expression of Wnt4, which subsequently acts as a paracrine factor to upregulate Bmp2 expression in the adjacent AVC myocardium to signal EMT (Wang et al., 2013). Verhoeven et al. demonstrated Wnt/b-catenin signaling is both sufficient and required for the induction of *Bmp4* and *Tbx2b* expression in the AVC and consequently the proper patterning of the myocardium. negative feedback loops were discovered between these two major signaling pathways: Id4 (from the Id family of transcriptional bHLH repressors) could restrict endocardial Wnt/β-catenin signalling in response to BMP signal (Ahuja et al., 2016)

Recent study reveals that one WNT members, the endocardial Wnt9b, "*is required to slow mesenchymal cell proliferation and drive the sub-endocardial mesenchymal cell condensation by which a bulky cushion is remodeled into a mature valve*" (Goddard et al., 2017). Further details about the role of *wnt9b* in AVC valve formation and its interaction with another important transcription factor, Klf2/Klf2a, will be presented in the following part IV.3. I also studied in this thesis the regulation of *wnt9b* by Egr1.

IV.3. The role of biomechanical forces in cardiac valve development

First hint that mechanical stimuli are essential for AVC development was the study of the early cardiac phenotype of *silent heart (sih)* mutant embryos - which establish neither a heartbeat nor blood flow due to mutation of *cardiac troponin T (tnnt2/tnnt2a*) gene, but do undergo looping morphogenesis (Sehnert et al., 2002) and fail to form an endocardial ring at the AV boundary (Bartman et al., 2004). Bartmann et al. noticed a specific absence of EC formation in these embryos. This clearly demonstrates that myocardial function is required for EC formation. However, they did not observe any change in the expression of *bmp4*; their hypothesis was that myocardial function was somehow required for another aspect of this signaling event.

It has been then shown that valve formation relies on a coordinated interplay with mechanical forces generated by the blood flow (Hove et al., 2003). Blood flow seems to not be necessary for endocardial ring formation, but rather critical for cell shape change and leaflet invagination (Vermot et al., 2009). Numerous studies demonstrated that ECs can both sense and transduce biomechanical stimuli, such as wall and fluid shear stress (WSS & FSS) and transmural pressure caused by pulsatile blood flow (first demonstration by (Helmlinger et al., 1991) into internal signals by sensing different patterns of blood flow. How hemodynamic forces are converted to molecular information that are relayed across the endothelium to control sub-endocardial mesenchymal cell, which are away from the blood flow, are not yet completely understood.

Some flow-responsive genes have been identified, revealing specific molecular/ mechanical pathway in heart valve formation. One such gene is the zinc-finger transcriptional regulator Krüppel-like factor 2a, *Klf2a*. Vermot et al. (2009) demonstrated that the endothelial ring assembly causes an increase in amplitude of reversing flows in the AVC. *Klf2a* expression then becomes detectable in this specific region. Its expression dramatically reduces if flow decreases. When *klf2a* is knocked-down valve genesis is altered. *Klf2a* appears to be an important mediator of mechanosensitive responses to blood flow within ECs and thus necessary for normal valve formation. The action of this endocardial transcription factor was shown to be mediated primarily by WNT9B/Wnt9b, a secreted protein which acts as a paracrine effector on underlying mesenchymal cells in mice and zebrafish (Goddard et al., 2017). In mice, the KLF-WNT signalling pathway is required by endocardial cells to convert

hemodynamic forces to paracrine WNT signals that orchestrate the behaviour of cells not in contact with blood to form a mature heart valve, as illustrated in Figure 8. In zebrafish, the endocardial expression of *wnt9b* is absent in *klf2a* mutants (Goddard et al., 2017), suggesting that *klf2a-wnt9b* axis in conserved in vertebrates.

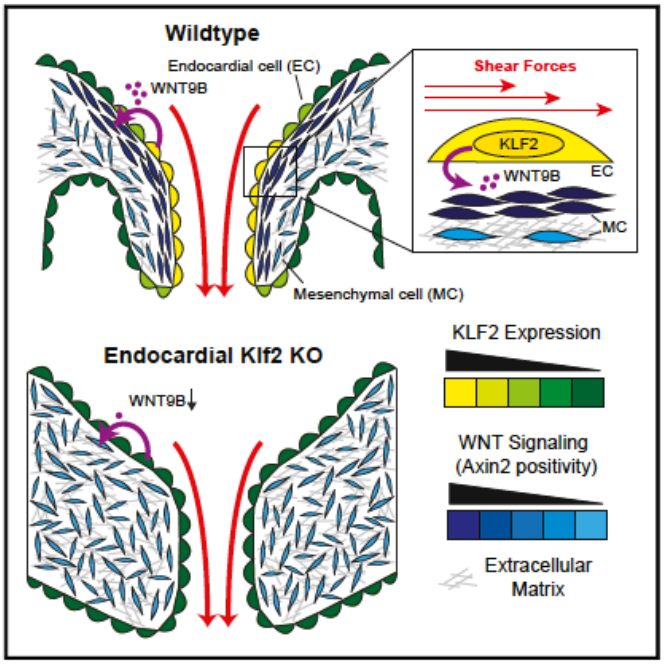


Figure 8: Role of biochemical forces in AV valve development through KLF-WNT signalling.

Hemodynamical forces activate the expression of KLF2/Klf2a which in turn regulates endocardial *wnt9b* expression in the mesenchymal cells. Wnt9b paracrine signal provide primary instructions for the remodeling of the underlying mesenchymal cushion.
(*From (Goddard et al., 2017).*

V. *klf2a*, an important gene for valvulogenesis

Figure 9 illustrates this part.

V.1. Klf2a, a member of Krüppel-life factor family of transcription factors

Krüppel-like factor 2a (*Klf2a*) has been identified by (Oates et al., 2001) as the zebrafish ortholog of KLF2 gene discovered previously in humans (Wani et al., 1999) & (Kozyreva 1999) and mice (Anderson et al., 1995). *Klf2a* has a paralog, *Klf2b*, reflecting the partial duplication of the zebrafish genome after divergence of the teleost lineage during evolution (Taylor et al., 2002), partial phylogenetic tree in Oates et al., 2001), as detailed previously.

Klf2 is part of a multigenic, evolutionarily conserved family of transcription factors (Krüppel-Like Factors or KLF), either activators or repressors, discovered in mammals (Turner and Crossley 1999). The genomic sequences are highly conserved among all members of the family and they are all zinc-finger transcription factors (Anderson et al., 1995). The percentage of identity is much higher in the region of the three tandem zinc fingers, DNA binding domains, common to every KLF member. Each has three tandem C2H2-zinc finger domains in the C-terminus of the protein (Figure 10). The N-termini have some common conserved motifs that act as transactivation or repression domains but generally exhibit much less similarity in their sequence. In some KLF proteins, N-termini contain sites of phosphorylation and acetylation that may contribute to regulation of KLF activity (Zhang and Bieker 1998). The zinc-finger motifs are able to bind to GC-rich sites of general structure (C)CN CNC CCN such as canonical common CACCC-boxes in various promoters and enhancers (Jiang et al., 2008).

Regarding KLF2 homology, the murine and human genes are identical at 85% and their proteins share 90% of identity in their primary structure. The homology between human KLF2 and zebrafish Klf2a is 78% for nucleotide content and 90% regarding the primary structure (Anderson et al., 1995) ; (Oates et al., 2001).

Published and studied zebrafish Klf2a protein contains 380 amino acids (43kDa). However, another form of Klf2a protein could be present in cells. Indeed, there is a second putative ATG in the coding sequence, similar to human which also contains two putative TATA boxes (Kozyreva 1999). The second putative zebrafish start codon is in the same reading frame and would code for a shorter isoform of 347 amino acids (39kDa). To date, it is not known if both forms are expressed in human or zebrafish cells. This will be explored in this thesis.

V.2 Functions of Klf2/*klf2a*

Even if they share common elements in their structural organization, all members of the KLF family present highly restricted patterns of expression in tissues and across development (Oates et al., 2001) controlling cell growth and differentiation.

In adult organisms, KLF2 is involved in a broad range of functions including the control of endothelial identity, vascular integrity, haematopoiesis (Wang et al., 2011, Sangwung et al., 2017). It is essential for cardiovascular physiology in general: vasculogenesis, angiogenesis, cardiac valve morphogenesis, regulation of vascular tone, and control of vasoprotective gene expression (Dekker et al., 2006) ; (Lee et al., 2006) ; (Nicoli et al., 2010) ; (Parmar et al., 2006) ; (Vermot et al., 2009). Furthermore, KLF2 is implicated in the endothelial pro-inflammatory process and over-expression the gene induces key endothelial inflammatory factors such as eNOS (Kumar et al., 2005). This is evidenced by its implication in pathologies such as atherosclerosis and cerebral cavernous malformations (Zhou et al., 2015).

Despite its implication in adult organism functions, KLF2 also plays an important role during embryogenesis. In mice *KLF2* is highly expressed in the vascular endothelium and is critical for lung development and blood vessel formation (Dekker et al., 2006). Kuo et al. (1997) observed that *KLF2* mutant mice die *in utero* due to failure to form stable tunica media in the blood vessels and subsequent haemorrhaging. *KLF2* mutant embryos were shown to present abnormal endocardial cell morphology and hypoplastic AV endocardial cushions. They die before complete valve formation (Chiplunkar et al., 2013). In zebrafish, *klf2a* is expressed in the early epidermis and becomes restricted to cells associated with blood vessels in the head, trunk, and tail (Oates et al., 2001). klf2a is important for hematopoietic stem cell formation (Wang et al., 2011). Klf2a is also essential for angiogenesis regulating some micro RNA (as miR-126) (Nicoli et al., 2010). During the first steps of valvulogenesis, *klf2a* is expressed in valve endocardium and in particular in the AVC (Vermot et al., 2009) (Heckel et al., 2015). The Knock-Down of *klf2a* using morpholino (MO) results in thicker and less flexible valves and increased regurgitation in the developing heart, compared to WT (Vermot et al., 2009). Approximately 10% of *klf2a*^{ig4} mutant embryos (TALEN-KO mutant, published in (Steed et al., 2016)) was missing any kind of valve structure, and 70% demonstrated a range of valvular defects. However, and contrary to mice, adult homozygous fish survive, highlighting

some compensatory mechanisms that should happen to compensate with *klf2a* loss. *Klf2b* gene, the paralog of *klf2a*, could be involved in this compensatory mechanism. Despite its central functions in both mice and zebrafish for valve development, to date, KLF2 has not been linked to valvulopathies in humans.

V.3 Klf2/*klf2a* is a flow-responsive gene

Klf2 was originally described as LKLF (lung KLF) (Kozyreva 1999) (Dekker et al., 2002). It is the first endothelial transcription factor that is uniquely induced by flow in human vascular endothelial cells. was shown to be a key molecular effector of blood flow during valve formation in the developing fish and mice heart (Vermot et al., 2009) (Heckel et al., 2015).

The circulation of blood flow in the vessels create so-called fluid shear forces, tangential forces on the vessel walls. This fluid shear stress is crucial for maintenance of a properly functioning endothelium. It was demonstrated in mice that *KLF2* expression is greatly induced by pulsatile shear stress in microvascular endothelium (Huddelson et al., 2004). In zebrafish, Vermot et al. (2009) and then Heckel et al. (2015) demonstrated that *klf2a* expression is activated and upregulated in response to retrograde flow in the AVC at 48hpf (*a contrario* a downregulation was observed when this reversing flow is experimentally decreased). *Klf2a* is consequently a flow-responsive gene. Its Knock-Down or Knoc-Out results in valvular defects which are similar to those observed in mutants with decreased retrograde flow: thicker and less flexible valves and increased regurgitation in the developing heart (Vermot et al., 2009) ; (Steed et al., 2016) suggesting a role for this transcription factor in blood flow-dependent valvulogenesis as a mechanosensitive transcriptional regulator.

Beside the regulation of *KLF2/klf2a* by CCM/MKK3 pathway (see part IV. of this introduction), recent studies have then demonstrated that *klf2a* expression is also under the control of the protein kinase D2 (PrKD2), which phosphorylates the class II histone deacetylase HDAC5 (Wang et al., 2010). HDAC5 acts as a chromatin modifier and usually represses gene expression when not phosphorylated. The mechano-transduction pathway leading to *klf2a* activation is thought to involve a series of posphorylation events that finally phosphorylate HDAC5 and de-repress *klf2a* gene. *Klf2a* expression was observed to be absent in *prkd2* mutants (also known as *cup* mutants) which do not form valves (Boselli et al., 2015). Heckel

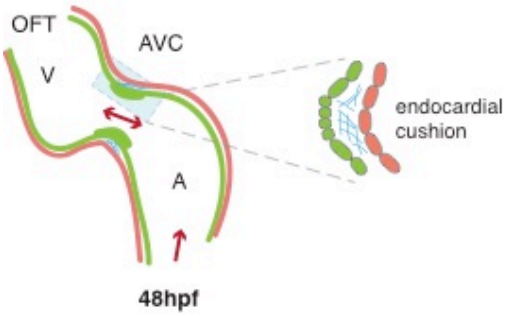
et al. discovered the roles of two membrane-bound mechanosensitive channels (polycystin 2 or Trpp2 and Transient receptor potential cation channel subfamily V member 4 or Trpv4) which lead to the activation of a calcium-activated intracellular cascade resulting in *klf2a* expression (Heckel et al., 2015).

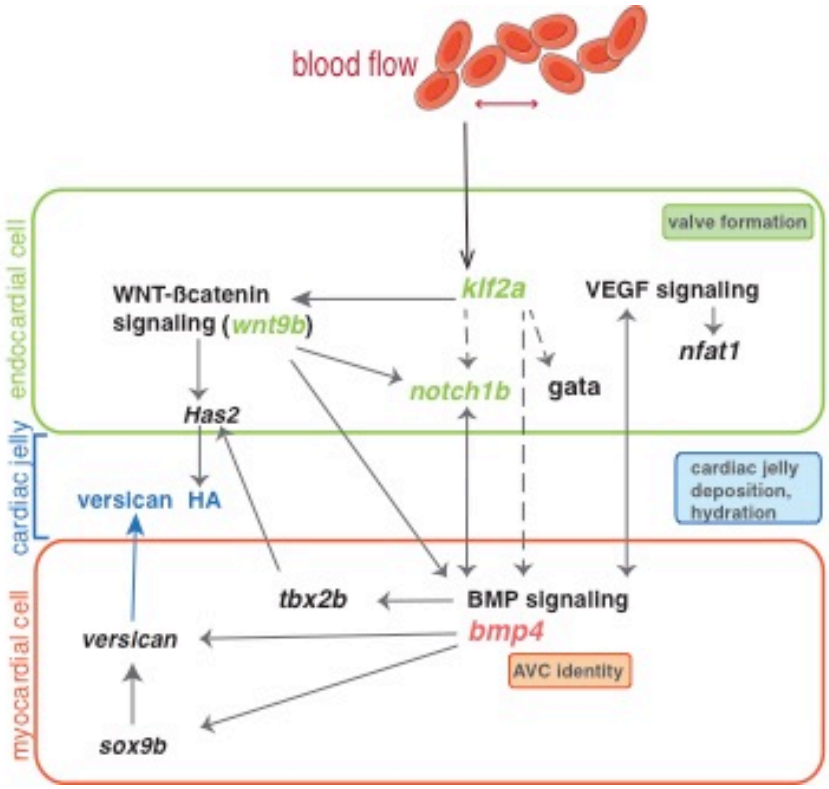
V.4 Genetic pathway regulated by Klf2/*klf2a*

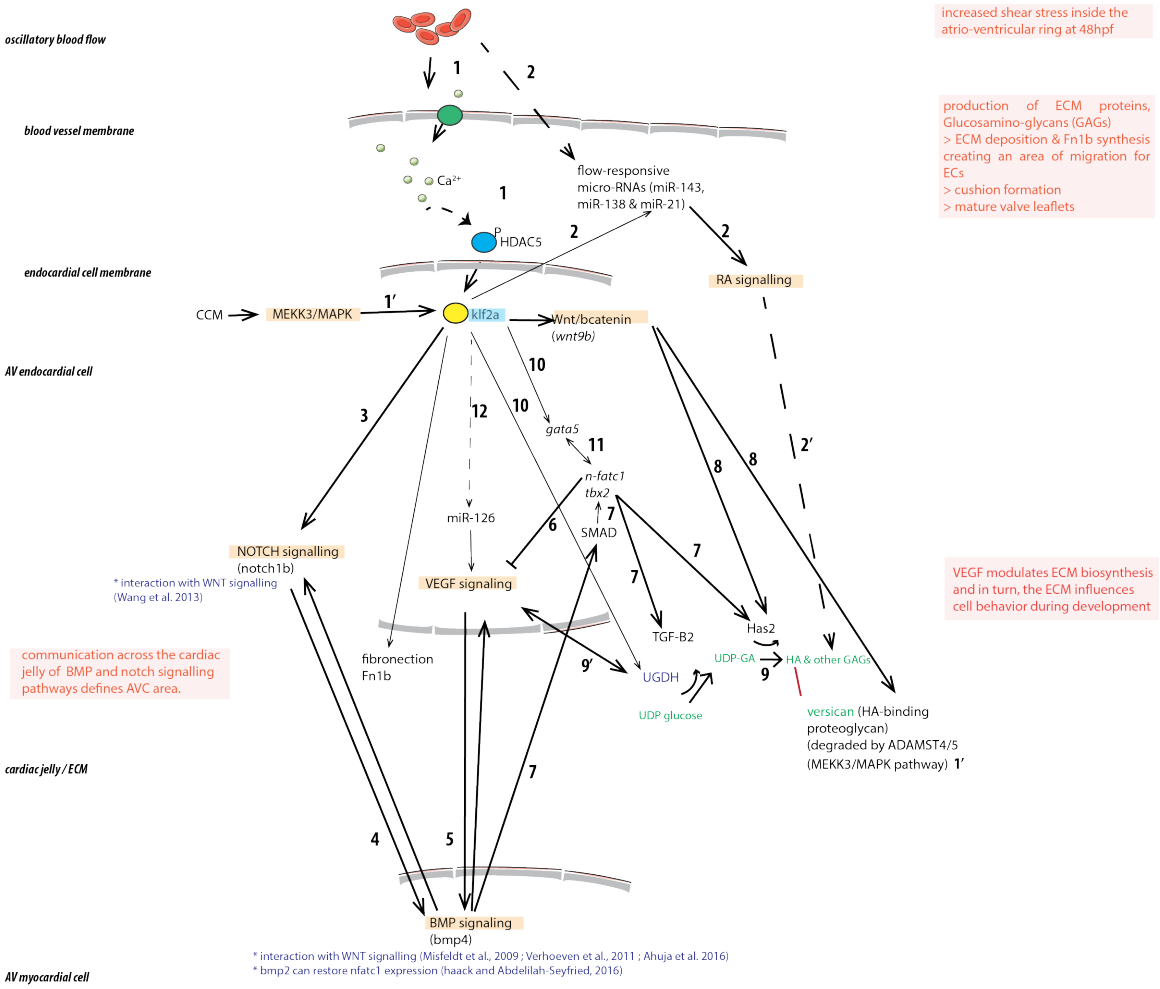
KLF2/Klf2a by transducing fluid shear forces was shown to regulate, directly or indirectly, endothelial gene expression (Chiplunkar et al., 2013) and consequently control EMT and atrial septation, through the activation of multiple genes involved in cardiovascular development: *gata4, tbx5* and *ugdh,* and indirect regulation of *sox9* (Dekker et al., 2006, Chiplunkar et al., 2013). *Klf2a* has also been shown to directly affect heart chamber growth and morphogenesis in the endocardium (Dietrich et al., 2014).

The molecular basis of KLF2/Klf2a diversity of functions and its direct transcriptional targets are still partially unknown. Steed et al. discovered *fibronection 1b (fn1b)* as a downstream gene target of Klf2a in developing AVC. Enriched *klf2a* expression on the ventricular side of the AVC at 48Hpf was shown to subsequently cause an enrichment of the expression of the gene *fn1b* to the same cells (Steed et al., 2016). Klf2a/klf2b were shown to controls Notch activity and Yap1 expression to coordinate OFT valve morphogenesis (Duchemin et al., 2018 in press)

It was also shown that within developing aortic arch blood vessels, Klf2a activates a signalling cascade involving the endothelial-specific microRNA miR-126a that promotes VEGF-induced angiogenesis (Nicoli et al., 2010). Blood flow is also essential for Hematopoietic Stem Cell (HSC) development and is mediated by a klf2a-Nitrogen Oxyde (NO) signalling cascade; Klf2a was shown to directly regulate NO synthase (nos) genes (Wang et al., 2011).


OFT
AVC
V
A
endocardial
cushion
48hpf


blood flow
valve formation
endocardial cell
WNT-ßcatenin
signaling (wnt9b)
klf2a
VEGF signaling
nfat1
notch1b
gata
Has2
cardiac jelly
versican
HA
cardiac jelly
deposition,
hydration
tbx2b
BMP signaling
bmp4
myocardial cell
versican
AVC identity
sox9b



1: PrKD2 phosphorylates of HDAC5 and nuclear export (Wang et al. 2010) / Activation of a calcium-activated intracellular cascade via blow-flood sensitive channels Trpp2 and Trpv4 (Heckel et al., 2010) which leads to the expression of *klf2a* gene.
1' : MEKK3/MAPK signaling also controls *klf2a and ADAMST proteases* expression via CCM complex regulation (Zhou et al., 2015)

2: Response to flow pattern is also mediated by small RNAs, often referred to as mechano-miRs (Kumar et al., 2014). miR-21 (banjo et al. 2002), miR-143 which targets retinoic acid (RA) signalling pathway components. Klf2 binds to promoter of miR-143 (Hergenreider et al., 2012).
2': cregulation by RA sigalling of GAGs synthesis

3: activation of *notch1b* gene by Klf2a (Vermot et al., 2009, Duchemin et al., 2018 in press)
4: crosstalk between notch and bmp factors (Vermot et al., 2009 ; Garside et al. 2013)
5: notch1b and bmp4 is abolished when VEGF-R signalling was chemically disrupted (Lee et al., 2006).

6: calcineurin/NFAT represses VEGF expression from 20–33 hpf (Lee et al., 2006).
7: BMP activates cytoplasmic SMAD proteins that regulate gene expression in the endocardium of tbx2/tbx2b gene. This gene directly bound Has2 and Tgf-β2 promoters and increased their transcriptional activities required for EMT / cushions formation (Beets et al., 2016)

8: Wnt/b-catenin signalling is operative in myocardial and endocardial cells only at AVC and regulates bmp4, has2 and versican (Hurlstone et al., 2003). wnt9b was shown as a target of Klf2a (Goddard et al., 2017)

9: Has 2 synthetizes HA, hyaluronic acid, which was shown to be important for cell migration into the cardiac jelly and for mice EMT in combination with Ras signalling and also involving CD44, and ErbB receptors/ligands (Camenisch et al., 2000). Some GAGs synthetize by HA interacts with versican protein (Hurlstone et al., 2003).
9': UGDH cooperates with VEGF signaling (Walsh et al., 2001)

10: KLF2 binds to and positively regulates the UDP-glucose dehydrogenase (Ugdh) gene, the Tbx5 and Gata4/5 genes in the mouse E10.5 AV region and indirect regulation of Sox9 (Chiplunkar et al., 2013)
11: cooperative interaction between GATA5 and NF-ATc regulates endothelial-endocardial differenciation of cardiogenic cells. (Nemer and Nemer, 2002)

12: Additional functions of Klf2/klf2a reported in angiogenesis, but to date, not proved in heart development. It was also shown that within developing aortic arch blood vessels, Klf2a activates a signalling cascade involving the endothelial-specific microRNA miR-126a that promotes VEGF-induced angiogenesis (Nicoli et al., 2010).

Figure 9: General genetic and molecular pathways leading to the formation of AVC valve leaflets.

Schematic representation of the mechanisms of activation of KLF2/klf2a and of the signaling pathways activated by this transcription factor.

V.5 Role of *klf2b*, paralog of *klf2a*

As mentioned before, in zebrafish, *klf2a* has a paralog gene called *klf2b*. The precise functions of this second gene are not yet completely known and could be redundant with *klf2a* own functions (Kotkamp et al., 2014). It is not known as well is *klf2b* is a hemodynamic-sensitive gene like *klf2a*. These specific questions will be addressed in this thesis.

Note about *klf4* and *klf17* genes:

Klf2a and *klf2b* form with *klf4* (former *klf4a*) and *klf17 (*former *klf4b* or *biklf) genes* a subfamily inside the KLF members - as mentioned in the phylogenic trees reviewed by (Oates et al., 2001) and (Xue et al., 2015) and illustrated in Figure 11.

Humans *Klf4* is also up-regulated by FSS in ECs *in vitro* and regulated by the same shear-responsive MEKK3-ERK5 pathway as KLF2 in endocardial cells *in vivo* (Cullere et al., 2015); (Zhou et al., 2016). Klf2 and Klf4 were shown to be both required in the endocardium of the developing mouse heart during the remodelling of cardiac cushions to mature heart valves (McCormick et al., 2001) ; (Zhou et al., 2015). *Klf17, klf2a* and *klf2b* function in a regulatory network to promote zebrafish periderm differentiation, acting downstream of both *Pou5f1* and an additional transcription factor speculated to be *Irf6* (Kotkamp et al., 2014) (Liu et al., 2016). *Klf4* was shown to be activated in the same manner as *klf2a* by HDAC5 pathway (Just et al., 2011).

The putative cardiac expression of these genes between 30h and 72hpf will be presented in the result part.

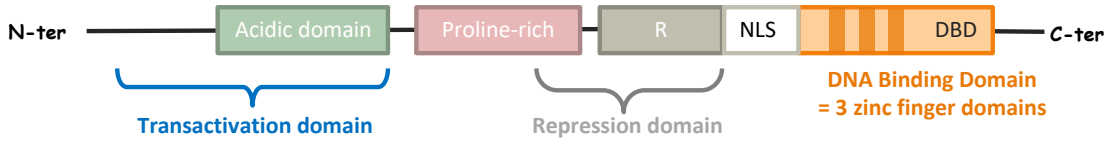


Figure 10: zebrafish Kf2a protein domains.

Klf2a contains an cidic domain (D20 to T37), a proline-rich domain (P75 to R106), a conserved block (F218 to D246) immediately N-terminal to a basic domain (K254 to R266).
Elaborated from data reviewed in (Oates et al., 2001) ; (Kaczynski et al., 2003) and (Novodvorsky et al., 2015)

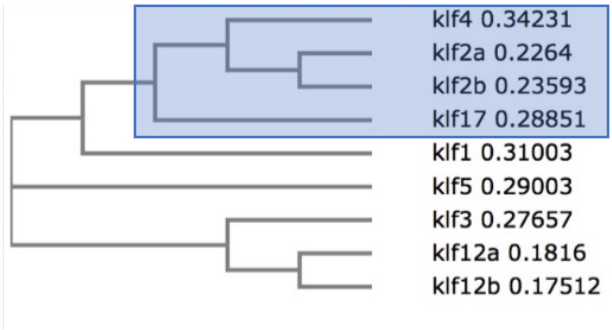


Figure 11: Phylogram of KLF zebrafish proteins.

An alignment of KLF protein sequences was realized with Clusal W program and a phylogram was generated (branch length: cladogram representation).

Figure 12: zebrafish Egr1 protein domains.

Adapted from (Burmeister and Fernald 2005)

Not surprisingly, other genes have been postulated to be flow-dependently activated and involved in valvulogenesis, in particular due to their restrictive area of expression during AVC formation. Amongst them the early growth factor 1 (egr1) is an interesting candidate.
Through In Situ Hybridization experiments, it was reported to be expressed in the heart from 30hpf and more restrictively in the AVC at 48hpf. Egr1 was also shown to respond to the flow in developing zebrafish AVC (Banjo et al., 2013)

VI. *Egr1*, another important gene for valvulogenesis

VI.1 Egr1, a member of the transcription factor "Early Growth Factor" family

Early Growth Factor 1 (*egr1*), also known as *NGFI*-A, *zif268*, *TIS8* or *krox-24*, is part of the transcription factor family of immediate early genes. They are named due to their rapid inducibility. They are master regulators controlling the expression of a wide variety of genes involved in cell growth, development and stress responses in many tissues. *Egr1* was originally identified in bones and cartilage as a rapid early stress-activated gene responding to a variety of proliferation stimuli, including NGF, (Sukhatme et al., 1998) ; FGF and PDGF (Christy et al., 1988) and general serum proteins (Lemaire et al., 1988), and mechanical stimuli in tendon formation and healing (Gaut et al., 2016)

EGR proteins are in a phylogenetic relationship with the 12 human KLFs, they are called "Krüppel-like related-proteins" (Dang et al., 2000). As KLF, EGR members are zinc-finger transcription factors, containing three DNA binding domains of the Cys2-His2 class (Sukhatme et al., 1988; (McMahon et al., 1990) (Figure 12). Egr1 binds to a consensus recognition element, conserved between human, mice and zebrafish: GCGGGGGCG (Drummond et al., 1994). Once bound to DNA, Egr1 can act as either an activator or a repressor of transcription, through mechanisms that depend on interactions with or regulation of distinct cofactors (SP1, NAB1/2) (reviewed in (Thiel and Cibelli 2002) but are to date not completely discovered.

The high degree of amino acid identity between the zebrafish and human *egr1* genes, the presence of well-characterized vertebrate regulatory cis-elements in the zebrafish promoter and conservation of the functional domains indicate that this gene has been highly

conserved during vertebrate evolution, implicating egr1 as an important regulator of growth and development (Drummond et al., 1994) and (Burmeister and Fernald 2005).

VI.2 *Egr1* expression and functions

Egr-1 is a central actor mediating the transcription of genes involved in a multitude of signaling cascades vital for growth, differentiation and apoptosis. (Thiel and Cibelli 2002) reported through the observation that Egr1 biosynthesis is strongly stimulated by activation of the mitogen-activated protein (MAP), a role for this factor in controlling cell proliferation and growth factor synthesis. Egr1 was postulated as a ''pro-apoptotic protein'' (Liu et al., 1998) and to be an active part of the apoptotic-signaling cascade. It was also shown to be associated with tissue injury, being an important mediator of fibroblast activation activated in tissue repair mechanisms, and in general playing a central role in regulation of fibrogenesis. (Bhattacharyya et al., 2013). Egr1 is important for brain and ocular development (Hu et al., 2006).

Egr1 has been linked to several aspects of cardiovascular pathology including intimal thickening following acute vascular injury (Khachigian 2006), cardiac hypertrophy (Rayner et al., 2013) atherosclerosis (Harja et al., 2004) and angiogenesis (Fahmy et al., 2003). Egr1 was shown to be involved in human valvulopathies. It was indeed detected in calcific human aortic valve cusps compared with non-calcified normal cusps (Ghazvini-Boroujerdi et al., 2004).

In zebrafish, expression of *egr1* starts at 4-somite stage (10hpf) with a strong expression in posterior adaxial cells, and from 40hpf a weak expression in the heart can be detected, the signal is maintained at 48hpf (Close et al., 2002) and (Banjo et al., 2013). *Egr1* is important in early zebrafish retinogenesis (Hu et al., 2006) and (Zhang et al., 2013). In human, *Egr1* has a role in angiogenesis and was shown to be activated by acute mechanical injury and other vascular stresses (Khachigian 2006). Khachigian demonstrated that *Egr1* – as *Klf2a* – is a flow-responsive gene: it is activated in ECs exposed to Fluid Shear Stress (Khachigian et al., 1997). In my thesis, I studied *egr1* functions during valvulogenesis and addressed if its expression is mechanosensitive. As binding-sites for *egr1* were found in mice KLF2 (Huddelson et al., 2004) suggesting a first crosstalk regulation of both genes, I also explored this specific question. Finally, I analysed the gene targets of *egr1* in the developing endocardium.

.

Objectives of this thesis

As suggested in the introduction, the present thesis aims at studying AVC valve formation in zebrafish through the comprehension of the roles of three transcription factors: klf2a, klf2b and egr1. We worked on the understanding of their specific roles during valvulogenesis, to complete the knowledge regarding their activation by blood flow and to decipher the target genes they are regulating.

1) The first aim was to validate the loss of Klf2a protein in a *klf2a* Knock-Out mutant lines generated in the lab, to be certain of the total disappearance of the protein, validating the mutant line status. For this purpose, polyclonal and monoclonal antibodies against Klf2a, klf2b and Egr1 proteins were generated. The homozygous state of mutant lines was assayed though Western blot assays.

2) Then the cardiac phenotype of these lines (and the double mutant lines obtained by crossing the single mutant fish) was characterized. We observed defects in valve formation and analysed different gene expression patterns.

3) A third part aimed at understanding the genetic pathway in which these three genes are involved, and at deciphering in particular which are their downstream gene targets. For this purpose, different omics-approaches were developed: ATAC-seq, mRNA-seq and Chromatin ImmunoPrecipitation (CHiP).

Results

Chapter 1

Generation of zebrafish polyclonal and monoclonal antibodies to validate Knock-Out mutant lines

The first part of my thesis aimed to generate specific antibodies to confirm the disappearance of proteins in the single KO-mutant *klf2a*^{ig4}*, klf2b*^{ig5} *(and the double mutant klf2a-klf2b)* and *egr1*^{s64} lines. Single mutant *klf2a*^{ig4} *and klf2b*^{ig5} were generated by our laboratory using TALEN technology (*see Materials & Methods for further explanation about their creation*). *klf2a*^{ig4} line was published in (Steed et al., 2016). *Egr1*^{sa64} was received as a heterozygous line from ZIRC (*see Materials & Methods for information about the generation of a stable line*). Figure 13 presents the localization of each mutation and general data about the endogenous proteins and information about the native proteins is presented in table 1.

All these lines should constitute genetic-KO mutant. Each respective mutation leads to a premature stop codon early in the coding sequence resulting in theory to a non-produced/non-encoded protein. The mRNA is nevertheless still produced as it can be detected by In Situ Hybridization assay with a probe complementary to them (*data not shown here but realized on egr1*^{sa64} *and klf2a*^{ig4} *mutant samples*).

To confirm the absence of gene product (here the protein) in these mutants, some polyclonal and monoclonal antibodies against the proteins of interest were generated. Their efficiency was determined on a Western Blot (WB) assay. Some of them were also used on an Immuno-Fluorescence assay (IF) to localize the proteins.

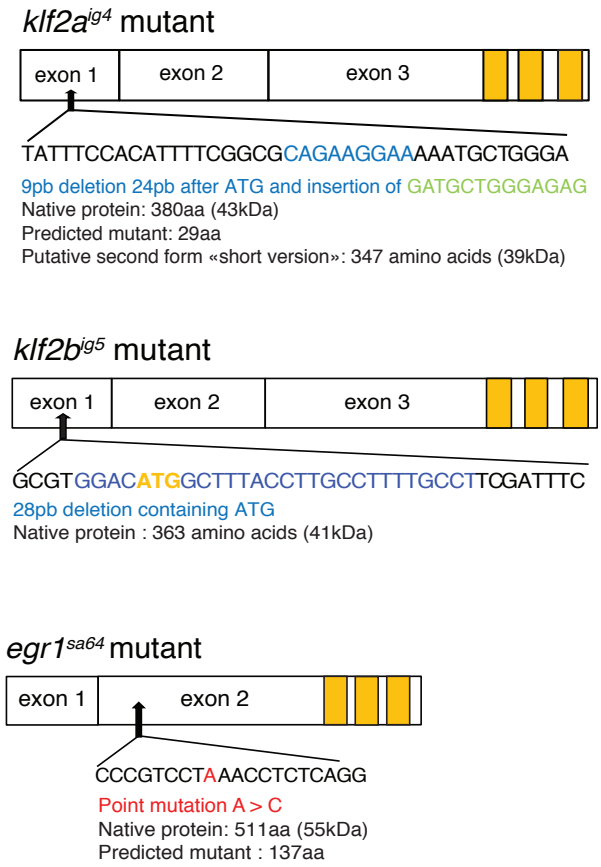


Figure 13: Presentation of the zebrafish KO-mutant lines for *egr1, klf2a* and *klf2b* genes used in this thesis.

The mutation is indicated on the genomic representation of each gene.

klf2a^{ig4} and *klf2b*^{ig5} were obtained using TALEN endonuclease system (*see Materials and Methods for further details*) in the lab. *Egr1*^{sa64} is a line available on ZIRC database, the point mutation resulted from ENU treatment.

Information about the molecular weight of each respective zebrafish protein.

Table 1: Native MW and Uniprot references of proteins of interest studied in this manuscript

Protein name	Molecular Weight	Uniprot reference
Klf2a LV	42,88kDa	Q1LXI7
Klf2a SV	39kDa	Q6P3M5
Klf2b	41,2 kDa	Q90XE7
egr1	55,1 kDa	P26632
Klf4 (previous name Klf4a)	43,6kDa	A9X6Q5
Klf17 (previous names Klf4b, biklf)	46,9kDa	Q9DFS2

I. Preamble about the optimization of zebrafish protein lysis and Western Blot assay

I.1. Experimental procedure to get zebrafish protein lysates

In order to get enough material of good quality and to obtain nice bands on immunoblots, I optimized a protein extraction protocol. The method used previously in the lab was not reproducible in my hands and did not bring satisfying results.

I.1.a) Choice of the extraction lysis buffer

To maximize the efficiency of the protein extraction, in particular by preventing sample degradation, protein aggregation and to enhance their solubility, I first tried and compared various extraction buffers commonly used in laboratories worldwide (the list is available in Table 2). The common approach in the zebrafish field involves the use of detergents like Sodium dodecyl sulfate (SDS) to solubilize the proteins before the separation in SDS-PAGE gel. It is one of the most common and helpful surfactants for protein extraction (Wang J. et al., 2014). Urea, a chaotropic agent[7], is also commonly added to lysis buffer, it competes with the protein native interactions, resulting in a better unfolding and solubilization. Sodium deoxycholate (DOC), a bile salt surfactant, is more used for the purification of membrane proteins (Wang J. et al., 2014). It is contained into the common protein lysis buffer (RIPA buffer) with a small percentage of SDS. β-mercapto-ethanol could be used to replace DTT. Although it has a higher volatility and toxicity than this latter, it was shown to be less prone to oxidation and less unstable. Unlike DTT, it drives the reaction forward to completion (from: VP-DSC MicroCalorimeter User's Manual, MicroCal, LLC).

[7] A chaotropic molecule helps disrupting the hydrogen bonds between water molecules thus impairing the stability of molecules (like proteins) in the solution, by weakening the hydrophobic effect.

Table 2: Overview of the different lysis buffers tested to optimize the protein lysis of 48hpf-zebrafish embryos

Protocol	Buffer components
1) RIPA lysis - Lysis performed 15 min in RIPA lysis buffer on ice - Addition of 1X SDS buffer - heat-denaturing step (5 min 95°C)	RIPA 50mM Tris-HCl pH7.5 150mM NaCl 0.25% Sodium deoxycholate 1% NP40 add immediately before use: protease inhibitor cocktail 6X SDS blue buffer 250mM tris-HCl 10% SDS 0.02% bromophenol blue 30% glycerol pH6.8 add immediately before use: 500mM DTT
2) Urea lysis same protocol as for buffer 1)	add 6M urea inside RIPA buffer
3) β-mercapto-ethanol lysis *(Dr Samantha Carrillos-Rosas from Dr Yvon Trottier's lab)* - Homogenize with a microfuge pestle until the mix can be pipetting easily, add the buffer in 2 times and boil 5 min between each -heat-denaturing step (5 min 95°C)	63mM 1M Tris-HCl, pH 6.8 0.1% glycerol 0.05 % β-mercapto-ethanol 1% SDS
4) Laemmli lysis buffer *(Following the advice of Dr Benjamin Vitre, CRBM – CNRS, Montpellier)* - Homogenize until the mix can be pipetting easily. - heat-denaturing step (5 min 95°C)	0,25M Tris pH 6,8 40% glycerol 4% SDS 10% β-mercapto-ethanol 0,3% bromophenol blue

I.1.b) Sample preparation: embryos dechorionating and deyolking steps

At 48hpf, not all embryos have hatched. Thus, their eggshell, the chorion, should be removed. Dechorionation could be done by enzymatic removal using Pronase enzyme (Roche), a manual operation was nevertheless preferred using a pair of forceps. This greatly increased the extraction by preventing the damaging of proteins.

48hpf-embryos should also be deyolked. Indeed, the large amount of lipids contained into the yolk will distort band pattern migration and may mask proteins of interest (Yuan et al., 2003).

Another important point to consider is the protein conservation. It should be possible to keep the denatured lysis extracts at -20 or -80°C few days after fast freezing. However, a degradation of the samples (i.e. transcription factors of interest) was observed already two days after storage. To keep reproducible data, samples should be prepared fresh and proteins run on Western Blot on the same day.

I.2. Western Blot optimization

Other parameters also require special attention and can be optimized to get a good resolution on the immunoblots.

I.2.a) The quantity of proteins loaded on gels

Loading too much protein could lead to signal saturation in western blots, yet too little produces weak signals. I compared three methods to measure protein concentrations in order to find a reliable method that gave me an accurate estimation of the quantity I load on gels.

BCA protein assay: it corresponds to a Copper-based protein assay based on the traditional Lowry assay, except bicinchoninic acid is used. Peptides containing three or more amino acid residues react with cupric ions (Cu2+) to form a colored chelate complex. A product of the reaction from an excess of urea, biuret, reacts with copper to form a light blue to purple tetradentate complex that absorbs at 560 nm. The intensity of the color produced is proportional to the number of peptide bonds participating in the reaction. BCA has a broad dynamic range – capable of measuring protein concentrations of 0.5µg/mL to 1.5mg/mL. Unlike other methods available, the BCA Protein Assay is compatible with samples that

contain up to 5% surfactants (detergents). In addition, the BCA Assay was shown to respond more uniformly to different proteins than the Bradford method *(thermo fisher scientific data)*

Coomassie dye (Bradford) protein assay: Coomassie Brilliant Blue (or G-250[8]) interacts electrostatically but non-covalently with the amino and carboxyl groups of proteins. The formation of the dye-protein complex stabilizes the negatively charged anionic form of the dye producing the blue colour. (Chial et al. 1993). This results in a spectral shift from the reddish/brown form of the dye (absorbance maximum at 465 nm) to the blue form of the dye (absorbance maximum at 610 nm). The difference between the two forms of the dye is greatest at 595 nm, optimal wavelength chosen to make the measurement. It is less sensible than BCA to detergents present in the lysis buffer.

NanoDrop protein A280 measurement: this device measures the absorbance of mainly tryptophan and phenylalanine amino acids, and estimates to the average frequency of these amino acids inside the protein mix. Consequently, the given measurement of protein concentration is, usually off from Bradford or BCA measurements by a factor of two or more. It is very sensitive to interference from a number of compounds, including detergents.

I.2.b) The blocking solution

Nonfat-dried milk or BSA can be diluted in PBS-Tween 0.1% as blocking solutions. BSA is generally used with biotin and antiphosphoprotein antibodies. It is one protein, so less for the antibody to cross-react with compared to milk, which contains in particular casein, a phosphoprotein and biotin, thus interfering with the assay results (Mahmood and Yang, 2012). However, BSA is more expensive and it was used in this thesis only when no band could be detected on membrane for some antibodies after optimization.

[8] "G" stands for the blue colour of the solution, a "greenish" tint. The "250" originally denoted the purity of the dye.

I.2.c) The concentration of primary antibodies

Different concentrations were tested to be able to detect a band on Western Blot. The figures of this manuscript present the immunoblot with the optimal concentration.

I.2.d) The time of exposure

Each immunoblot was exposed during various times until saturation of the signal, using Clarity System for the revealing solution.

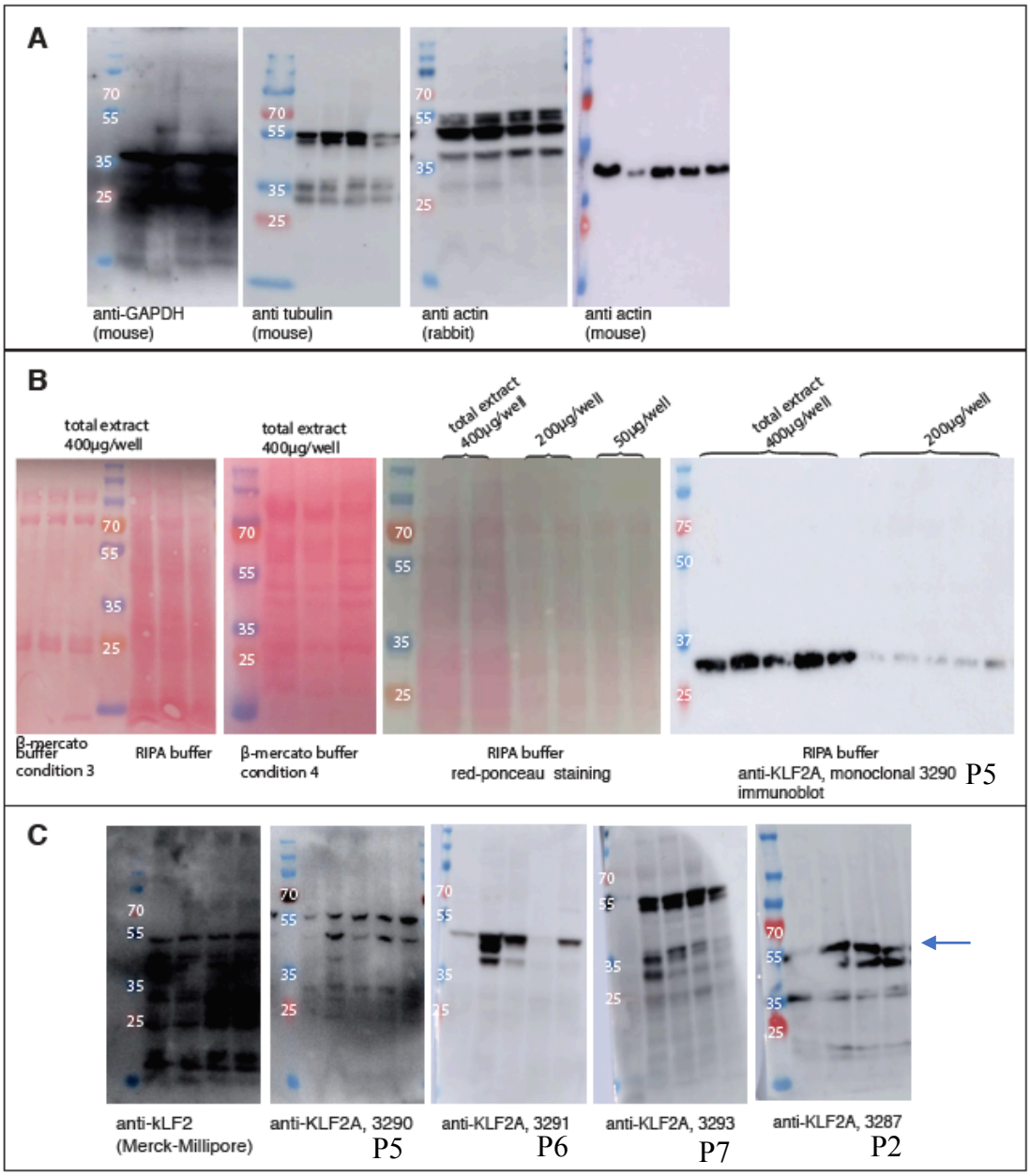


Figure 14: Optimization of a Western Blot protocol for protein lysates of 48hpf-embryos.

WT-protein lysate replicates were loaded on a 12%-polyacrylamide gel. Nitrocellulose membrane was incubated with antibodies in 5%milk-PBST solution (concentration 1/1000).
Secondary antibody: goat anti-rabbit HRP (ref: 111-035-003, Jackson ImmunoResearch Lab.), goat anti-mouse HRP (ref: 115-035-003, Jackson ImmunoResearch Lab.)
Revelation using Clarity System. Ladder: Pageruler plus Prestained protein 26619 from Thermo-Fisher.

A: Test of different control antibodies: anti-mouse Beta tubulin (ref: TUB-2A2, IGBMC), mouse anti-alpha actin (ref: ACT-2D7, IGBMC), rabbit anti-actin (ACT, IGBMC)
B: Test of different lysis buffers to extract protein from 48hpf-deyolked and dechorinated embryos. Buffer compositions are listed in Table 2.
C: Test of anti-KLF2/Klf2a antibodies: anti-KLF2 (commercial Merck-Millipore) ; polyclonal anti-Klf2a generated and purified at IGBMC.

I.3. Results and discussion

The final optimized protocol is available in the materials and methods part. It summarizes the best parameters which avoid sample degradation, protein aggregation and allow to perform clean and reproducible immunoblots in my hands with 48hpf-embryo protein lysates.

I.3.a) The choice of the lysis buffer

When embryos were lysed into RIPA buffer, following a protocol used previously in the lab, I never managed to obtain reproducible data. Most of the time, proteins stacked in the gel wells, certainly resulting from aggregation and degradation of proteins. In addition, the proteins of interest are not detected at their expected size - as illustrated on the immunoblot in Figure 14). Klf2a is a 43kDa-protein and a band around 30kDa was detected. A slight discrepancy between the theoretical MW and the MW observed during SDS-PAGE is however not uncommon. It may happen that a protein does not run on a gel at its expected size. It can be the result of protein modification (glycosylation, phosphorylation…), in this case the detected band would correspond to a higher molecular weight (Guan et al., 2015). They may either be not completely denatured or equally likely do not bind SDS micelles well due to its amino acid composition (Rath et al., 2009), resulting from a faster migration on the gel. It could also come from protein degradation. As with the other lysis protocols, a band corresponding more to the theoretical size can be detected, we assumed it was indeed due to protein degradation with RIPA buffer.

Extraction in RIPA buffer supplemented with 6M-urea was slightly better than in RIPA alone, as evident from the larger number of protein bands as well as higher intensity bands which can be seen on red Ponceau stained membranes (*data not presented here*). However, correct gel migration and visually correct bands on immunoblots were obtained with an extraction based on ß-mercapto-ethanol (buffers conditions 3 and 4, cf table 2). However, with the buffer 3 from time to time, really sticky sample solutions were obtained, not easy to pipet. Finally, only the buffer 4 was used, which contained a higher concentration of SDS, for the following protein extractions.

Additionally, the results presented here were based on whole embryo extract rather than nuclei proteins. A special lysis protocol for keeping the nucleic protein fraction could have also been interesting to try, especially when working with transcription factors. Nevertheless, as we managed to get good visible bands of proteins on immunoblots with the current protocol.

I.3.b) The quantity of loading material

BCA and Coomassie methods gave the same range of protein concentration measurements. As expected, with a Nanodrop the values are biased by a factor of two. However, Nanodrop method was preferred when a very accurate estimation of the protein concentration is not required. This method is fast, cheap and requires only one microliter per sample lysate. The range of concentrations of the lysates was thus determined once for all with BCA or Coomassie, and the appropriate amount to load on gels was selected. Then for further extractions, a Nanodrop measurement was performed in order to normalize the lysates before loading. Different amounts of whole protein extracts were tested on gel: 400, 200 and 50ug per gel well. As illustrated on Figure 14, panel B, the best resolution corresponds to 400ug of total proteins loaded per well, in order to detect the bands corresponding to the transcription factors of interest. To obtain enough quantity to perform up to 4 Western blot the same day, 50 embryos are required.

I.3.c) Finding a good control protein

The use of an antibody against a “control protein” ensures the same amount of protein sample was loaded in each lane, they transferred well in the nitrocellulose membrane and the signal detection are uniform across different lanes after revelation. They are thus typically used to normalize the signals from the proteins of interest.

Some troubleshooting was encountered to find a good control antibody. Beta-Actin, beta-Tubulin and GAPDH proteins present high homology between corresponding proteins from species as diverse as human, mouse, zebrafish and chicken. For that, antibodies against them are commonly used as controls for Western Blot. However, with anti-GAPDH and anti-Beta-tubulin unexpected bands were obtained (Figure 14), which were not due to non-specific binding caused by a too high concentration of primary or secondary antibodies. It may be due to a cross-reactivity with similar epitopes on other proteins. Only anti-beta-actin from mouse

gave satisfying results. As Beta-actin is also present in the nucleus, as a component of chromatin remodeling complexes (Olave et al., 2002), it is not recommended in theory to be used as a control for nuclear protein samples. Nevertheless, after some trials, no change in protein expression was observed for control versus mutant samples used in this study. As an alternative to antibody control, several authors have proposed to forego the use of a protein as a loading control, and to rely on the dye staining of proteins before (by Coomassie blue) or after (by Ponceau) the transfer step during Western blotting (Romero-Calvo et al., 2010). They do not fulfill the advantages of "antibodies" as an ideal control, as they are controls for sample loading and protein transfer but not for antibody incubation/signal detection. However, in practice, for some samples, they are necessary and a combination of both when possible if presented on the figures of this manuscript.

I.3.d) The time of exposure

Depending on the time of exposure some bands may be masked. Immunoblots were then always exposed during different times to get to the over-exposure and be able to detect residual bands.

This optimized method was applied to the validation of Knock-Out mutant lines by searching for protein disappearance. This aim required first to find good antibodies specific for the protein of interest.

II. The need of home-made anti-Klf2a/Klf2b antibodies

A commercial mouse anti-KLF2 antibody is available on the market (Merck-Millipore, reference 09-82). (Wang et al., 2011) used it in Western blot assays to show that Klf2a protein was significantly down-regulated in Klf2a-morpholino (MO)-injected embryos and to confirm *nos* genes as downstream targets of *klf2a* on Chromatin Immuno-Precipitation Assay (CHiP).

While testing this antibody on Western Blot, lots of bands can be detected, as shown on Figure 14, Panel A. This antibody does not seem specific to KLF2/Klf2a. According to the manufacturer, the epitope used to design the antibody targets the common DNA-binding domain region (*emails exchanged with Merck Millipore R&D department*). Thus, this antibody is not specific for one KLF member, in particular not only specific for KLF2/KLF2A. It could detect KLF2A, since mouse and zebrafish KLF2 proteins are highly conserved, but could also detect KLF2B and even KLF4, 17 and other KLF. Moreover, the antibody is in an un-purified serum. In order to detect specifically Klf2a and Klf2b, the production of "home-made" antibodies against these proteins was started.

Due to this lack of specific commercial zebrafish antibodies against Klf2a, Klf2b and Egr1 when the study was started, we decided to generate our own antibodies, starting with polyclonal antibodies.

II.1 Generation of polyclonal antibodies against Kf2a/Klf2b

II.1.a) Presentation of the chosen peptides

This project was launched by Dr Emily Steed and Dr Karim Hnia in 2012. They designed the peptides, which were synthetized by M. Pascal Eberling at the sequencing IGBMC platform. Peptides were injected to immunize rabbits and four bleed sera were collected. The sequences of the peptides are available in the Table 3, their localization in the protein sequence is highlighted in the alignment on Figure 17. As mentioned in the introduction, KLF protein sequences are highly similar, Klf2a and Klf2b form with Klf4 and Klf17 a sub-protein family inside the KLF members. For this purpose, the peptides were designed to be not found in the most homologous sequences. We can notice P5 is not completely specific to Klf2a and a partial redundancy in sequence with Klf2b is visible on Figure 17.

This peptide was designed to be specific to Klf2a long version, the published protein sequence available for Klf2a, since the second putative methionine (*see the introduction for further details*) corresponding to the short Klf2a version is located downstream. However, P5 sequence can also be found in Klf2b. The same can be said about P6 specificity for Klf2a.

II.1.b) Purification and test of the polyclonal antibodies

My PhD project started with the test of the polyclonal sera on Western Blot. Unpurified sera gave similar results to the one obtained with the commercial anti-KLF2 antibody. Lots of bands were detected on immunoblots (*data not presented here*)

The sera were then purified. The peptide corresponding to its related serum was bound to a protein A sepharose column. The complete protocol is detailed in the Materials and Methods part. Table 3 and Table 4 summarize the concentration of antibody obtained for each purified serum. Each purified polyclonal antibody was tested on Western Blot with 48hpf-whole embryo protein lysates. Two different concentrations (1/500 and 1/1000) were tested. Figure 14 presents the results obtained with the concentration 1/1000.

According to the sub-optimal results, despite a purification step, the production of monoclonal antibodies was undertaken, choosing carefully interesting epitopes.

II.2. Monoclonal antibodies against Klf2a, Klf2b, Klf4, Klf17 and Egr1

II.2.a) Monoclonal antibodies production

To generate some monoclonal antibodies against Klf2a, the most promising peptides used for polyclonal antibodies were selected (P2 and P7) and some new peptides against Klf2b, Egr1, Klf4 and Klf17 proteins were also designed. This was performed using Abdesigner website. This website does not contain links to the zebrafish protein bank. It is only possible to copy a fasta sequence of interest and then carefully select interesting peptides based on the Ig-score rank. Table 5 summarizes the designed sequences.

Hybridoma culture supernatants from the immunized mice (*see materials and methods for the protocol*), positive in Elisa test, were collected and pre-tested to select one or two good candidates in Western Blot. The selected hybridomas were established and ascites fluids collected to obtain the monoclonal. A purification step was performed before testing the antibodies on Western Blot assay.

II.2.b) Results and discussion

Production of a good anti-Klf2a antibody

On the immunoblots incubated with the 2KLF anti-klf2a antibody, no bands were detected in the samples “single klf2a mutant” and “double mutant klf2a-klf2b” (Figure 16, panel A’). This complete disappearance of Klf2a protein in mutants would be in favor of a real KO of *klf2a* gene in the mutant line ig4.

The putative second form of Klf2a is produced in fish

A band at a very low intensity can be detected in the “double mutant klf2a-klf2b” sample when the immunoblot - incubated with the 1KLF anti-klf2a antibody - was more exposed (Figure 13, panel A). This second short version could be recognized by this antibody (Figure 17). It should have a slightly lower molecular weight (39kDa vs 43kDa for the long version of Klf2a). As the 12%-protein gel is not resolutive enough, both forms cannot be discriminated according to their sizes with our assay. This result would be in favor of the production of a second putative form of Klf2a.

Does this form is functional in cells? Trying to answer this question, we generated in the lab two transgenic constructs allowing us to over-express both versions of Klf2a proteins: Tg(UAS:klf2a LV; cmlc2:eGFP) for the Long Version and Tg(UAS:klf2a SV; cmlc2:eGFP) for the Short Version (*See Materials and Methods for further information*).

The fish which over-express the long version, corresponding to the known functional version of Klf2a, they present some cardiac defects, some cardiac edema develop from 72hpf and they do not survive. No phenotypical defect was detected with the fish over-expressing the "short version". Moreover, when using these lines on WISH with anti-egr1 probe to detect the expression of egr1 mRNA, no difference in expression was observed with the fish expressing the "short version", but egr1 is up-regulated in the fish over-expressing Klf2a-long version (Figure 15 and *Figure 1 of the paper in submission, cf chapter 4*).

Are egr1^{sa64} and klf2b^{ig5} real KO-mutant lines?

Regarding the immunoblots against Egr1 protein, the commercial antibody (which was received after the production of the monoclonal antibodies) detected a band at the expected size (55kDa) but in both control and mutant samples (Figure 16, panel C). With the monoclonal antibody, surprisingly, it is not the same band which is detected (slightly bigger than 55kDa). The question of the specificity and cross-reactivity of this antibody is then raised. We did not get any answer about the epitope targeted by the commercial anti-egr1, it may target the common binding part of the protein and thus recognize other Egr proteins as it was described previously with the commercial anti-KLF2, or also other proteins of the same molecular weight.

For Klf2b, the antibodies 3KLF and 4KLF still cross-react with other proteins, since several bands were detected (Figure 16, panel D).

What does it mean regarding the mutant *egr1^{sa64}* and *klf2^{ig5}* lines used in our studies and the relevance of the data? To assay completely the KO-state of this line, it would be necessary to know what is exactly recognized by the generated antibodies. To determine this point, it could be possible to perform an immunoprecipitation assay and analyze by mass-spectrometry the immune-precipitates. Optimization could be performed with these two antibodies, and for example, other epitopes could be chosen to produce new monoclonal antibodies.

However, clear cardiac phenotypes were observed for all mutant lines, which indicate the mutation has an effect (*data presented in the following chapters*). Which is relevant even if the "KO"-state of the lines still remain to be completely validated.

Monoclonal against Klf17 and Klf4

A single band higher than 55kDa was detected for the immunoblots incubated with anti-KLf17 (5KLF), as shown on Figure 16, panel B. However, the expected size for this protein is 46,9kDa (Uniprot reference: Q9DFS2). It is possible this protein has some modifications (phosphorylation, glycosylation …) and may run at a slightly higher size. This have been elucidated yet.

Sera against Klf4 remain to be tested.

To conclude, we can notice that producing monoclonal antibodies was forth doing. Cleaner immunoblots were obtained. We managed to generate a specific antibody against Klf2a which detect, named 1KLF in this study. The characterization of Klf2a^{ig4} mutant was published in Steed et al. (2016) ; this paper is available in Annex 1.

Even if were not successful in generating efficient antibodies against Egr1 and Klf2b, a method was optimized to analyze 48hpf-zebrafish protein lysates and produce polyclonal and monoclonal antibodies against zebrafish proteins.

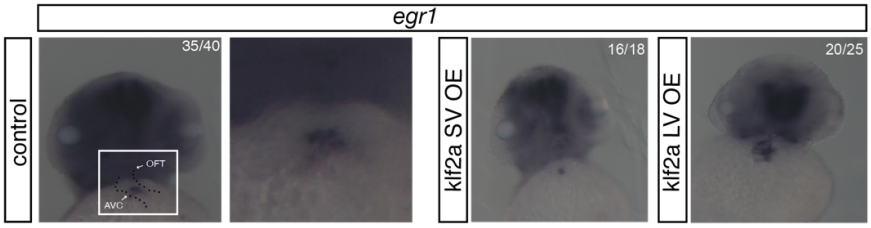


Figure 15: *egr1* WISH on the transgenic lines over-expressing *klf2a* in endocardial cells.

Fish carrying either the transgenic Tg(UAS:klf2a LV; cmlc2:eGFP) for the Long Version (LV) of Klf2a protein or the Tg(UAS:klf2a SV; cmlc2:eGFP) for the Short Version (SV) of the protein (corresponding to a putative second form of Klf2a) were crossed with Tg(fli:Gal4FF;UAS:kaede) fish. The fish over-expressing klf2a/GFP were selected under fluorescence at 48hpf (see Materials and Methods) and used for In Situ Hybridization with anti-egr1 probe.

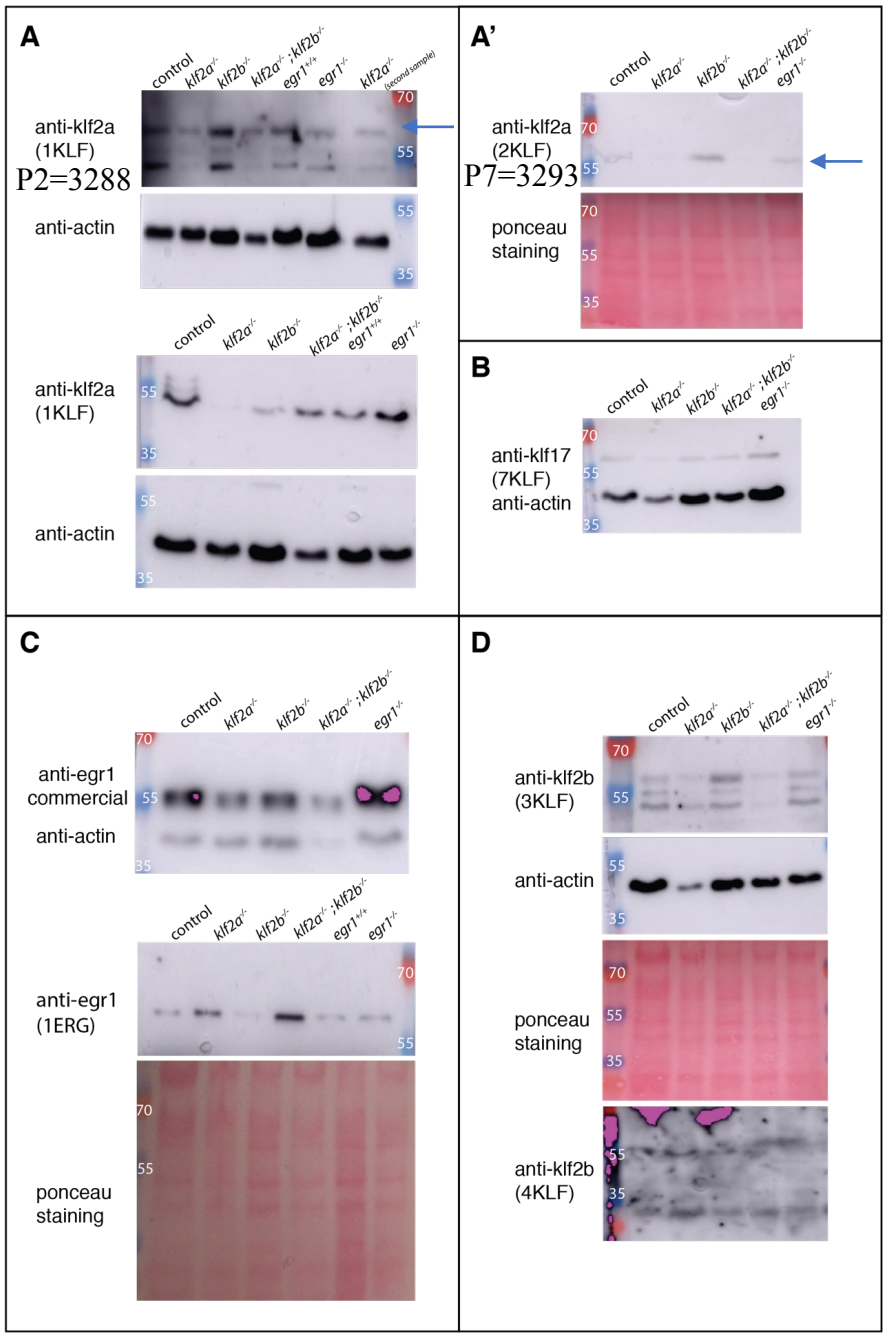


Figure 16: test of zebrafish monoclonal anti-Klf2a, anti-Klf2b, anti-Klf17 and anti-Egr1 antibodies, generated and purified at IGBMC.

Each primary antibody was incubated at 1/250 in 5%milk-PBST and 1/1000 for anti-actin control antibody. 50 embryos were lysated in β-mercapto-ethanol lysis buffer (condition 3) 400ug/well were loaded on gel.

For some blots, the anti-actin incubation was not realized and Ponceau staining is presented to control the global amount of protein loaded per gel.

Table 3: Sequences used to design peptide-directed antibodies

For some sequences, a cysteine residue (C highlighted in blue in the table) was added, required for the forward purification step. For the peptides with an asterisk (*), only the immunoblot for one of each couple are presented in Figure 14.

Peptide number	Sequence	Directed against	IGBMC Serum n°
P1	AELINTDCQSHSC	Kl2a	Not used
P2	SQSYQGNTGFPHC	Klf2a	3287 – 3288 *
P3	NVALNTDTCRGRKC	Klf2b	Not used
P4	HQFAMYEEAMGMQPSC	Klf2b	Not used
P5	SISTSAQKEKCWENC	Klf2a	3289 – 3290 *
P6	CMRSYEQPRLANS	klf2a	3291 – 3292 *
P7	KDELDRSMHLSC	klf2a	3293 – 3294 *
P8	IEEHKPLNASC	Klf2b	3295 – 3296

Table 4: Concentration and volumes of purified polyclonal antibody obtained after purification of the rabbit sera.

Concentration was determined by *Nanodrop* measurement.

Peptide n°	Serum n°	Concentration	Volume
P2	3288	2.02 mg/mL	1 mL
P5	3289	2.07 mg/mL	1 mL
P5	3290	1.66 mg/mL (03.12.2014) 1.67mg/ml (05.05.2015)	500 µL 2mL
P8	3295	3 mg/mL	2 mL
P8	3296	2.6 mg/mL	1.5 mL
P6	3291	3.92 mg/mL	2.5mL
P2	3287	0.60mg/mL	2mL
P7	3293	0.66mg/mL	1.5mL

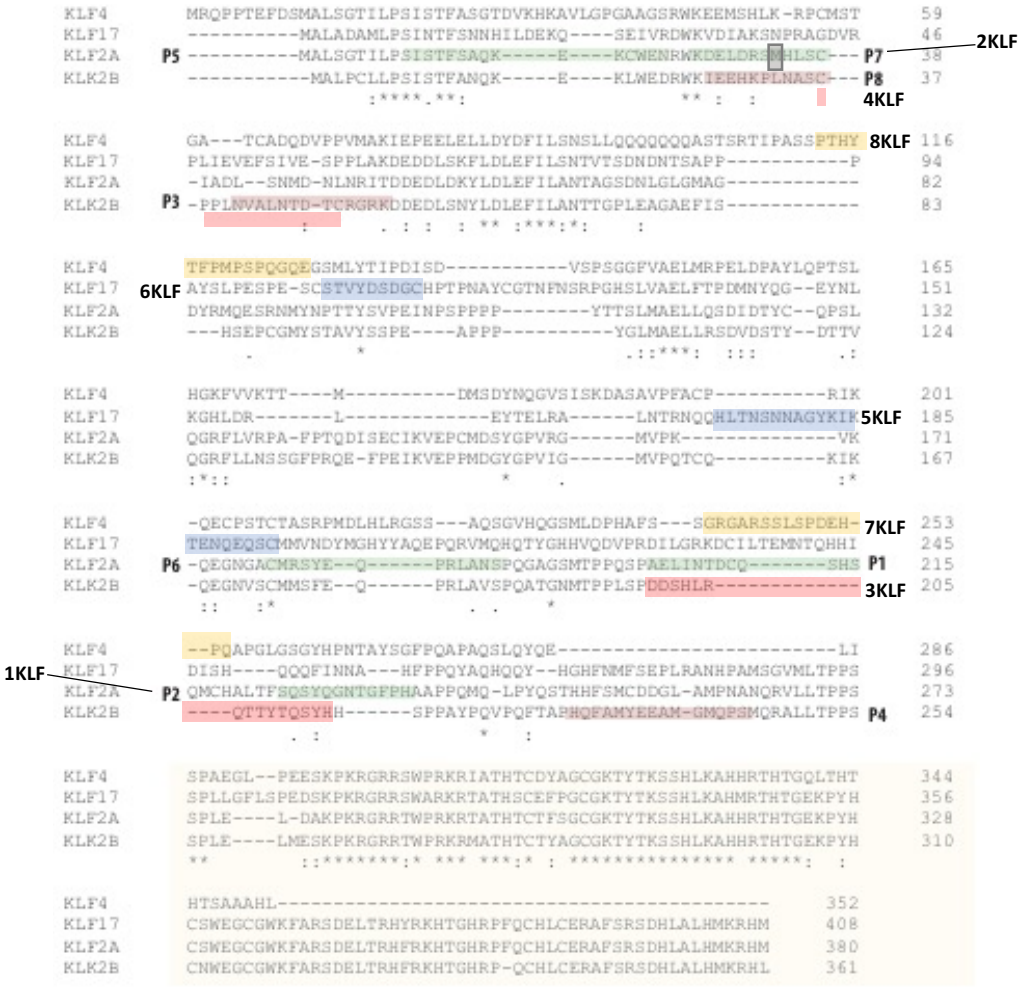


Figure 17: CLUSTAL O(1.2.4) multiple sequence alignment of Klf2a, Klf2b, Klf4 and Klf17 zebrafish proteins.

In green are highlighted peptides designed against Klf2a protein, in red those against Klf2b, in blue against Klf17 and in yellow against Klf4 (see Figure 16 and Table 5 for the correspondence peptides-names).

From the amino acid 344 starts the zinc-finger DNA-binding motif common to all KLF members. The second putative methionine (M) in Klf2a protein sequence is highlighted in grey.

Table 5: Peptidic sequences used to generate monoclonal antibodies

For some sequences, a cysteine residue (C highlighted in blue in the table) was added, required for the forward purification step.

Peptide n°	Polyclonal	Against	Sequence	Supernatant selected
P2 pg77LB1 PK270	serum 3287 (P2)	Klf2a	SQSYQGNTGFPHC	1KLF
P7 PK275	serum 3293 (P7)	Klf2a	KDELDRSMHLSC	2KLF
PM8	/	Klf2b	DDSHLRQTTYTQSYHC	3KLF
PM7	/	Klf2b	CPPLNVALNTDTC	4KLF
PM10	/	Egr1	SSSTYPSAKTC	1ERG
PM9	/	Egr1	IPSSTSQATHPSSSSTC	2ERG
PM11	/	Klf17	HLTNSNNAGYKIKTENQEQSC	5KLF
PM12	/	Klf17	STVYDSDGC	6KLF
PM13	/	Klf4	GRGARSSLSPDEHPQC	7KLF
PM14	/	Klf4	PTHYTFPMPSPQGQEC	8KLF

Chapter 2

Use of the home-made monoclonal antibodies to study Klf2b and Egr1 protein localization by Immuno-Fluorescence

It is possible that an antibody efficient to detect an epitope on a denaturated protein (on Western Blot assay) is not able to bind to native protein or inversely. That is why, we wanted to test some monoclonal antibodies in native conditions on fixed samples by Immunofluorescence. This requires as well optimization steps in order to ensure the penetration of the antibody inside our region of interest (the heart) and to choose the proper concentration of antibody to add to detect a visible signal.

I. Problematic and optimization steps

We were interested in the tissue and cellular localization of Egr1 and Klf2b inside the heart. We expected to find them inside endocardial cells of the valve forming area at 48hpf according to the preliminary results obtained with ISH *(data presented in the submitted paper).*

The principal conditions to optimize were the following:

- the time of permeabilization: from 1h at RT we switch to o/n 4°C under rocking
- the piercing of embryo in the neck near the heart: this should allow the antibodies to penetrate more easily the tissues
- the composition of the blocking solution: we add 10% of normal goat serum to the common basal blocking solution (PBS-Tween 0,5% ; 0.5% TritonX-100 ; 1% BSA)
- the time of the blocking step: from 2h at RT to o/n 4°C under rocking
- the concentration of the primary antibody: from 1:1000 to 1/250

II. Preliminary results

Even after the attempts to optimize the protocol, no clear results have been obtained. Both anti-egr1 (1ERG) and anti-klf2b (3KLF) do not recognize native proteins in tissues. The permeabilization steps have improved the entry of the antibodies inside the heart but they seem to bind un-specifically red blood cells has illustrated on Figure 18 (white arrow). A very faint signal could be detected in few endocardial cells only when increasing the laser detection power, but it is more likely an artefact.

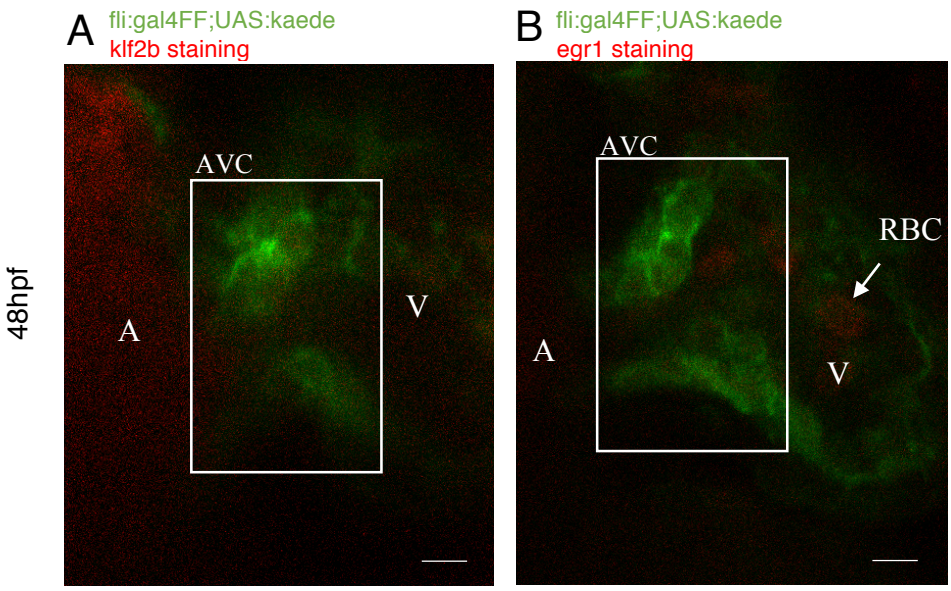


Figure 18: Immuno-Fluorescence assay with anti-Klf2b and anti-Egr1 antibodies

Incubation of home-made monoclonal anti-Klf2b (3KLF) (**panel A**) and of anti-Egr1 (1ERG) (**panel B**) in fixed-48hpf WT (AB) zebrafish embryos. SP8 confocal imaging. Scale bar: 10um

III. Perspectives

This assay would require more time to be fully optimized. It would be interesting to test other anti-Klf2b and anti-Egr1 sera. Some antibodies could work better on IF than other. Their concentration could be increased too to be able to discriminate between real signal and background/ artefacts.

The current data do not permit to conclude regarding the cellular type of localization of Egr1 and Klf2b proteins.

Chapter 3

Elucidating the roles of *egr1*, *klf2a/klf2b* and *flt1* genes in valve formation

The results of this chapter - and those of the following chapter 4 - are part of a scientific paper which will be published. The first draft of the paper is available on the following page 89. It will be modified before sending to editors. In particular the analysis of the data about the double mutant *klf2a/klf2b* will be further investigated.

Chapter 3 and 4 will bring some complementary information not included in the manuscript.

The *egr1-flt1* axis defines a mechanosensitive pathway required for valve morphogenesis

Nathalie Faggianelli-Conrozier[1,2,3,4], Katerina Polizou[5], Renee Chow[1,2,3,4], Eirini Trompouki[5,*], Julien Vermot[1,2,3,4,*]

[1]Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67404 Illkirch, France
[2]Centre National de la Recherche Scientifique, UMR7104, 67404 Illkirch, France
[3]Institut National de la Santé et de la Recherche Médicale, U964, 67404 Illkirch, France
[4]Université de Strasbourg, 67404 Illkirch, France
[5]Department of Cellular and Molecular Immunology, Max Planck Institute of Immunobiology and Epigenetics, 51 Stübeweg, 79108, Freiburg, Germany

Correspondance: trompouki@gmail.com, julien@igbmc.fr

[pages 90 to 110 were removed for confidentiality until the manuscript will be accepted and published and also Annex 2 for the same reasons]

Data availability-Accession numbers

All sequencing data will be deposited in the Short Read Archive (SRA https://www.ncbi.nlm.nih.gov/sra) under BioProject accession codes: xxxx

Author Contributions

N.F. and J.V. designed the experiments. N.F. performed the experiments, with the help of the following persons for some experiments: R.C. for the photoconversion experiments, S. R. for the qPCR experiments, K. P. for the bioinformatics analysis for the treatment of raw ATAC and mRNAseq data, footprinting analysis and E.T. for guidance and help on this analysis, and for providing reagents and critical insights. J.V. and N.F. wrote the paper together.

Acknowledgements

We thank E. Troumpoki, K. Polizou and Vermot laboratory for discussion and thoughtful comments on the manuscript. We thank the IGBMC fish facility (S. Geschier and S. Gredler) and the IGBMC imaging center, in particular E. Guiot, E. Grandgirard and D. Hentsch. We thank M. Philipps, B. Jost, C. Thibault-Carpentier in the IGBMC Microarrays deep sequencing platform, a member of the 'France Ge´nomique' consortium (ANR-10-INBS-0009), for performing mRNA sequencing and ATAC sequencing experiments. This work was supported by ERC Consolidator Grant - European Research Council (ERC) – 2015 and an FRM PhD grant 2018.

Materials and Methods

Zebrafish accession gene numbers
klf2a: ENSDARG00000042667
klf2b: ENSDARG00000040432
egr1: ENSDARG00000037421

flt1: ENSDARG00000019371

Zebrafish husbandry, fish strains and embryo treatments

The zebrafish (Danio rerio) lines used in the experiments were the following: wild-type AB, Tg(egr1^{sa64}) - from ZIRC, received as an heterozygous line ; Tg(klf2a^{ig4}) - generated in Vermot's lab using TALEN technology, published in Steed et al. 2017 ; Tg(klf2b^{ig5}) - generated in Vermot's lab using TALEN technology (further information below), Tg(flt1^{sa1504}) - from ZIRC, received as an heterozygous line, Tg(pkd2^{tc321}) or cup mutants from ZIRC, Tg(vlt^{m651}) or vald tepes, gata1 mutant, from ZIRC, Tg(tnnt2a^{b109}) or silent heart (sih) mutant from Stainier's lab (Sehnert et al. 2002). Tg(fli:gal4FF^{ubs}; UAS:kaede) (Herwig et al. 2001) from Markus Affolter's lab (Biozemtrum, Basel) and Tg(flk1:mCherry) (Bertrand et al., 2010),

Embryos were staged according to hours (hpf) and days postfertilization (dpf). They were incubated at 28.5°C in 0.3% Danieau medium supplemented 5h after birth with 0.003% (wt/vol) 1-phenyl-2-thiourea (PTU) (Sigma Aldrich) to inhibit pigment formation. All zebrafish strains were maintained at the IGBMC fish facility under standard husbandry conditions (14h light/10h dark cycle). The Animal Experimentation Committee of the Institutional Review Board of IGBMC approved all animal experiments performed in this project.

klf2a^{ig4} mutant line: Mutant described in Steed et al. 2017. The INDEL generated using TALEN technology leads to premature stop codon in the protein. A second ATG could be a putative protein-coding start site, this site will be conserved in *klf2a^{ig4}* mutant. Its effective presence and cellular activity have still to be proved.

klf2b^{ig5} mutant line: Tg(Klf2b^{ig5}) line was generated in the same way as Tg(Klf2a^{ig4}), using TALEN system. A TALEN pair (left and right arms: 5'GGACATGGCTTTACCT-3' and 5'AACGTTTGCAAACCAG-3') were designed to target exon 1 of the klf2b gene and injected into single cell wild-type (AB) first cell. We identified the alleles generated and confirmed that potential targeting events could be transmitted through the germline by out-crossing the F0 fish with AB animals and sequencing genomic DNA from pools of 6 F1 embryos. We After screening the first generation, we focused on a 28pb-deletion mutation (5'-GGACATGGCTTTACCTTGCCTTTTGCCT-3') leading to a premature stop codon in klf2b transcript. Studies were performed from F4 fish and later generations, and on transgenic lines resulting of out-crossings. A PCR-based genotyping strategy was established using the following primers to identify the wild-type and mutant alleles, the length of the deletion allowing their visualization directly on a 3%-agarose DNA gel: forward 5'-GGAAAGCGCGTATATTTGGA-3', reverse 5'- CAAGTAGGAAATGCAAGTGT-3' and sequencing forward primer 5'- AGAGCGCACTGTGCCTTATA-3'.

egr1^{sa64} mutant line: Egr1^{sa64} mutant comes from Stemple line (ZIRC), adult males were treated with ENU. It contains a C>T point mutation in the exon 2 of the gene leading to a premature stop codon in the predicted translation product. Studies on Tg(egr1^{Sa64}) line were realized - after cleaning to obtain a stable line - on F4 fish and later generations, following out-crossing to transgenic lines of interest. Genotyping was performed by sequencing of the PCR product generated with the following primers: forward 5'-ATATCCTACACAGGCCGTTTCAC-3', reverse 5'-CACTGGGATATGTTGATGAGGAG-3', and sequenced with the primer 5'-TGGTTGAGAAGTTAGGGCCAGAC-3'.

Klf2a^{ig4}, *klf2b*^{ig5} and *egr1*^{sa64} fish were viable and kept as homozygous fish to be able to work with maternal zygotic embryos in this study. Double mutant lines (*klf2a-klf2b, egr1-klf2a*) came from crossings between single mutant lines and following genotyping.

Flt1 ^{sa1504} mutant line: Flt1^{sa1504} mutant comes from Stemple line (ZIRC), adult males were treated with ENU. It contains a T>G point mutation in the exon 9 of the gene, which in theory leads to a stop codon preventing to form both soluble and membrane flt1 forms (
Studies on Tg(flt1^{Sa1504}) line were realized from F4 fish and later generations, following out-crossing to transgenic lines of interest. Genotyping was performed by sequencing of the PCR product generated with the following primers: forward for PCR and genotyping: 5'-TTAGGCTGAAGGATGGGATG -3', reverse 5'- TGGTCCTCTTTGAACAACCA-3'.

Genotyping

Genotyping of adult fish was performed three months after birth on genomic DNA extracted from a small piece of the caudal fin lysed at 55°C in 100µL of SDS-lysis buffer (10mM Tris-HCl pH8, 200mM NaCl, 10mM EDTA, 0.5% SDS and 100 µg/mL Proteinase K), and then purified using isopropanol/70%-ethanol. Genotyping of embryos was either realized on whole embryos (after live imaging or ISH) or from dissected tails (before immunofluorescence), lysed in 50µL of 50mM NaOH at 95°C for 10min. Lysis reaction was stopped with 10µL Tris-HCl pH8. 1µL used for PCR reaction.

In Situ Hybridization (ISH)

ISH assay was performed as in Thisse (2008) using the following anti-sense probes:
klf2a probe: obtained by PCR amplification of the plasmid IRBOp991B0734D, provided by RPDZ, Berlin, using forward primer: 5'- CAGGCGACTACAGAATGCA -3' and reverse primer: 5'-TAATACGACTCACTATAGGGAGTGAC. Transcription with T7 polymerase.
Klf2b probe: obtained by PCR amplification of the plasmid #1343 pSCB-klf2b, provided by Mrs Cecile Otten, Seyfried group), 640pb fragment amplified with reverse primer 5'-CTACGGTCCGGTGATAGGCATG-3' and forward primer 5'-AGCATTTAGGTGACACTATAGTCACAGGTGTCTCTTCATGTGCAG-3'. Transcription by SP6 polymerase.
Egr1 probe: from pBSK plasmid, PCR amplification using forward primer: 5'-ATGACCCGTGAGTCAGTAA-3' and reverse primer: 5'-ATTAACCCTCACTAAAGGGACTTGGTGCCCTGAGTTCTGAT-3'. Transcription using T3 polymerase.
Flt1 probe: from a partial cDNA fully sequenced ordered at Science Biosource (IRCYp5023H065D, 9038840 IMAGE ID), pCR4-TOPO plasmid, linearized by NCOI (NEB). Transcription using SP6 polymerase.
Wnt9b probe: forward primer: 5'-TATTGCCCTCTGCATCCTTC-3' and reverse primer: 5'-TGACATTCAACGTGACAGCA-3'.
To optimize the imaging of the heart after staining, in certain cases the embryos were made more transparent using fructose. A first bath in 45%-fructose solution (D-fructose, Sigma, ref F0127) containing 1/100 1-Thioglycerol (Sigma M6145) was performed for 15 min followed by a second bath in 90%-fructose solution from 30min to 1 hours depending the stage of the embryo. Imaging was performed using a Leica M165 macroscope with a TrueChrome Metrics (Tucsen) with a Leica 1.0X objective (10450028).

High throughput dissection of 48hpf-hearts

Hearts were dissected from *klf2a*^{ig4}*;klf2b*^{ig5}, Tg(egr1^{sa64}) and Tg(klf2a^{ig4}) lines in background Fli:lifeact-eGFP; flk:nls-mcherry and respective control embryos at the desired stage (48hfp) using high-throughput extraction technique, optimized from Lombardo et al. (2015) protocol. Hearts were separated mechanically by pipetting up and down a batch of 200 embryos in culture medium (L-15 Leibovitz, 10% FCS 9150**,** 1.25mM CaCl2**,** 800mg/L glucose**,** 50microg/mL penicillin**,** 0.05mg/mL streptomycin), medium prepared fresh for each experiment. Hearts were then separated from debris after passages on two different filters: first purification though a 100um nylon cell stainer (Falcon, 352360), and further collection on a pre-separation filter 30um (Miltenyi Biotech). After centrifugation 10 min at 2.5g, culture medium was replaced by cold FACS medium (PBS with 2% Fetal Calf Serum 9150, 1% Penicilline/streptavidine, 1mM EDTA) and final filtration was performed with tip strainers (40um) (Scienceware, Bel-Art, flowmi tip strainers)

development

Fluorescence-activated cell sorting (FACS)

Further separation of endocardial cells from the other cardiac cell types was achieved directly after heart extraction by FACS. FACSAria Fusion (BD Biosciences) device was used. Cells were sorted on a FITC-A detector and flow cell passed through 70um-nozzle.

Sorted cells were collected at 4°C in PBS: 9,5uL cold PBS + 0,24uL RNAsin and immediate freezing in dried ice of 1000 cell-samples for mRNA sequencing assay; or 50uL cold PBS for the samples directly treated after sorting with the transposase for ATACseq assay.

mRNA-sequencing

1000 endocardial cells were collected in 9.4 uL PBS-RNase free supplemented with 0,24uL RNAsin and immediately frozen after FACS sorting in dried ice and stored at -80°C. cDNA transcription was performed using Clontech SMART-Seq v4 Ultra Low Input RNA Kit. Paired-end 100 bp reads for control and experimental samples were generated with Illumina Hiseq 4000. RNA-seq raw sequencing data from zebrafish were trimmed and aligned to zebrafish genome version GRCz10/danRer10, with the tophat algorithm (version 2.1.1) (Kim et al., 2013) and the use of «--b2-very-sensitive» parameter. Samtools (version 0.1.19) (Li et al., 2009) were used for data filtering and file format conversion while the HT-seq count (version 0.5.4p3.) (Anders et al., 2015) algorithm was applied to assign aligned reads to exons using the following command line «htseq-count –s no –m intersection -nonempty». Differentially expressed genes were identified with the use of the DESeq R package (Anders and Huber 2010), and genes with fold change cut-off 1.5 and P<=0.05 were considered to be differentially expressed (DEGs). Heatmaps that were constructed to depict DEGs or selected genes were generated with R/Bioconductor. *A summary of the differentially expressed genes can be found in Supplementary Tables* *****

ATAC-sequencing

After FACS sorting, cells harvested in PBS were treated according to Buenrostro et al. (2015) protocol. Transposition reaction and purification were directly performed the day of cell collection, DNA resuspended in Buffer EB from the MinElute kit was stored at -20°C upon further treatment. Custom Nextera PCR Primer were used during PCR amplification to label the samples. Library quality control was done using Bioanalyser, a purification step was

performed using MiniELute Qiagen kit and size separation on Agencourt AMPure XP beads (A63882, Beckman Coulter, Inc) to remove PCR Nextera primers.
Sequencing was performed on one lane on Illumina HiSeq 4000 sequencing, Nextera, HS-2x50 millions bases for control and mutants. Bowtie2 algorithm (version 2.1.0) (Langmead and Salzberg 2012) and «--very-sensitive» parameter were used for aligning ATAC-seq data to the zebrafish genome version GRCz10/danRer10. Samtools (version 0.1.19) (Li et al., 2009) were used for data filtering and file format conversion. Duplicate reads were removed before peak calling. The MACS2 (version 2.1.0) algorithm (Zhang et al., 2008) was used for ATAC-seq peak identification with default p-value 1-E05. All .bam files were converted to bedgraphs with genomeCoverageBed and MACS2 bdgdiff command was used in order to identify differential enrichment in the accessible regions between control and mutants, with default options. Gene annotation (100 kB upstream and 10kB downstream ftom the tss) and genomic distribution of accessible regions identified by MACS2 was performed with bedtools (Quinlan and Hall 2010). *A summary of ATAC-seq peaks can be found in Supplementary Data xxxxxx.*

Gene ontology, pathway and network analysis
Gene ontology and pathway analysis of the differentially expressed genes from RNA-seq and ATAC-seq was performed with DAVID knowledgebase (Huang da et al., 2009) (Huang da et al., 2009) and Ingenuity Pathway Analysis software (IPA, Ingenuity® Systems, www.ingenuity.com) with the default settings. Only pathways and biological processes with p-value <=0.05 were considered to be significantly enriched.

Network construction
Cytoscape (Shannon et al., 2002) and Metascape (Tripathi et al., 2015) were used for network construction.

Motif analysis
Motif analysis for ATAC-seq peaks was performed with "findMotifsGenome.pl" from HOMER software (Heinz et al., 2010).

Digital genomic footprinting for ATAC-Sequencing
To attain sufficient sequencing depth to perform digital genomic footprinting (DGF) on ATAC-Seq data produced in this work, aligned bam files were merged using samtools merge (version 1.3.1) (Li et al., 2009) following sorting via samtools merge and subsequently indexed using samtools index. DGF was
performed using dnase_footprints of the Wellington pyDNase package (version 0.2.4) (Piper et al., 2013) on total merged ATAC peaks using -A as a parameter to enable ATAC mode, resulting in coordinate shift 5' and 3' by +4 and -5 bp, respectively. Motif overrepresentation and average profile analyses were performed with dnase_average_profile.py of the Wellington pyDNase package (Piper et al., 2013), on WT-only footprints versus all mutant footprints and vice-versa. Briefly, motif bootstrapping was performed in WT-only and mutants-only footprints using HOMER motifs corresponding to de novo found motifs found in WT- and mutants-specific footprints.

qPCR
Whole hearts where extracted as explained previously. RNA was extracted directly after heart collection inside RNA lysis buffer from Quick RNA micro-prep (Zymo kit, R1050). Around 1mg of each RNA mix was used to synthetized cDNA using SuperScript VILO kit (Invitrogen, ref: 11754-050) in 20uL final volume according to the manufacturer's instructions. Products

were amplified in a real-time PCR reaction with Light Cycler 480 Real-Time PCR System (Roche) using a UPL Probes Master mix (Roche) according to the manufacturer's instructions. Sequence of primer pairs were as follows:

gene	Forward primer	Reverse primer
egr1	agtttgatcaccttgctggag	ggtgaaacggcctgtgtaag
flt1	agtcacgtccacccacaatc	gccaactgtcagaactccaac
gata5	ggacgccagggaactctac	acacggcaggtcatccag
has2	agcatccctgttcaactaacg	gctgaccgctttatcacatct
klf2a	ccgtctatttccacattttcg	tccagttcatccttccacct
klf2b	cgtggacatggctttacctt	ttgtgctcctcaatcttcca
sox17	caacccacagttctggtcaa	ccccaaaagagagacagtgc
tbx20	gacttatgctggagatgaagagact	tgcagtgaacgctgaacc
wnt9b	cttattgccctctgcatcct	gaggctcgcgacctgtaa
polr2d-2	aacgcaagtgggagatgtg	agcgtctctgcgttctcaa

Valve imaging

Embryos were incubated with 4mM BODIPY-FL C5-ceramide (Molecular Probes from ThermoFischer, D3521) overnight and then anesthetized with 0.02% Tricaïne (Sigma Aldrich, Saint-Louis, USA) solution and were mounted on Petri dishes embedded in 0.7% low melting point (LMP) agarose (Sigma Aldrich). To visualize the valve shape and to observe their motion, confocal imaging was performed on a Leica SP8 confocal microscope. Images were acquired with a low-magnification water immersion objective (Leica HCX IRAPO L, 25X, N.A. 0.95). For four-dimensional imaging, time series were acquired at a random time in the cardiac cycle at 35fps for 3 s. The optical plane was moved 2 mm between the z-sections until the whole AVC was acquired. AVC area was visually defined as in (Steed et al., 2016). The number of cells within each AVC was analyzed using Imaris software (Bitplane).

AVC photoconversion

Photoconversion was performed at 48 hpf and follow-up imaging experiments were performed at 98 hpf using the protocol described in Chow et al., 2018 (Jove Paper). Briefly, Tg(fli1a:Gal4FF; UAS:Kaede) embryonic hearts were induced to stop beating by treating embryos with 50 mM 2,3-Butanedione monoxime (BDM) (Sigma Aldrich), a myosin inhibitor, for 5-10 minutes. They were then mounted in 0.7% low melting-point agarose supplemented with 50mM BDM to inhibit heart contraction for the duration of the photoconversion procedure. 48hpf- ventricular cells were selected using the Region of Interest tool and exposed to 405 nm light (400 Hz, 25% laser power) using the FRAP module on a SP8 confocal microscope and a Leica HCX IRAPO L, _25, NA0.95 water immersion objective. For each selected region, one pre-bleach frame was acquired, followed by 3–6 bleach pulses with acquisition to convert kaede protein in the illuminated region from its green to red form. A z-stack of the photoconverted heart was then acquired in the standard confocal mode to record the starting point of each experiment. Embryos were then carefully removed from the agarose using a glass suction pipette, washed in fish water, and then placed in fresh fish water containing 0.003 % PTU. Heart contraction soon resumes, and the fish are allowed to develop normally at 28.5 °C under standard conditions until the second timepoint of interest (98hpf).

References

Anders, S. and W. Huber, (2010). "Differential expression analysis for sequence count data." Genome Biol 11(10): R106.

Anders, S., P. T. Pyl and W. Huber, (2015). "HTSeq--a Python framework to work with high-throughput sequencing data." Bioinformatics 31(2): 166-169.

Banjo, T., J. Grajcarek, D. Yoshino, H. Osada, K. Y. Miyasaka, Y. S. Kida, Y. Ueki, K. Nagayama, K. Kawakami, T. Matsumoto, M. Sato and T. Ogura I, Banjo, T., J. Grajcarek, D. Yoshino, H. Osada, K. Y. Miyasaka, Y. S. Kida, Y. Ueki, K. Nagayama, K. Kawakami, T. Matsumoto, M. Sato and T. Ogura (2013). "Haemodynamically dependent valvulogenesis of zebrafish heart is mediated by flow-dependent expression of miR-21." Nat Commun 4: 1978.

Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, (2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position." Nat Methods 10(12): 1213-1218.

Buenrostro, J. D., B. Wu, H. Y. Chang and W. J. Greenleaf, (2015). "ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide." Curr Protoc Mol Biol 109: 21 29 21-29.

Chappell, J. C., K. P. Mouillesseaux and V. L. Bautch I, Chappell, J. C., K. P. Mouillesseaux and V. L. Bautch (2013). "Flt-1 (vascular endothelial growth factor receptor-1) is essential for the vascular endothelial growth factor-Notch feedback loop during angiogenesis." Arterioscler Thromb Vasc Biol 33(8): 1952-1959.

Chen, J. N., F. J. M. van Eeden, K. S. Warren, A. Chin, C. Nüsslein-Volhard, P. Haffter and M. Fishman I, Chen, J. N., F. J. M. van Eeden, K. S. Warren, A. Chin, C. Nüsslein-Volhard, P. Haffter and M. Fishman (1997). "Left-right pattern of cardiac BMP4 may drive asymmetry of the heart in zebrafish." Development 124: 4373-4382.

Francois, M., P. Koopman and M. Beltrame, (2010). "SoxF genes: Key players in the development of the cardio-vascular system." The International Journal of Biochemistry & Cell Biology 42(3): 445-448.

Gassmann, M., F. Casagranda, D. Orioli, H. Simon, C. Lai, R. Klein and G. Lemke, (1995). "Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor." Letters to nature 378.

Ghazvini-Boroujerdi, M., J. Clark, N. Narula, E. Palmatory, J.-M. Connolly, S. DeFelice, J. Xu, B. Jian, S. Hazelwood and R.-J. Levy I, Ghazvini-Boroujerdi, M., J. Clark, N. Narula, E. Palmatory, J.-M. Connolly, S. DeFelice, J. Xu, B. Jian, S. Hazelwood and R.-J. Levy (2004). "Transcription factor Egr-1 in calcific aortic valve disease." J heart Valve Dis. 13(6): 894-903.

Goddard, L. M., A. L. Duchemin, H. Ramalingan, B. Wu, M. Chen, S. Bamezai, J. Yang, L. Li, M. P. Morley, T. Wang, M. Scherrer-Crosbie, D. B. Frank, K. A. Engleka, S. C. Jameson, E. E. Morrisey, T. J. Carroll, B. Zhou, J. Vermot and M. L. Kahn I, Goddard, L. M., A. L. Duchemin, H. Ramalingan, B. Wu, M. Chen, S. Bamezai, J. Yang, L. Li, M. P. Morley, T. Wang, M. Scherrer-Crosbie, D. B. Frank, K. A. Engleka, S. C. Jameson, E. E. Morrisey, T. J. Carroll, B. Zhou, J. Vermot and M. L. Kahn (2017). "Hemodynamic Forces Sculpt Developing Heart Valves through a KLF2-WNT9B Paracrine Signaling Axis." Dev Cell 43(3): 274-289 e275.

Goishi, K., P. Lee, A. J. Davidson, E. Nishi, L. I. Zon and M. Klagsbrun, (2003). "Inhibition of zebrafish epidermal growth factor receptor activity results in cardiovascular defects." Mechanisms of Development 120(7): 811-822.

Harris, T. A., M. Yamakuchi, M. Kondo, P. Oettgen and C. J. Lowenstein, (2010). "Ets-1 and Ets-2 regulate the expression of miR-126 in endothelial cells." Arterioscler Thromb Vasc Biol 30(10): 1990-1997.

Helmlinger, G., R. V. Geiger, S. Schreck and R. M. Nerem I, Helmlinger, G., R. V. Geiger, S. Schreck and R. M. Nerem (1991). "Effects of Pulsatile Flow on Cultured Vascular Endothelial Cell Morphology." journal of biomechanical engineering 113: 123.

Heckel, E., F. Boselli, S. Roth, A. Krudewig, H. G. Belting, G. Charvin and J. Vermot, (2015). "Oscillatory Flow Modulates Mechanosensitive klf2a Expression through trpv4 and trpp2 during Heart Valve Development." Curr Biol 25(10): 1354-1361.

Hoffmann, J. and S. Kaplan I, Hoffmann, J. and S. Kaplan (2002). "The Incidence of Congenital Heart Disease." Journal of the American College of Cardiology 39(12): 1890-1900.

Huang da, W., B. T. Sherman and R. A. Lempicki, (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc 4(1): 44-57.

Huddelson, J. P., S. Srinivasan, N. Ahmad and J. B. Lingrel I, Huddelson, J. P., S. Srinivasan, N. Ahmad and J. B. Lingrel (2004). "Fluid shear stress induces endothelial KLF2 gene expression through a defined promoter region." Biol chem 385: 723-729.

Hurlstone, A. F. L., A.-P. G. Haramis, E. Wienholds, H. Begthel, J. Korving, F. van Eeden, E. Cuppen, D. Zivkovic, R. H. A. Plasterk and H. Clevers I, Hurlstone, A. F. L., A.-P. G. Haramis, E. Wienholds, H. Begthel, J. Korving, F. van Eeden, E. Cuppen, D. Zivkovic, R. H. A. Plasterk and H. Clevers (2003). "The Wnt/b-catenin pathway regulates cardiac valve formation." Nature 425(6958): 628-633.

Kappas, N. C., G. Zeng, J. C. Chappell, J. B. Kearney, S. Hazarika, K. G. Kallianos, C. Patterson, B. H. Annex and V. L. Bautch, (2008). "The VEGF receptor Flt-1 spatially modulates Flk-1 signaling and blood vessel branching." J Cell Biol 181(5): 847-858.

Khachigian, L. M. I, Khachigian, L. M. (2006). "Early growth response-1 in cardiovascular pathobiology." Circ Res 98(2): 186-191.

Khachigian, L. M., K. R. Anderson, N. J. Halnon, M. A. J. Gimbrone, N. Resnick and T. Collins I, Khachigian, L. M., K. R. Anderson, N. J. Halnon, M. A. J. Gimbrone, N. Resnick and T. Collins (1997). "Egr-1 is activated in endothelial cells exposed to fluid shear stress and interacts with a novel shear-stress-response element in the PDGF A-chain promoter." Arterioscler Thromb Vasc Biol 10: 2280-2286.

Kim, I., T. L. Saunders and S. J. Morrison, (2007). "Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells." Cell 130(3): 470-483.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley and S. L. Salzberg, (2013). "TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." Genome Biol 14(4): R36. Langmead, B. and S. L. Salzberg, (2012). "Fast gapped-read alignment with Bowtie 2." Nat Methods 9(4): 357-359.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing, (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics 25(16): 2078-2079.

Lyons, S. E., N. D. Lawson, L. Lei, P. E. Bennett, B. M. Weinstein and P. P. Liu I, Lyons, S. E., N. D. Lawson, L. Lei, P. E. Bennett, B. M. Weinstein and P. P. Liu (2002). "A nonsense mutation in zebrafish gata1 causes the bloodless phenotype in vlad tepes." PNAS 99(8): 5454-5459.

Pacini, L., S. Suffredini, D. Ponti, R. Coppini, G. Frati, G. Ragona, E. Cerbai and A. Calogeroa, (2013). "Altered calcium regulation in isolated cardiomyocytes from Egr-1 knock-out mice." Canadian Journal of Physiology and pharmacology 91(12): 1135-1142.

Pestel, J., R. Ramadass, S. Gauvrit, C. Helker, W. Herzog and D. Y. Stainier I, Pestel, J., R. Ramadass, S. Gauvrit, C. Helker, W. Herzog and D. Y. Stainier (2016). "Real-time 3D visualization of cellular rearrangements during cardiac valve formation." Development 143(12): 2217-2227.

Piper, J., M. C. Elze, P. Cauchy, P. N. Cockerill, C. Bonifer and S. Ott, (2013). "Wellington: a novel method for the accurate identification of digital genomic footprints from DNase-seq data." Nucleic Acids Res 41(21): e201.

Quinlan, A. R. and I. M. Hall, (2010). "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics 26(6): 841-842.

Renz, M., C. Otten, E. Faurobert, F. Rudolph, Y. Zhu, G. Boulday, J. Duchene, M. Mickoleit, A. C. Dietrich, C. Ramspacher, E. Steed, S. Manet-Dupe, A. Benz, D. Hassel, J. Vermot, J. Huisken, E. Tournier-Lasserve, U. Felbor, U. Sure, C. Albiges-Rizo and S. Abdelilah-Seyfried, (2015). "Regulation of beta1 integrin-Klf2-mediated angiogenesis by CCM proteins." Dev Cell 32(2): 181-190.

Rottbauer, W., S. Just, G. Wessels, N. Trano, P. Most, H. A. Katus and M. C. Fishman I, Rottbauer, W., S. Just, G. Wessels, N. Trano, P. Most, H. A. Katus and M. C. Fishman (2005). "VEGF–PLC gamma1 pathway controls cardiac contractility in the embryonic heart." Genes Dev 19(1): 1624-1634.

Sehnert, A. J., A. Huq, B. M. Weinstein, C. Walker, M. Fishman and D. Y. Stainier I, Sehnert, A. J., A. Huq, B. M. Weinstein, C. Walker, M. Fishman and D. Y. Stainier (2002). "Cardiac troponin T is essential in sarcomere assembly and cardiac contractility." Nat Genet 31(1): 106-110.

Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, (2002). "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." Genome Res 13(1): 2498-2504.

Steed, E., N. Faggianelli, S. Roth, C. Ramspacher, J. P. Concordet and J. Vermot I, Steed, E., N. Faggianelli, S. Roth, C. Ramspacher, J. P. Concordet and J. Vermot (2016). "klf2a couples mechanotransduction and zebrafish valve morphogenesis through fibronectin synthesis." Nat Commun 7: 11646.

Tripathi, S., M. O. Pohl, Y. Zhou, A. Rodriguez-Frandsen, G. Wang, D. A. Stein, H. M. Moulton, P. DeJesus, J. Che, L. C. Mulder, E. Yanguez, D. Andenmatten, L. Pache, B. Manicassamy, R. A. Albrecht, M. G. Gonzalez, Q. Nguyen, A. Brass, S. Elledge, M. White, S. Shapira, N. Hacohen, A. Karlas, T. F. Meyer, M. Shales, A. Gatorano, J. R. Johnson, G. Jang, T. Johnson, E. Verschueren, D. Sanders, N. Krogan, M. Shaw, R. Konig, S. Stertz, A. Garcia-Sastre and S. K. Chanda, (2015). "Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding." Cell Host Microbe 18(6): 723-735.

Vermot, J., A. S. Forouhar, M. Liebling, D. Wu, D. Plummer, M. Gharib and S. E. Fraser I, Vermot, J., A. S. Forouhar, M. Liebling, D. Wu, D. Plummer, M. Gharib and S. E. Fraser (2009). "Reversing blood flows act through klf2a to ensure normal valvulogenesis in the developing heart." PLoS Biol 7(11): e1000246.

Wong, K. S., K. Rehn, S. Palencia-Desai, V. Kohli, W. Hunter, J. D. Uhl, M. S. Rost and S. Sumanas I, Wong, K. S., K. Rehn, S. Palencia-Desai, V. Kohli, W. Hunter, J. D. Uhl, M. S. Rost and S. Sumanas (2012).

"Hedgehog signaling is required for differentiation of endocardial progenitors in zebrafish." Dev Biol 361(2): 377-391.

Zhu, D., Y. Fang, K. Gao, J. Shen, T. P. Zhong and F. Li I, Zhu, D., Y. Fang, K. Gao, J. Shen, T. P. Zhong and F. Li (2017). "Vegfa Impacts Early Myocardium Development in Zebrafish." Int J Mol Sci 18(2).

In zebrafish, valvulogenesis starts between 24 and 48hpf with the formation of an endocardical ring creating a region called AVC. Within the AVC, the pulsatile flow plays critical roles, activating in particular key transcription factors, as Klf2a (Vermot et al. 2009), which is expressed inside the AVC. We were first interested in proving a similar expression of egr1 and klf2b in this area, and a flow-dependent activation of egr1 as postulated by (Khachigian et al., 1997). We also studied klf4 and klf17 gene expression patterns, trying to find putative cross-regulations during valvulogenesis between the sub-family klf2/klf4/klf17 discovered by (Oates et al., 2001). In addition, we studied the flt1^{sa1504} mutant and discovered an egr1/klf2b-dependent expression of flt1 gene.

I. Cardiac gene expression patterns of *egr1* in zebrafish

I.1. Egr1 cardiac expression

(*Supplementary Figure 1 of the paper)*

In the heart, the expression of *egr1* started at 30hpf in a small subset of the endocardial cells at the AV initiation of constriction. At 48hpf, this expression became stronger in AVC with an additional spot in the OFT. Expression is still maintained at high level at 72hpf. We can note a strong expression of *Egr1* in brain and tail vasculature since 30hpf similar as published by (Banjo et al., 2013).

(figure 1 of the paper)

Egr1 expression is blood flow- and klf2-dependent.

I.2. Additional data about *egr2* gene

On mice, another *egr* gene, *egr2/krox20*, is important for cardiac formation (Odelin and Zaffran, 2014). *egr2* mutant mice show aortic valve dysfunction associated with disorganization of the ECM, hypertrophic aortic valve defects and in 27% of them bicuspid aortic valve. The *krox20*-expressing cells constitute a subset of neural crest cells which contribute to arterial cells and OFT (Odelin et al., 2018). *Egr2b* (one of the two *Danio rerio* paralogs of *Egr2*) expression was investigated in zebrafish 30, 48 and 72hpf-embryo. However, no cardiac expression has been observed, as shown on Figure 19. The other paralog, *egr2a*, could be nevertheless expressed and recapitulates the data obtained in mice. However, the lab does not possess the WISH-probe for *egr2a* and time was missing to study further down this point.

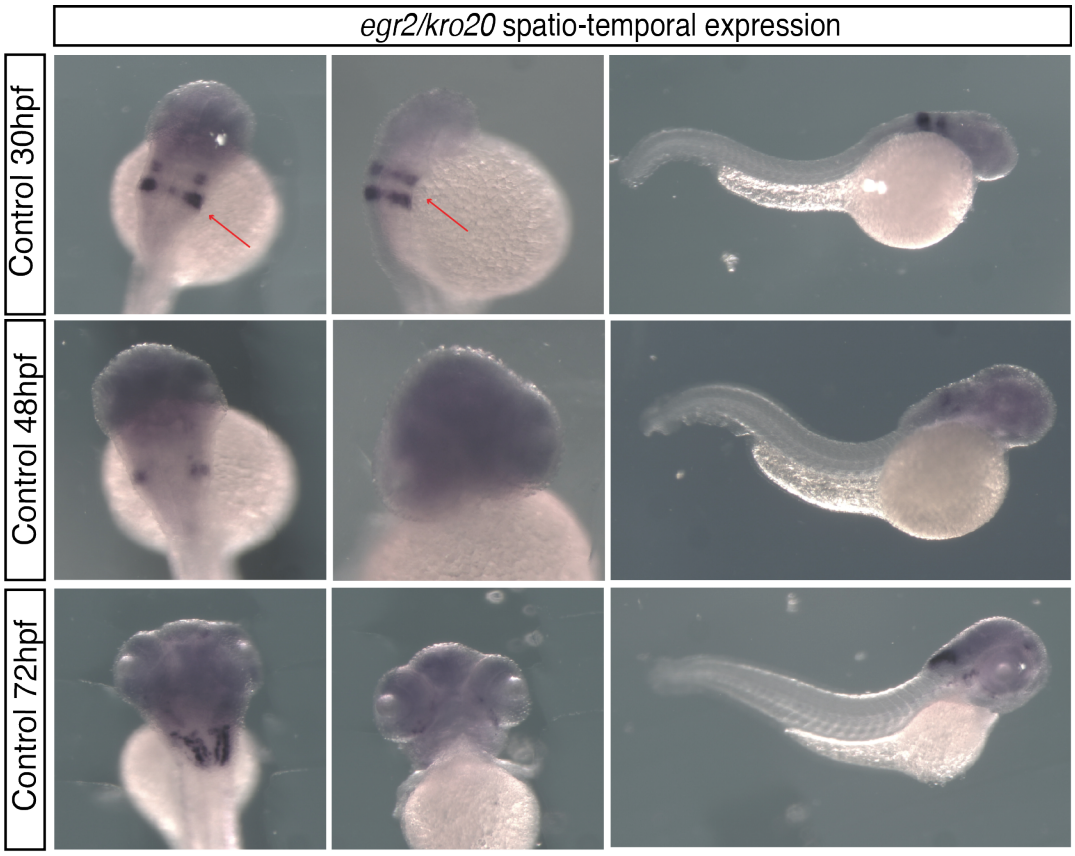


Figure 19: *egr2b/krox20* expression pattern between 30 and 72hpf in WT zebrafish embryo.

Expression was detected in rhombomeres 3 and 5 from 30hpf (red arrows), this expression decreased at 48hpf and is maintained only in hindbrain nuclei at 72hpf. This recapitulates results published by Thisse et al. 2001 (Zfin). No cardiac expression was reported.

II. *Some klf members* cardiac expression

II.1 *klf2b* AVC-expression would be *klf2a*-dependent

As *klf2a* (Vermot et al., 2009), *klf2b* was expressed in the heart tube from 30hpf (Figure 20). At 48hpf, the expression was restricted to AVC and OFT area.

We then compared expression in blood flow mutants (Figure 21, panel B). *Klf2b* expression is reduced in *pkd2*$^{-/-}$ mutant and absent in *silent heart / tnnt2a*$^{-/-}$ mutants. But the expression of *klf2b* in *gata1 / vlad tepes*$^{-/-}$ mutant was not changed. This would suggest different mechanotransduction pathway would regulate *klf2b* expression. As *klf2b* expression is down in *klf2a* mutant and up-regulated in fish where *klf2a* is over-expressed (klF2a OE), it is possible that *klf2a* regulates *klf2b* expression (Figure 21, panel C).

qPCR data showed a downregulation of *klf2b* in *klf2a* mutant (fold change: 0,54, p.value: 0,018). However, this was not recapitulated in the mRNA seq data. We have to bear in mind that both *klf2a* and *klf2b* mRNA are still produced in the mutants in a lower extent. Consequently, it will be interesting to do a qPCR in the *klf2a* OE line to confirm the trend observed.

II.2. *klf4* and *klf17* are not expressed in AVC and OFT between 30 and 72hpf

WISH probes were designed for *klf4* and *klf17* genes. Figure 22 shows the results obtained after incubation of these probes with WT (AB) embryos. Klf4 was highly expressed in brain of 30 to 72hpf-embryo but no signal was detected in the heart. Klf17 seems to be expressed in the heart tube at 30hpf but then the area of expression at 48hpf does not show a specific expression in heart. Some yolk cells were stained around the cardiac area.

In the mRNA-seq data, no deregulation of these two genes was detected at 48hpf in endocardial cells in mutants compared to controls. No feedback regulation seems to happen between the KLF subgroup family of Klf2 / Klf4 and Klf17 in zebrafish. Interestingly, we can nevertheless notice a high expression of *klf17* mRNA in the mRNA-seq data of the controls (around 17, 700 reads against 3,000 for *klf2b* and 600 for *klf2a*). This could be explained by the fact that surprisingly a great number of erythroid markers are present in the mRNA-seq data (hemoglobin is one of the first hit – *this will be discussed later*) and *klf17* was shown to be involved in erythroid cell differentiation and no heart expression of this factor was mentioned to date (Kawahara and Dawid 2001).

In literature, *klf4a*/*klf4* was shown to be involved in heart valve formation with *klf2a*. (Just et al., 2011) studied the zebrafish *bng* mutant (a mutation within the kinase domain of PKD2 protein which abrogates the PKD2-HDAC5-KLF pathway). *Klf2a* and *klf4* expression were down in *bng* mutant hearts at 72hph. The WISH in our study were realized earlier (from 30 to 56hpf) and could be done at 72hpf trying to recapitulate the cardiac expression observed by Just et al. *Klf4* is above all important for primitive erythropoiesis (Gardiner et al., 2007).

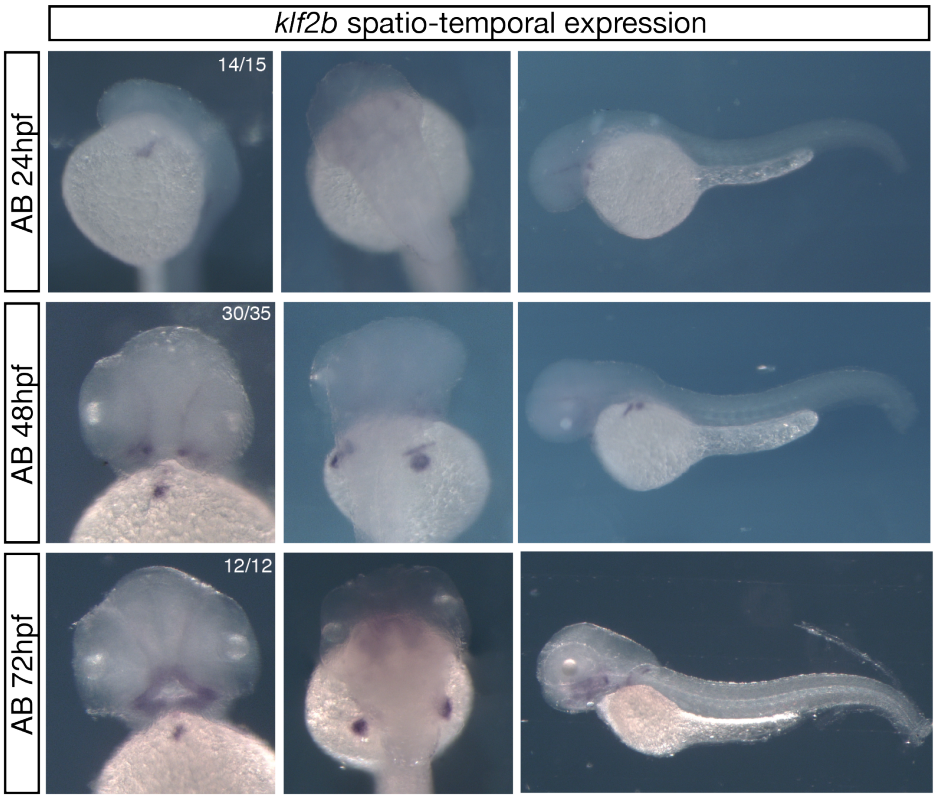


Figure 20: *klf2b* spatio-temporal expression (WISH assay) on 30-48 and 72hpf zebrafish embryos

klf2b is expressed from 30hpf in the heart tube and then its expression becomes restricted similarly to *klf2a* expression in AVC and in a small extend in OFT at 48hpf. These area of expression are maintained at 72hpf. Klf2b is also detected in the pectorial fin buds at 48 and 72hpf. No expression in the tail vasculature.

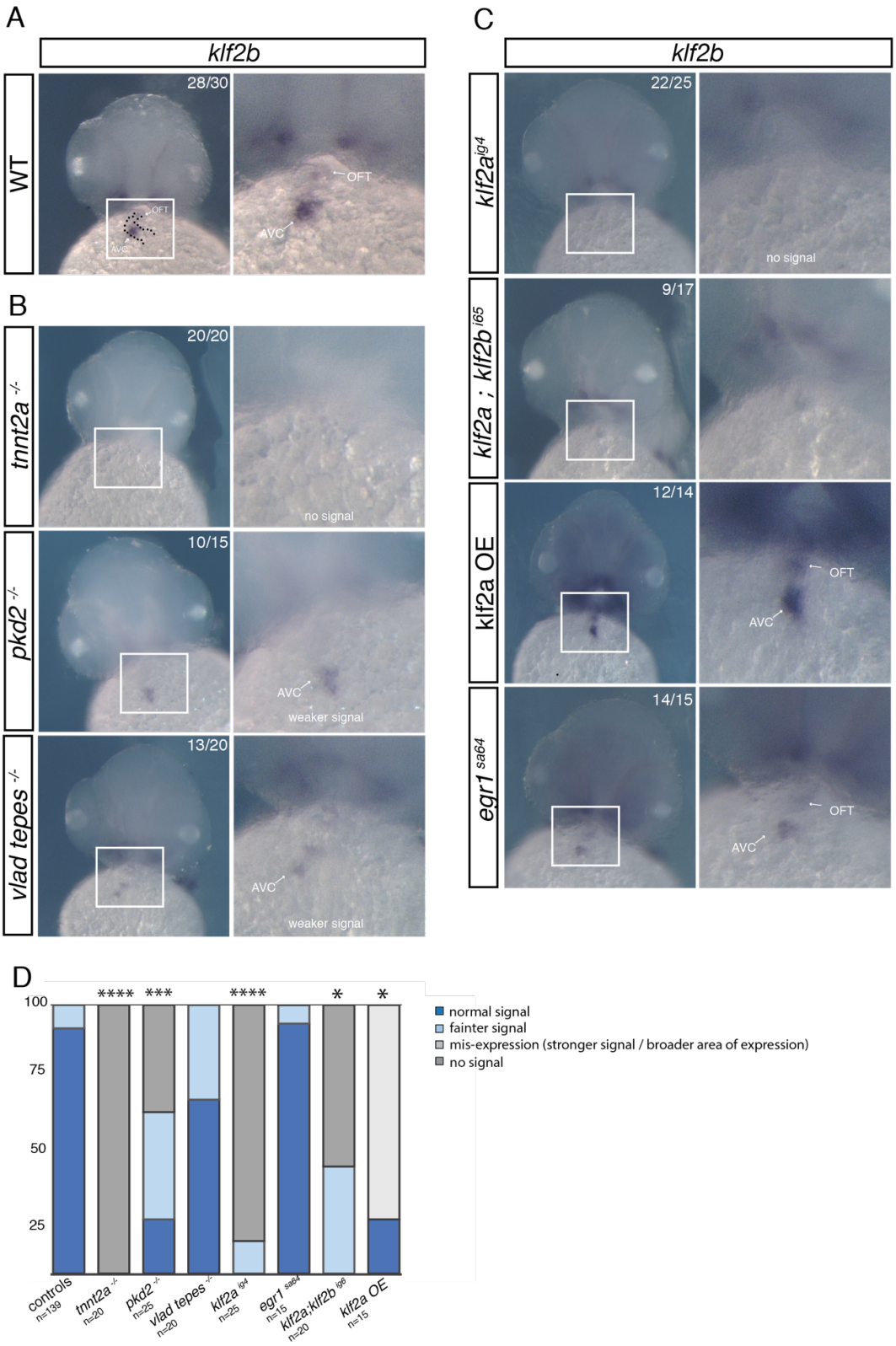


Figure 21: *klf2b* WISH on blood-flow, *klf* and *egr1*^{sa64} mutants at 48hpf

Panel A: *klf2b* cardiac expression at 48hpf is restricted to valve forming area (AVC and OFT) in WT (AB) embryos. **Panel B:** *klf2b* expression is abnormal in mutants where blood flow is altered (tnnt2) and where mechano-sensors trpp2 is knocked out (pkd2 mutant). This indicates that *klf2b* expression would be flow dependent, but in a different manner of *klf2a* expression. **Panel C**: *klf2b* expression in klf mutants (single mutant *klf2a*, double mutant *kl2fa; klf2b;* the over-expressing *klf2a* line). *Klf2b* expression seems abnormal in *klf2* mutants. This indicates that *klf2b* expression might be modulated by klf2a. No change in expression in *egr1*^{sa64} mutant. **Panel D:** statistical analysis of *klf2b* WISH in the different mutants. The control presented here is the sum of all respective controls for the mutants. *p<0.05, **p<0.01, ***p<0.001, **** p<0.0001 chi-squared analysis for ISH

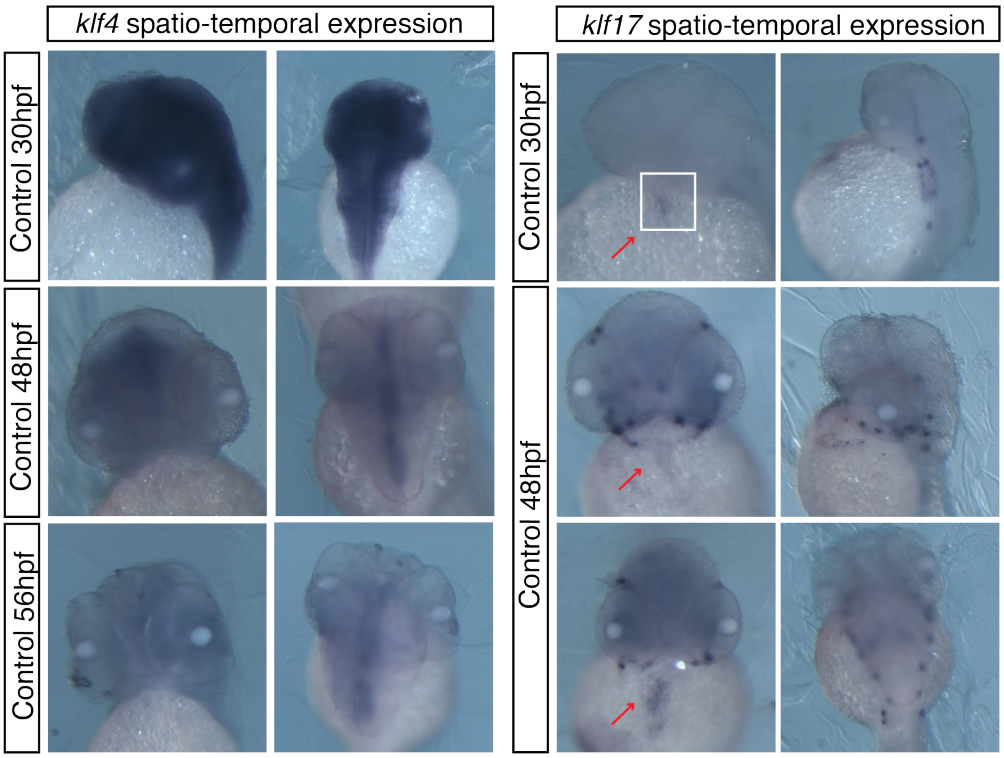


Figure 22: *klf4* and *klf17* WISH expression in 30-48 and 72hpf WT zebrafish embryos.

III. *Additional information about flt1 and flt4 gene expression*

III.1 *flt1 cardiac expression is klf2a- and egr1-dependent*

As described in the main introduction of this thesis, in zebrafish *flt1* gene encodes for both membrane and soluble proteins (mFlt1 and sFlt1). sFlt1 is formed by alternative splicing at the exon 10—intron 10 boundary. It was shown by (Wild et al., 2017) that *mflt1* mutant alleles did not reveal any obvious vascular malformations or alterations in vascular branching morphogenesis compared to mutants which involved mutation of soluble Flt1. These mutants presented indeed a hyper-branching of the trunk vasculature at 3–4 dpf (Wild et al., 2017).
If the mutation is located in the first exons, more likely the mutant will be lacking both Flt1 proteins.

Our mutant *flt1*^{sa1504} presents a point mutation in the exon 9, in an essential splice site. However, this mutant has not been characterized yet. Probably both forms are not encoded but this has to be checked. We haven't finished the complete analysis of this mutant to look for potential defects in the vasculature which could recapitulate the data of (Wild et al., 2017).

We were focusing on the heart defects of this mutant. Zebrafish MZ homozygous *flt1*^{sa1504} displayed severe heart defects, oedema appeared from 48hpf and the mutation penetrance observed was 60%. The complete study of the valve defects is presented in the draft of the paper.

III.2 No *flt4 cardiac expression at 48hpf*

To test the possible implication of other VEGF receptors during valve development, we performed a WISH with *flt4* probe. Figure 23 presents the results obtained, no cardiac expression was detected at this time-point. No change in transcriptomic data was neither observed. This gene was only reported to date in vascular development, in particular it was shown to be important for developing zebrafish intersegmental arteries (Hogan et al., 2009).

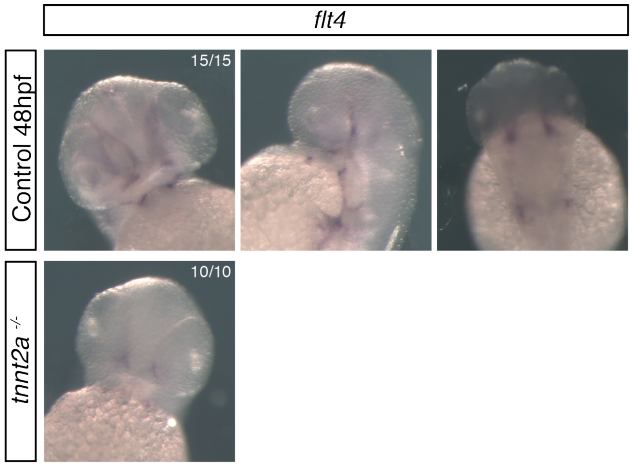


Figure 23: *flt4* WISH on 48hpf-zebrafish embryos

No cardiac expression of *flt4* was reported at 48hpf neither in WT or *silent heart* (tnnt2a) mutants contrary to the results of our collaborators.

Figure 24: *wnt9b* WISH in 30 to 72hpf zebrafish embryos

Wnt9b was detected in the heart tube at 30hpf and its expression was then observed in the ventricle and AVC/OFT area. The ventricular expression became stronger et 48hpf. At 72hpf cardiac expression was absent.

IV. *Additional information about wnt9b spatio-temporal expression in zebrafish heart*

We analysed the spatio-temporal of *wnt9b* gene in zebrafish from 30hpf to 72hpf Figure 24. Cardiac expression was detected in the linear tube at 30hpf and was maintained in the ventricle and in the AVC-OFT areas. The expression seemed even more intense in OFT and AVC. The signal in ventricle became more intense from 46hpf to 48hpf. At 72Hpf it seemed wnt9b was not more expressed in heart. Additional data about the expression of *wnt9b* in *klf* and *egr1* mutants are presented in the paper.

Chapter 4

Deciphering the cardiac downstream transcriptional network of *klf2a* and *egr1* at 48hpf

Development of a Chromatin Immuno-Precipitation (ChIP)

Integration of Assay for Transposase Accessible Chromatin-Sequencing (ATAC-seq) and mRNA-seq data

NB: This part is included in the scientific paper (page 87) and only additional data will be added in this chapter.

I. Preamble: introduction to integrative genomic studies used to understand regulatory networks in development

I.1. Chromatin and gene expression

In the cells, the entire genetic information is stored and packaged into a compact and dense structure called chromatin. This reinforces the DNA macromolecule during mitosis, preventing its damage, and also controlling gene expression and DNA replication. The local structure of chromatin changes depending on the gene expression requirement of each cell: DNA which encodes actively transcribed genes ("turned on genes") and DNA corresponding to biologically active regions (promoter, enhances or other regulatory elements) is loosely packaged, associated with RNA polymerases and thus more accessible to the transcription machinery. This state is referred to as euchromatin (Gross and Garrard, 1988, Bell et al., 2011). While inactive genomic regions that contain inactive genes ("turned off genes") are sequestered in a condensed state - referred as heterochromatin. Very simplistically, this hierarchical packaging is composed of three levels: DNA is wrapped around histone proteins forming nucleosomes, the "beads on a string" structure which constitutes the euchromatin. Multiple histones can then be wrapped into a 30nm-fiber forming compact nucleosome arrays (heterochromatin). An even higher-level of DNA packaging can be observed, when the 30nm-fiber compacts into the metaphase chromosome during mitosis and meiosis (Kornberg, 1974). These different levels of chromatin compaction are summarized on Figure 25.

The molecular DNA compaction into chromatin provides the epigenetic control system that Conrad Waddington first postulated in the 1940s. “Epigenetics“ literally means “over or above genetics”. It refers to hereditary changes in genome expression that do not involve alteration of DNA sequences. Epigenetic code includes DNA methylation, nucleosome positioning, histone composition, and histone modifications. Transcription factors, chromatin remodelers, and non-coding RNAs are the main actors of these epigenetic mechanisms governing cellular processes (Kouzarides 2007); (Rinn and Chang 2012) ; (Chen and Dent 2014).

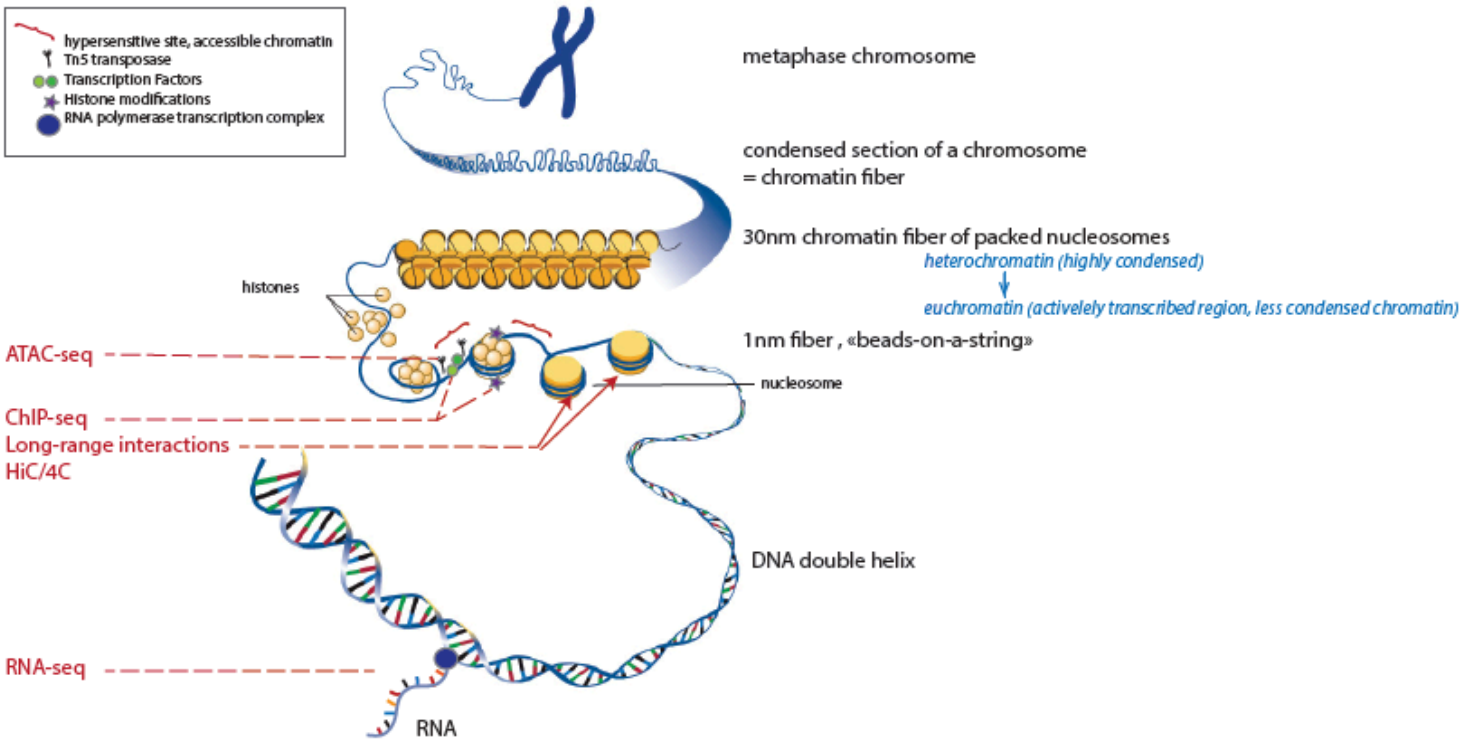


Figure 25: Hierarchical levels of DNA organization in the eukaryotic nucleus.

Nucleosomes are the basic structure unit in eukaryotic chromatin, where DNA wraps around a histone octamer. Post-translational modifications of the histone core affect DNA compaction and act as signaling components for various proteins. These modifications are part of the "epigenetics code". Different genomic techniques are used to study epigenetics and to provide information on gene regulation and expression *(mentioned in red letters in the figure).*

I.2. Genome-wide analysis techniques

Understanding the epigenetic regulation mechanisms and the related chromatin changes provides information on gene regulation and expression, and gene network. By identifying the binding regions of transcription or chromatin factors, but also the regions of open chromatin it is possible to study how this information is translated into particular phenotypes.

In parallel of these studies, the emergence of high-throughput sequencing technologies has changed the way to study the genome and its expression. Coupling molecular biology with sequencing permitted the implementation of new analysis tools. Among the variety of information-rich genome-wide analysis methods, genome-wide measurements of protein-DNA interactions by Chromatin Immuno-Precipitation (ChIP), of chromatin accessibility by Assay for Transposase Accessible Chromatin (ATAC) and quantitative measurements of transcriptomes (RNAseq) are increasingly used to link regulatory inputs with transcriptional outputs (Figure 25 and 26).

Integration of the data produced by the methods mentioned above allows to answer to this type of questions:

II. Which are the downstream transcriptional targets of a transcription factor of interest?
III. Are they direct targets?
IV. Is the transcription factor an activator, repressor, or both?
V. Does it have different binding partners depending on the type of regulation or the cellular environment?

In other words, a computational analysis combining data from each type of assay allows a simultaneous assessment of transcriptome (genome expression profile), epigenomics and/or chromatin dynamics at a given time for one particular cell type. Collective results revealed that changes in chromatin accessibility occurred concomitantly with changes in RNA expression and large-scale genome organization (Duren et al., 2017).

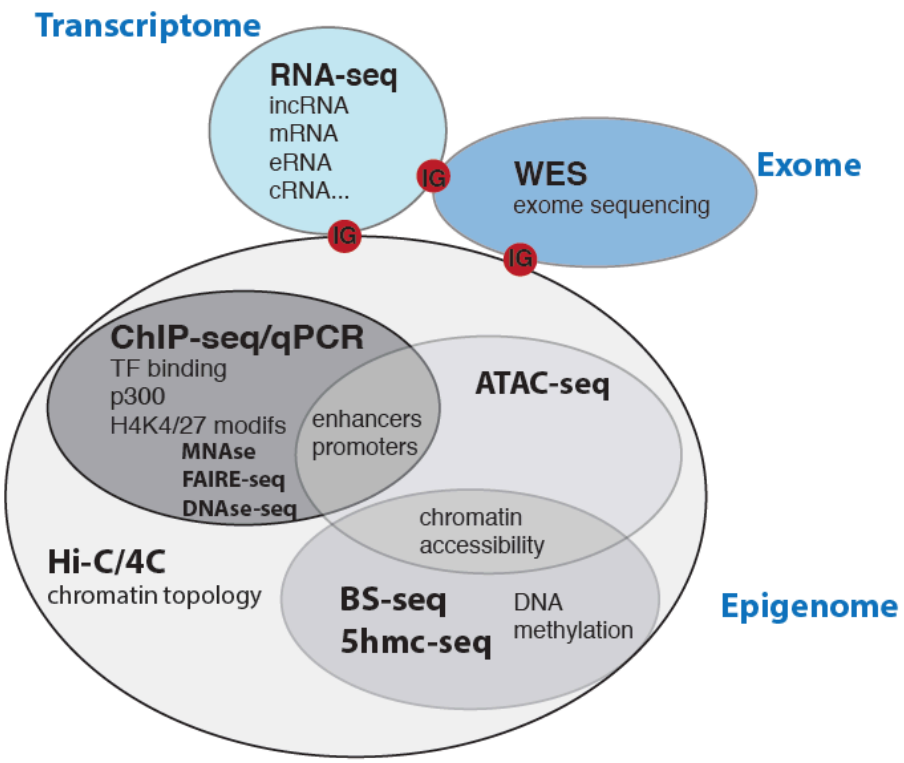


Figure 26: Integration of transcriptomic and epigenetic studies.

With the progress in genomics and in high-throughput sequencing, **I**ntegrative **G**enomics (IG) provides novel insights into the understanding of molecular developmental mechanisms. It would help in particular to decipher the ones driving heart development, by identifying genetic factors and investigating regulation pathways.

RNA-seq: Ribonucleic Acid sequencing (transcriptome sequencing)
WES: Whole Exon Sequencing / WGS : Whole Genome Sequencing
ChIP-seq: chromatin Immuno-Precipitation
ATAC-seq: Assay for Transposase Accessibility
BS-seq : Bisulfite sequencing, used to detect methylated cytosines in genomic DNA

Hi-C / 4C: two 3C technique (Chromosome conformation capture). Hi-C is used to find the nucleotide sequence of fragments. 4C for Chromosome conformation capture-on-chip (4C) captures interactions between one locus and all other genomic loci

MNAse-seq: technique used to distinguish nucleosome positioning based on the ability of nucleosomes to protect associated DNA from digestion by Micrococcal Nuclease. Sequenced fragments reveal nucleosome location information about the input DNA.

FAIRE-seq: Formaldehyde-Assisted Isolation of Regulatory Elements), method used for determining the sequences of DNA regions in the genome associated with regulatory activity.
DNAse-seq: sequencing of regions sensitive to DNase I nuclease, it helps identifying and localysing regulatory regions.

Three types of analysis performed in this thesis to discover downstream (direct or indirect) transcriptional targets of klf2a and egr1.

I.3 Chromatin Immuno-Precipitation followed by high-throughput sequencing or ChIP-seq

ChIP is a method to investigate the protein/chromatin interactions *in vivo*. ChIP can be employed to define the specific DNA binding sites of a particular protein / transcription factor and thus provides evidence for a DIRECT regulation of genes by a transcription factor (Buck and Lieb 2004) (Mardis 2007). A single ChIP experiment can in principle identify all the binding sites of a particular protein under the chosen experimental conditions.

It relies on antibodies to identify the presence of specific histones or histone modifications or transcription factors bound to DNA regions of interest. First, a fixative chemical, usually formaldehyde, will cross-link proteins to DNA and maintaining the chromatin architecture. Glycine is added to quench the formaldehyde and terminate the cross-linking reaction. Chromatin is then extracted from cells, covalent protein-DNA complexes are sheared into ~ 500 bp DNA fragments by sonication and incubated with antibodies against the protein of interest. The chromatin fragments bound to the protein-antibody complex are captured using protein A/G beads, and the immuno-precipitated DNA is reverse cross-linked. Enrichment of specific DNA sequences represents regions on the genome that the protein of interest is associated with *in vivo*. Then, it is possible to focus on specific genes and to investigate protein-DNA interactions at known genomic binding sites, in order to determine the abundance of a region of interest in the precipitated material. For this purpose, the purified DNA is amplified using primers designed at the loci of interest (ChIP-PCR or qPCR); or the entire set of fragments is sequenced and their distribution is mapped across the genome (microarray or ChIP-sequencing). This method is summarized on Figure 27.

I.4. Assay for Transposase-Accessible Chromatin with high-throughput sequencing or ATAC-seq

ATAC coupled with high-throughput sequencing is a powerful method for mapping chromatin accessibility in a genome-wide manner (Buenrostro et al., 2015). This method probes accessible DNA regions recognized by the hyperactive Tn5 transposase, which simultaneously cuts those sites and inserts sequencing adapters. Sequencing reads can infer regions of increased accessibility that can also indicate transcription factor binding sites. ATAC-seq can also lead to the identification of nucleosome positions.

The method is a fast and sensitive alternative to DNAse hypersensitivity assays or FAIRE-seq for assaying chromatin accessibility genome-wide, or to MNase for assaying nucleosome positions in accessible regions of the genome.

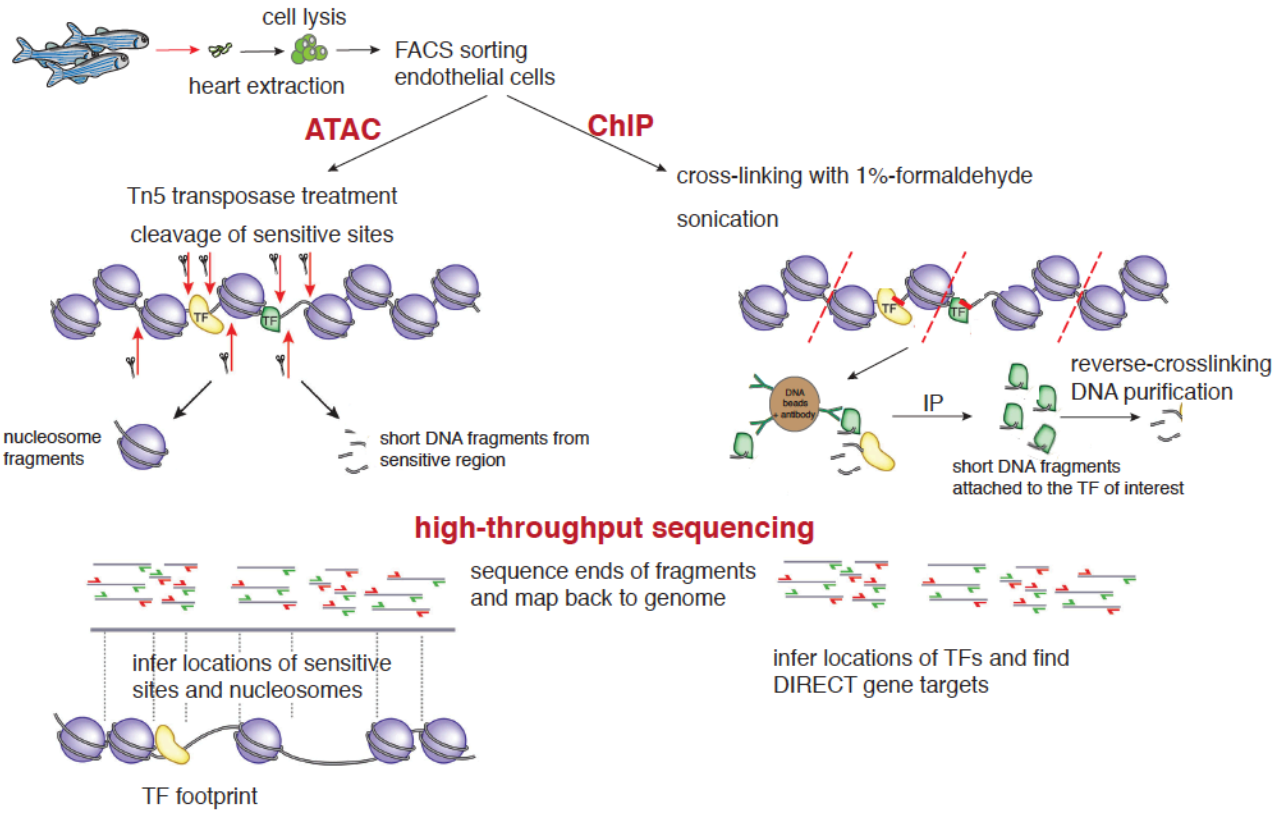


Figure 27: Comparative scheme presenting **ATAC-seq and ChIP-seq principles** and the method used in this thesis on 48hpf-zebrafish endocardial cells.

Adapted from the figure presenting ATAC-seq by (Raj and McVicker 2014)

I.5. Whole transcriptome shotgun sequencing or RNA-seq

RNA-seq refers to a "next generation sequencing" (NGS) approach that offers a snapshot of the entire transcriptome profile at a given time point. It reveals the presence and amount of RNA in a biological sample at a given moment. The shortcut term RNA-Seq is a bit confusing, as RNA is not directly sequenced. Single RNA strands are reverse transcribed into complementary DNAs (cDNA) and then converted to double stranded DNA before being sequenced. cDNA synthesis is performed using adapter primers followed by a PCR amplification step. So, while the initial material is RNA, the sample loaded on the sequencing instrument is DNA. This analysis is used to determine global expression levels of each transcript, identify exons, introns and map their boundaries. IT can also be used to quantify alternative splicing.

It is possible to look only at different populations of RNA and for example to sequence only the messenger RNA population. mRNA-seq protocol permits an enrichment of all polyadenylated (poly-A) transcripts of the transcriptome, which represents 1-2% of the entire transcriptome. By targeting mRNA, sequencing depth is improved as resources are dedicated to the sequencing of coding genes. This makes identifying rare variants and low expressed mRNA transcripts easier. *mRNA-seq will be used in this thesis to study gene expression profile and find differentially expressed genes between mutants and controls.*

During my thesis, I have first worked on the development of a ChIP-assay to investigate the direct targets of Klf2a in 48hpf-endocardial cells. The optimization of the protocol was a long process and the first results obtained were not totally satisfying. We decided to perform an ATAC-seq coupled to mRNA-seq assay, which were "easier" to implement. Integrative analysis of the data is part of the paper which will be sent to editors. In a second part of this chapter will be presented the work realized about the ChIP assay and the experienced difficulties.

II. Optimization of a ChIP method

In parallel to the generation of anti-Klf2a antibodies (*Chapter 1 on this thesis manuscript*), I tried to optimize a ChIP assay, using the commercial anti-KLF2 antibody (Merck Millipore) and a home-made polyclonal anti-Klf2a (n°3290). The aim was to determine new direct transcriptional targets of *Klf2a*.

As presented before, ChIP is a powerful method to identifying genome occupancy by proteins (histones and transcription factors). However, the protocol could be a bit tricky to settle in order to get results. The complete and detailed protocol from Dr. Eirini Trompouki (Max-Planck of Immunobiology and Epigenetics, Freiburg) is available in the Material and Methods part. This part will present the points which have required a specific optimization, the preliminary results obtained and encountered issues.

II.1. Optimization steps

A particular attention was paid on these different points:

II.1.a) Quantity of material

Working with living animals and no cells in culture constitutes a first challenge to overcome. We need to access to chromatin in early-stage zebrafish embryos, which are surrounded by a thick glycoprotein chorion and possessed a large amount of yolk (Lindeman et al., 2010). For transcription factor ChIP it is recommended to use 10^7 cells (Trompouki et al., 2011). A first ChIP assay was performed using 1000 48hpf-whole embryos and some modifications were introduced for the next assays, as presented later). Embryos were dechroniated by Pronase digestion (Roche) and crosslinked with the addition of 1%-formalhehyde for 5 min at RT.

II.1.b) Sonication step

An over-sonication will lead to chromatin degradation and an insufficient sonication to non-specific binding. Sonicated fragments should ideally have an average size of 500 bp ranging from 200 to 1000 bp. After sonication, a part of the sonicated chromatin was reverse crosslinked by boiling for 20 min and loaded on a gelto check for chromatin size. Bioanalyzer analysis (using Agilent 2100) was also performed to get a more accurate estimation of the sonication efficiency.

II.1.c) Antibodies efficiency

The antibody used should be evaluated to determine: 1- if it is able to bind the beads, 2- if it recognizes the native transcription factor of interest and 3- if it is able to pull it down efficiently during the immunoprecipitation step (IP step). A "ChiP-Western Blot" was performed to evaluate these points. A part of the beads after the IP was collected, boiled in loading buffer for 10 min and loaded on a gel.

II.1.d) Choice of controls

To ensure that each step of the experiment is working, different controls were used.

A*ntibody controls*: To determine if the results obtained are correct, we run in parallel a ChIP with antibodies against the histones H3K4me3 and H3K27ac. ChIP for these histones modifications were previously published and marks active regulatory elements in whole zebrafish embryos (Lindeman et al., 2009; Vastenhouw et al., 2010; Quillien et al., 2017). Histone H3 tri-methyl Lysine 4 (H3K4me3) is generally a mark for actively transcribed promoter regions. Histone H3 acetyl Lys27 H3K27ac is also commonly associated with active enhancers and promoters. These antibodies will be "positive controls". As a negative antibody control, we could have used an antibody that recognizes a non-chromatin epitope such as an anti-GFP antibody (*not done here*).

Control for the background of the assay: As a negative control, I use "beads only" or "beads with an isotype matched control immunoglobulin (Ig)".

"Input": A part of the DNA sample that has been cross-linked and sonicated but not immuno-precipitated. It is used to normalize signal from ChIP enrichment.

Controls for qPCR: Primers for both positive and negative loci should be designed. One couple of primers was designed where we know the protein of interest is present (positive control locus) and one where it is absent or significantly decreased (negative control locus). This will reveal whether the observed enrichment is specific. It is important to include these controls as some antibodies result in non-specific enrichment.

II.3. Results of the optimized steps

II.3.a) Sonication time

Sonication was realized using Covaris E220 sonicator (Sonolab 7.3. software). The parameters were the followings: duty cycle 20% - intensity 8 and frequency / cycles per burst 200 and I tried a range of sonication time between 5 and 30 min. A time course was performed using different numbers of sonication cycles (effective sonication during x minutes followed by x min of no sonication to avoid an overheating of the chromatin). The efficiency of the sonication was attested by agarose gel loading and Bioanalyser analysis (Agilent 2100). Results are presented on Figure 28. 30 min of sonication brought the best quality of fragments with the selected parameters.

II.3.b) Antibodies test

The Immublots of IP samples were first incubated with the secondary antibody (HRP anti-rabbit) to detect any false positive and to distinguish the bands of the proteins of interest from the bands corresponding to the antibodies bound to the beads. Only the heavy and light chains of antibodies incubated with the beads should be detected: a band around 55kDa for the heavy chain and a small one (sometimes not easy to detect) around 25kDa for the light chain. A band around 100 kDa is visible in some samples, as shown in Figure 29, which could correspond to the complex heavy and light chains or to a non-specific band. This figure also illustrates that anti-KLF2/KLF2A 3290 worked nicely, but surprisingly, not the IP using anti-H3K4me3. The “Input” immunoblot lanes are free from any bands probably because the sample is diluted.

The immunoblot was then incubated with anti-Klf2a 3290 and commercial anti-KLF2 antibodies: KLF2A protein could be detect in the inputs and in the IP samples.

A third incubation of the membrane with anti-H3K4me3 allows to attest the presence of H3K4 in the IP as well. This confirms the protocol was efficient to immune-precipitate the proteins of interest.

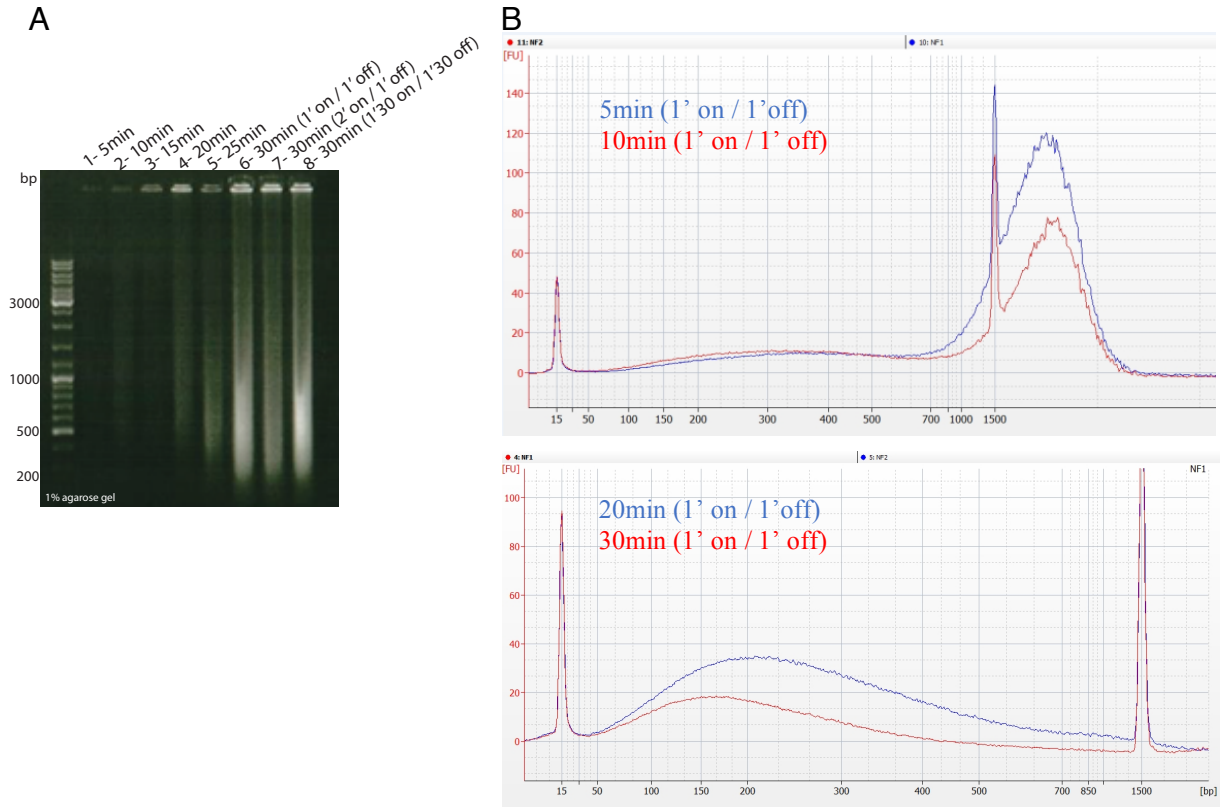


Figure 28: Optimization of the sonication step of the ChIP protocol.

Effects of different sonication times on chromatin fragment length. Chromatin was fragmented using Covaris E220 for the indicated times and conditions at 4°C. Panel A: Following crosslink removal, the DNA was extracted and analyzed on a 1%-agarose gel. 5ug of sonicated chromatin were loaded per lane (from 2000 48-hpf embryos).

Panel B: For four conditions the purified chromatin was analyzed with Bioanalyzer System (Agilent 2100, microchip Agilent High Sensitivity DNA Kit (reorder-no 5067-4626) and Agilent DNA 1000 Kit (reorder number 5067-1504)

Fragments above 600bp were considered under-sonicated and below 200pb over-sonicated.

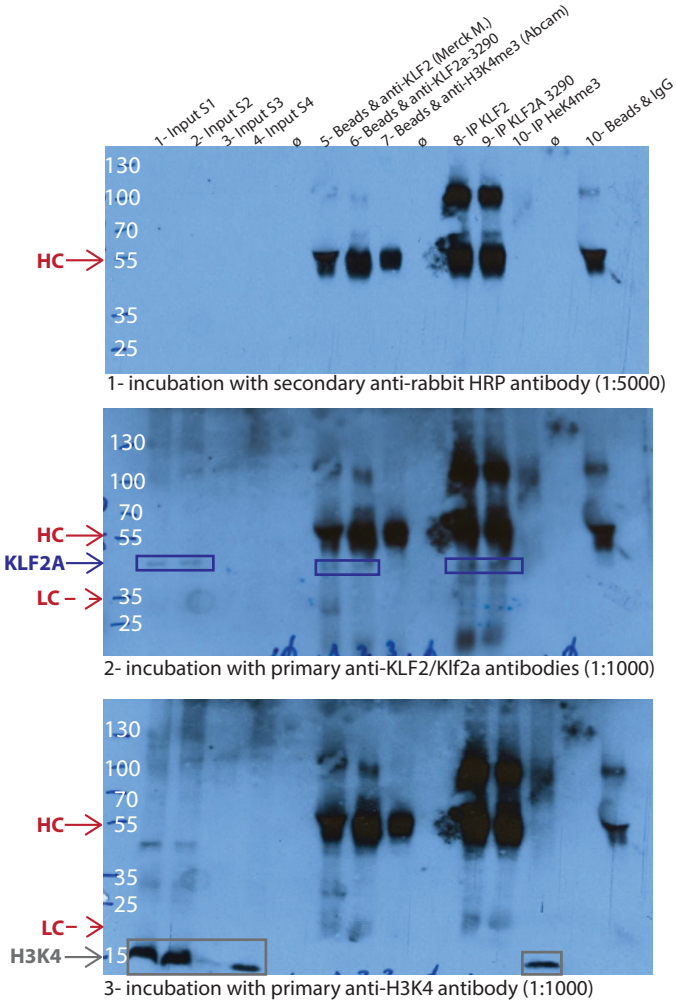


Figure 29: Test of the antibody efficiency to immune-precipitate Klf2a and H3K4 proteins in ChIP assay.

Immunoblots.

II.4. Preliminary results

Once these points were optimized, a first ChIP-qPCR was launched. IPs were performed with the commercial anti-KLF2 (Merck Millipore), the home-made polyclonal anti-Klf2A n°3290 and the anti-H3K4me3 (Merck Millipore, 07-473) as control.

II.4.a) ChIP-qPCR

2000 48hpf whole WT embryos were used and treated according to the optimized protocol. For the qPCR, putative *klf2a* target genes were tested. Among already validated targets of *klf2a/Klf2* in literature in endothelial cells, there are No synthase or *nos* genes (*nos1* and *nos2b*). (Wang et al., 2010) performed a ChIP-qPCR assay and showed Klf2a was bound to the promoters of *nos1* and *nos2b* genes *in vivo*, indicating a direct gene regulation. I was interested in recapitulating these data to validate the optimized ChIP protocol. Another gene was also tested in qPCR: *fibronectin 1b, fn1b,* a putative downstream target of *klf2a* - which was then validated as such by (Steed et al., 2016). Positive and negative primers were selected for each gene: a positive couple will amplify a fragment containing klf2a binding motif[9], and the negative couples not. Published primer couples were used for *nos1* gene (Wang et al., 2011) ; for *fn1b*, one putative Klf2a binding site (CACCC) was found inside *fn1b* proximal promoter and primers were designed according to its position, as illustrated on figure 31 (*primers sequences available in Table 6*). We analyzed two housekeeping genes, *glyceraldehyde-3-phosphate dehydrogenase (gapdh*) and S18 ribosomal protein (*rps18*) as positive controls for H3K4Me3 and H3K27ac for the qPCR (Kozera and Rapacz 2013) primers for these genes were gratefully provided by Dr Eirini Trompouki.

ChIP enrichment was calculated with the method “percent of input”. A standard curve per couple of primers was performed with the input sample and then used to calculate the enrichment of each immunoprecipitation (IP). Preliminary data (*not presented here*) presented an important background (the IgG read-out was high). A second ChIP-qPCR was

[9] *NB about klf2a binding motif:* KLK binding motifs are CG-rich sites of general structure CCN CNC CCN such as CACCC-boxes in various promoters and enhancers (Jiang et al. 2008). This motif is common to all KLF members and not only to Klf2a. Moreover, it is quite common in genome. I try to look for this motif not more than 1kb upstream of the transcription start site (TSS).

thus performed after modification of these points on the advice and experience of Dr Eirini Trompouki: addition of 1%-Triton x-100 just after sonication to obtain less stringent conditions, longer washes of the IP (15 min more), double amount of N-laurosarcosine in lysis buffer 3. These helps reducing the background. Figure 30 presents the data obtained.

Enrichment can be shown with each positive couple of primers for *fn1b*, *nos1* genes for IP with anti-Klf2a/KLF2 and for *gapdh* and *rps18* genes for IP with anti-H3K4 (Merck Millipore, 07-473). This will validate the ChIP assay efficiency and recapitulates data showing *klf2a* is regulating *nos1* gene and *fn1b*. Nevertheless, for some couples of negative primers, where no enrichment should be detected, a significant amplification can be detected. Moreover, the standard deviations are not acceptable, out of four qPCRs from four independent ChIPs, two gave non-optimal data explaining the discrepancy in the data obtained.

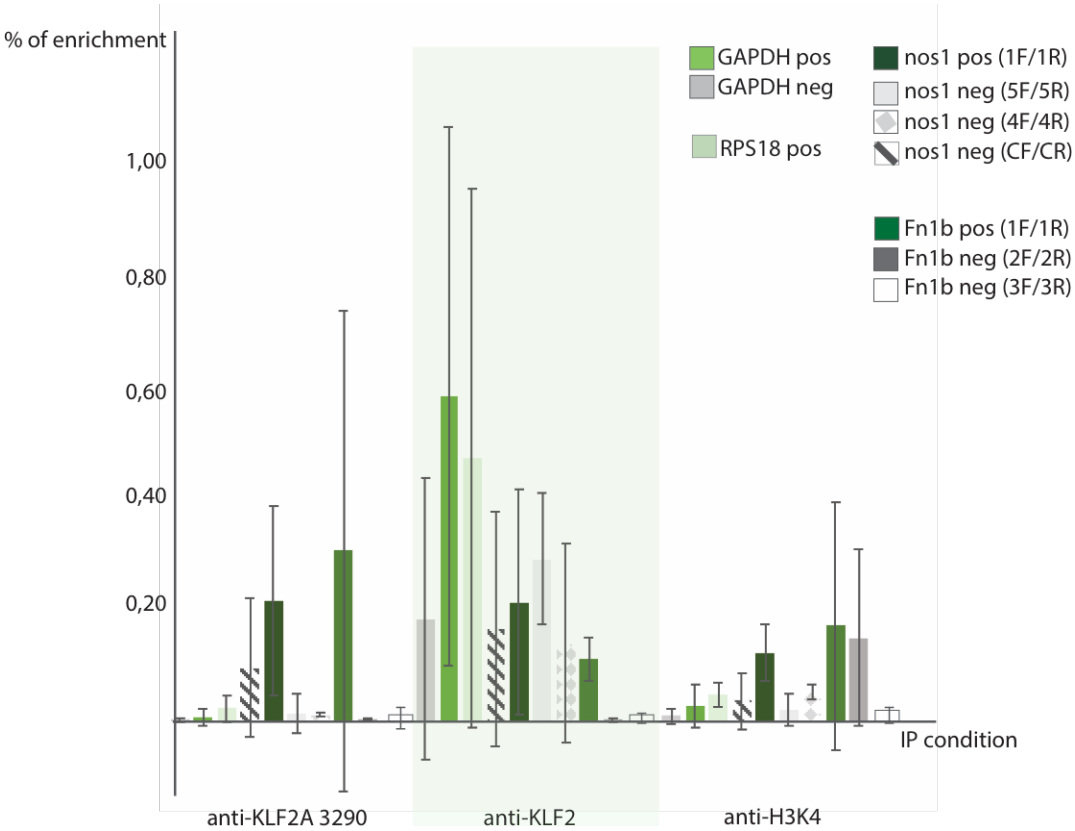


Figure 30: ChIP-qPCR

Quantitative RT-PCR analysis of *gapdh*, *rps18* (house-keeping control genes) and *nos1*, *fn1b* genes (*klf2a* downstream targets) expression in whole-embryo ChIP at 48hpf (mean of four replicate assays +SD).

Immunoprecipitation with anti-KLF2 (Merck Millipore), anti-Klf2a (home-made polyclonal antibody n°3290) and control antibody anti-H3K4 (Merck Millipore, 07-473).

Table 6: List of primers used for the ChIP-qPCR assay.

To validate each potential binding site, a couple of "positive primers" will amplify a fragment containing Klf2a binding motif (CACCC), these primers are specific to Klf2a binding site. Other couples are unspecific to it and called « negative primers ». Primers are designed within 1kb in the upstream region of each gene. *gapdh* and *rps18* are control genes for H3K4 assay. Nos primers were used by Wang et al. (2011).

Primer couple	Primer sequence	PCR fragment length
Fn1b_positive primers	1F: TCTTTGTGTGTGTGTGTGTGTG 1R: GCTCCAGATTTGCCTTCATTACTC	147pb
Fn1b_negative primers	2F: GTAATGAAGGCAAATCTGGAGC 2R: GCTCAGCATAAATGAGTACACC	359pb
Fn1b_negative primers	3F: ACTCTATTACCTGCGGGTTC 3R: TCCAGGAGAGAGAGAGAGAG	172pb
Fn1b_negative primers	4F: TTGCCAGATTTCCCGACAGTG 4R: TTCACCCGCGTTATGGAGAG	253pb
GAPDH_positive primers	FP1: TTGTGAAACTGCAACACCTC RP1: CTCAGATTGCGCTGGTGAAG	330pb
GAPDH_negative primers	FP2: TGTAGTTCACAGTAGCCTGTTG RP2: ATTGCGCTGGTGAAGCCTTTG	180pb
RPS18_positive primers	FP1: CATTTACAAACCCACCTCATCCTC RP1: CACGAACATTGATGGAAGACG	238pb
RPS18_negative primers	FP2: AGCTTTCCTCAGAACCACATG RP1: CACGAACATTGATGGAAGACG	178pb
RPS18_negative primers	FP3: GACACGAAGGATGTGCTGAAAC RP3: GGAGCCGTTACAATAATGTGC	170pb
Nos1_negative primers	Nos1_5F : GGCGTTGATTATGGGTGTAG Nos1_5R : GCCAGAGCACTAGGATTGT	191 pb
Nos1_negative primers	Nos1_CF : ATCCAGTTAAAGTCCTTCCC Nos1_CR : GGTTTAGGTTTTGGGTAGGATT	186 pb
Nos1_positive primers	Nos1_1F : GCAAACCACCTGTCATCGTC Nos1_1R : CTTCGTGGACCAGAGGGTTT	199 pb

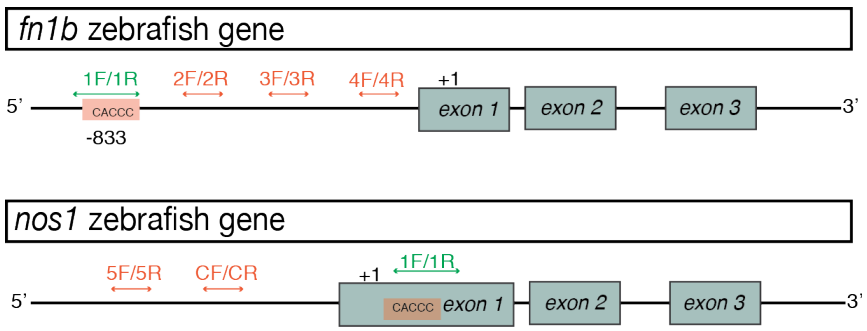


Figure 31: Positions of negative/positive couples of primers on *nos1* and *fn1b* zebrafish genes used in qPCR.

Nos1 primers were published by (Wang et al., 2011)

Table 7: General quality assessments for the samples sequenced in the ChIP-seq assay 1.

Sequencing performed with Illumina technology. Library prep: SQ00/SIL-01-SR, Read Length: SQ00/HS-1x50 — Hiseq Sequencing 1x50 bases standard.

In the column “raw reads”, it is estimated as good when the number of reads is equal or above 25 million. The "Aligned reads" column represents the number of reads aligned onto the reference genome. "%aligned" is the ratio (aligned reads / total reads).

Sample name	Concentration	Quantification method	Volume	Raw reads	Aligned reads	% aligned
input	96.6 ng/ul	Qubit	30uL	65,116,728	36,071,163	55.39
IP anti-H3K4me3	11.1 ng/ul	Qubit	18uL	73,740,824	41,623,425	56.45
IP anti-KLF2 (Merck Millipore)	6.7 ng/ul	Qubit	11uL	77,957,951	43,897,882	56.31

II.4.b) Results of the ChIP-seq assays

With the quite promising data obtained in ChIP-qPCR, we decided to redo a similar ChIP for immunoprecipitating Klf2a and H3K4 and to sequence the samples. Sequencing was performed using Illumina (Library prep: SQ00/SIL-01-SR, Read Length: SQ00/HS-1x50 — Hiseq Sequencing 1x50 bases standard).

Unfortunately, despite a good quality of the fragmented chromatin and sequencing (Table 7) the results were not satisfying: not enough coverage and an important background masking any interesting peak even for the IP with controls H3K4 antibody.

At 48hpf, *Klf2a* is mainly expressed in cardiac valve area and slightly in the tail vascular. While running a ChIP with whole embryos, we may "diluted" the few cells expressing the transcription factor among the different cell type populations. In order to enrich the samples with cells containing Klf2a and thus to maximize the chances to pull-down the transcription factor, it will be better to run a ChIP with endocardial cells and not with entire embryos.

FACS-sorting protocol was optimized to be able to collect these cells. Fli:gal4FF;UAS:kaede transgenic line was used to sort endocardial GFP positive cells. This requires to find an optimal lysis buffer which maintain cells alive during and after the sorting. Time of sorting was also a critical parameter and the quantity of starting material, its concentration should be optimized to be able to sort in 15 minutes maximum one sample.

A 48hpf zebrafish heart roughly contains around 200 cardiomyocytes (Bennett et al., 2013) meaning a ChIP on endocardial cells requires to collect more embryos (5000 to 10,000) to obtain enough cells after sorting. 10 FACS-sorts were required to reach 2 million cells, pulled together to obtain a sample for one ChIP condition. A second ChIP-seq assay using the cell population was run. IP was realized with anti-Klf2a n°3290 and another control antibody: anti-H3K27ac (Abcam, 4729), rather than H3K4me3 (Merck Millipore, 07-473).

Unfortunately, this second sequencing was not more successful as the previous one, even if again the quantity of chromatin and quality were nevertheless optimal. Figure 32 presents an example of data obtained on *egr1* gene. ATAC-seq on fli positive cells obtained in Dr Trompouki's lab as a comparison is included. No peak can be seen, just a baseline background.

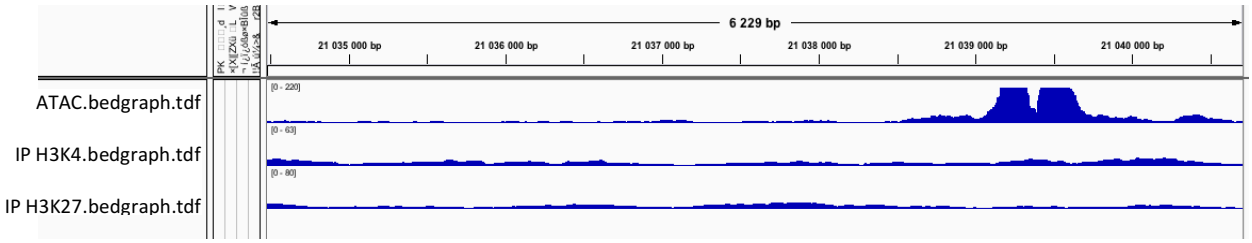


Figure 32: comparison of ChIP-seq data on *erg1* promoter

Data obtained on *egr1* promoter region (IPs using anti-H3K4me3 and anti-H3K27ac) with ATAC-seq obtained in Dr Trompouki's lab on the same type of cells (48hpf-endothelial cells).

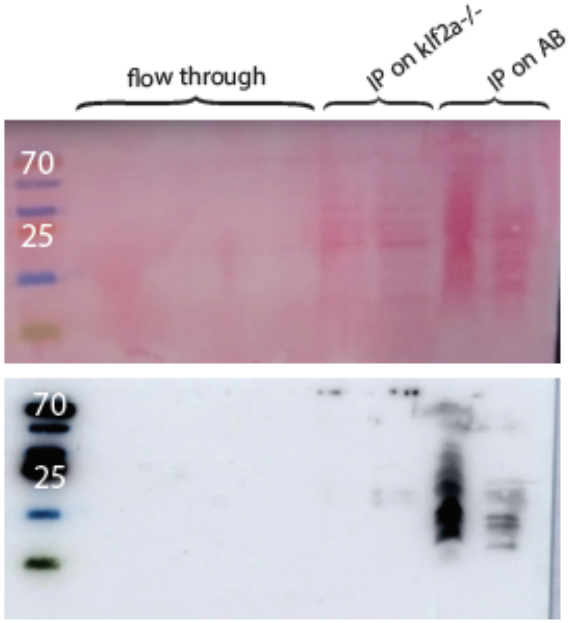


Figure 33: Immunoprecipitation of Klf2a protein

IP realized with 300 embryos WT (AB) and 300 embryos (klf2a^{ig4}) in 600uL lysis buffer. Incubation of the lysates with protein G-sepharose beads coated with anti-Klf2a (monoclonal 1KLF) and elution with 0.1M glycine pH2.8.

II.5. Discussion and perspectives

The low coverage and poor quality of the data could be explained in different ways.

II.5.a) The quality of the starting material

Removing the chorion using Pronase (Roche) could be detrimental to the efficiency of ChIP (Lindeman et al., 2010), notably by resulting in the degradation in modified histone epitopes such as H3K27me3. This will require to dechorionate manually around 5000-10,000 embryos per condition before starting the experiment, adding a considerable time-consuming step to the protocol. Moreover, the sorted and cross-linked cells were snap frozen in liquid nitrogen and stored at -80°C up to one month.

II.5.b) The cross-linking step

An excessive cross-linking reduces antigen accessibility and sonication efficiency, and epitopes may thus be masked. A shorter cross-linking step may help.

II.5.c) Antibody specificity and efficacy

It is also possible anti-Klf2a antibody is not efficient enough to pull down correctly the *in vivo* protein. The band on the immunoblot is indeed not consequent (Figure 29). Even if with H3K4 and H3K27ac, usual ChIP-grade antibodies, the results were not better. I tried to a protocol of immunoprecipitation of Klf2a with the purified monoclonal anti-Klf2a 1KLF antibody. Figure 33 attests this antibody is able to pull down the protein correctly. This means either there are not enough cells in the assay or enough protein to be detected.

This approach would have required more time to be developed completely. We decided to investigate another approach, an ATAC-seq assay coupled to mRNA-seq, which will be less time-consuming to optimize. They will bring us some information about the changes in chromatin accessibility in our mutants and we could thus infer some potential Klf2a and Egr1 downstream targets (direct or indirect, these methods are less precise to determine relations). Data obtained with this integrative analysis are presented in the draft of the paper which will be submitted to editors.

In parallel, we have also decided to generate zebrafish transgenic lines which will over-express Klf2a protein with a biotin-tag. It might be powerful to use it in a ChIP assay using anti-biotin antibody instead of anti-Klf2a antibodies.

III. Improvement: establishment of *in vivo* biotinylation tagging of Egr1 and Klf2a in zebrafish endocardial cells.

III.1. Presentation of the constructs designed

With the help of Dr Bernardo Reina San Marin and M. Vincent Heyer (IGBMC), two transgenic constructions were designed in order to express a biotin tag-version of Klf2a and Egr1 in endocardial cells of living fish. A schematic representation of the construction is presented in Figure 34.

The system is based on the bacterial BirA biotin ligase, which recognizes and biotinylates a short 23 amino acid peptide termed BioTag, allowing subsequent high stringency streptavidin affinity chromatography for purification of bait proteins, as illustrated on Figure 35. In this assay, the bait proteins will be Klf2a and Egr1. We combined biotinylation tagging with the versatile Gal4/UAS system for *klf2a* and *egr1* expression in myocardial cells. the approach has already been used to BioChIP–Chip applications to characterize the genome-wide distribution of histones, transcription factors, and chromatin-binding proteins (Lee et al., 2006) and (Mito et al., 2005).

The module cmlc2-eCFP will be a control for transgene. The transfection would be efficient if the myocardial cells present cyan fluorescence. Klf2a or Egr1 production will be followed by mCherry expression produced in parallel under UAS/gal4 system control (*presented in the general introduction of this thesis*).

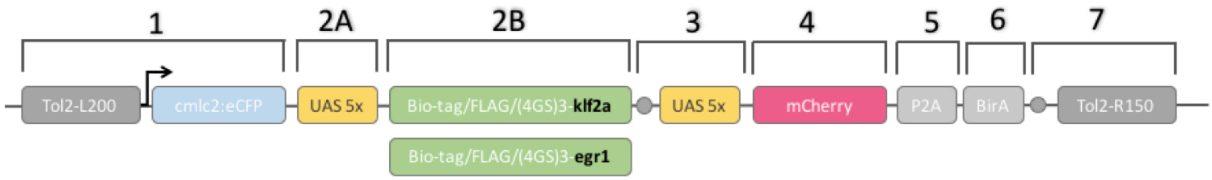


Figure 34: schematic representation of the transgenic constructions designed to over-express a biotin-tag version of Klf2a and Egr1 in endocardial cells of living fish.

7 modules called megaprimers were amplified by PCR using Q5 enzyme with addition of adapters containing BsaI restriction site and sequences for further insertion into pUC plasmid. Megawhoop cloning inside this vector was performed by PCR.

1: fusion of tol2-L200 – cmlc2 promoter and eCFP cassette
2A: fusion of SV40 and 5x UAS
2B: klf2a and egr1 sequences with tags
7: fusion of tol2-5150 and SV40 polyA.

Tol2-L200 and Tol2-R150: cis-sequences of the Tol2 transposon system
pA = polyA
P2A: fusion peptide, proteins linked by this peptide will be produced in the same time
BirA gene: code biotin ligase enzyme which regulates biotin biosynthesis
Bio-tag: 23 amino acid peptide which is recognized by BirA and undergoes biotinylation modification
FLAG: polypeptide protein tag, having the sequence motif DYKDDDDK (where D=aspartic acid, Y=tyrosine, and K=lysine). It is an artificial antigen to which specific, high affinity monoclonal antibodies have been developed and hence can be used for protein purification or used for locating proteins within living cells.
(4GS)3 = spacer between the tag and the sequence of the gene of interest

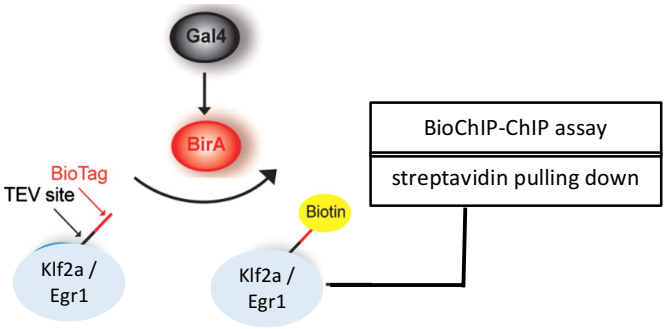


Figure 35: Presentation of birA biotinylation system used to create tagged proteins

(adapted from (Strubbe et al., 2011). Klf2a and Egr1 are fused at their C-terminus to a TEV-cleavable BioTag allowing proteolytic elution from streptavidin-coated beads. Gal4 induction of BirA biotin ligase expression leads to the biotinylation of the BioTag. Biotinylated Klf2a/Egr1 proteins are expressed in cells. Streptavidin pull-down will allow in future the immunoprecipitation of DNA-Klf2a/Egr1 complexes by ChIP using anti-biotin antibodies.

III.2. Construction protocol

The assembly of the different modules was realized using Golden Gate method (adapted from Miyazaki protocol, 2011) based on Megawhoop cloning.

This method allows the efficient, seamless and orderly assembly of multiple inserts into a final plasmid vector using the simultaneous activities of single type II S restriction enzyme (BsaI) (NEB) and T4 DNA ligase (NEB). Unlike standard Type II restriction enzymes like EcoRI and BamHI, enzymes from type S bind to their recognition sites but cut DNA outside at a downstream positional, not sequence-specific, cut site (*NEB website*). Therefore, they can create non-palindromic and unique overhangs. BsaI has a recognition site of GGTCTC(N1/ N5), where the GGTCTC represents the recognition/binding site, the N1/ N5 indicates the cut site is one base downstream on the top strand, and five bases downstream on the bottom strand. Ordered assembly of digested fragments is possible through annealing of complementary four base overhangs on adjacent fragments different from each other. The digested fragments and the final assembly no longer contain Type IIS restriction enzyme recognition sites, so no further cutting is possible. Figure 36 illustrates the principles of this method.

The different sequences to assemble were first PCR-amplified using Q5 enzyme (NEB) to form so-called "megaprimers". Some megaprimers are composed by the fusion of different fragments. The sequences used are listed on Table 8.

Adapters containing BsaI restriction sites and pUC-sequences of insertion were added to the extremities. Cloning of modules 2A and 3 containing UAS promoter were not straightforward, as UAS is composed of repetitive sequences. After several unsuccessful attempts to PCR-amplify five repetitions in one final fragment, the complete sequence was ordered as an ultramer at IDT. Regarding the fragment containing *egr1* genomic sequence, additional PCRs were required to mutate BsaI restriction site, which will have impaired the module assembly.

Megaprimers were then purified using NucleoSpin columns (Gel and PCR clean-up, MN) before to be inserted into pUC plasmid by a second PCR with Q5 (megawhoop reaction). Plasmid were then digested one hour 37°C-digestion by Dpn1 (NEB) and transformed into TOP10 competent cells (Sigma) by heat shock. Bacteria were plated on LB supplemented with

ampicillin and Xgal for selection. Mini-prep from white colonies were performed and their sequences verified by sequencing (GATC/Eurofins).

The final assembly was realized into pXpA vector (unpublished data, kindly provided by Dr Reina San Marin's lab, IGBMC) in one reaction (digestion by BsaI and ligation by T4 DNA ligase (NEB)) as described on figure 34. Final plasmids were checked to be error-prone by sequencing (Table 10 presents the primers used). For each construction one clone was selected: pXpA_633 mini prep n°6 for *klf2a* construction and pXpA_634 mini prep n°9 for *egr1* construction.

III.3. Fish mutant generation

To deliver the construction to all embryo cells, one-cell stage WT (AB) embryos were injected with 10ng/uL of pXpA-633/634 plasmid and 12ng/uL of mRNA tol2 (pCS2+ tol2 plasmid from Kawakami) diluted in RNAse-free water. Expression of eCFP under cmlc2 promoter was used to control the efficiency of transfection. CFP-positive embryos were raised to adulthood. A first out-crossing of this F0 generation with AB fish was performed to raise a stable line.

III.4. Preliminary results and perspectives

CFP-positive F1 fish were then crossed with the *fli:gal4FF;UAS:kaede* line. To confirm the overexpression of Egr1 and Klf2a in the offspring, we measured the expression of the mcherry reporter. This is part of an on-going work, data will not be shown in this thesis. We did detect mcherry fluorescence in CFP and GFP positive embryos from 48hpf, but I still have to validate the expected over-expression in valve areas. CFP and GFP fluorescences were detectable under regular fluorescent macroscope, but not mCherry fluorescence. Therefore, mCherry fluorescence signal was assessed using a spinning disk confocal microscope at 48, 96 and 120hpf.

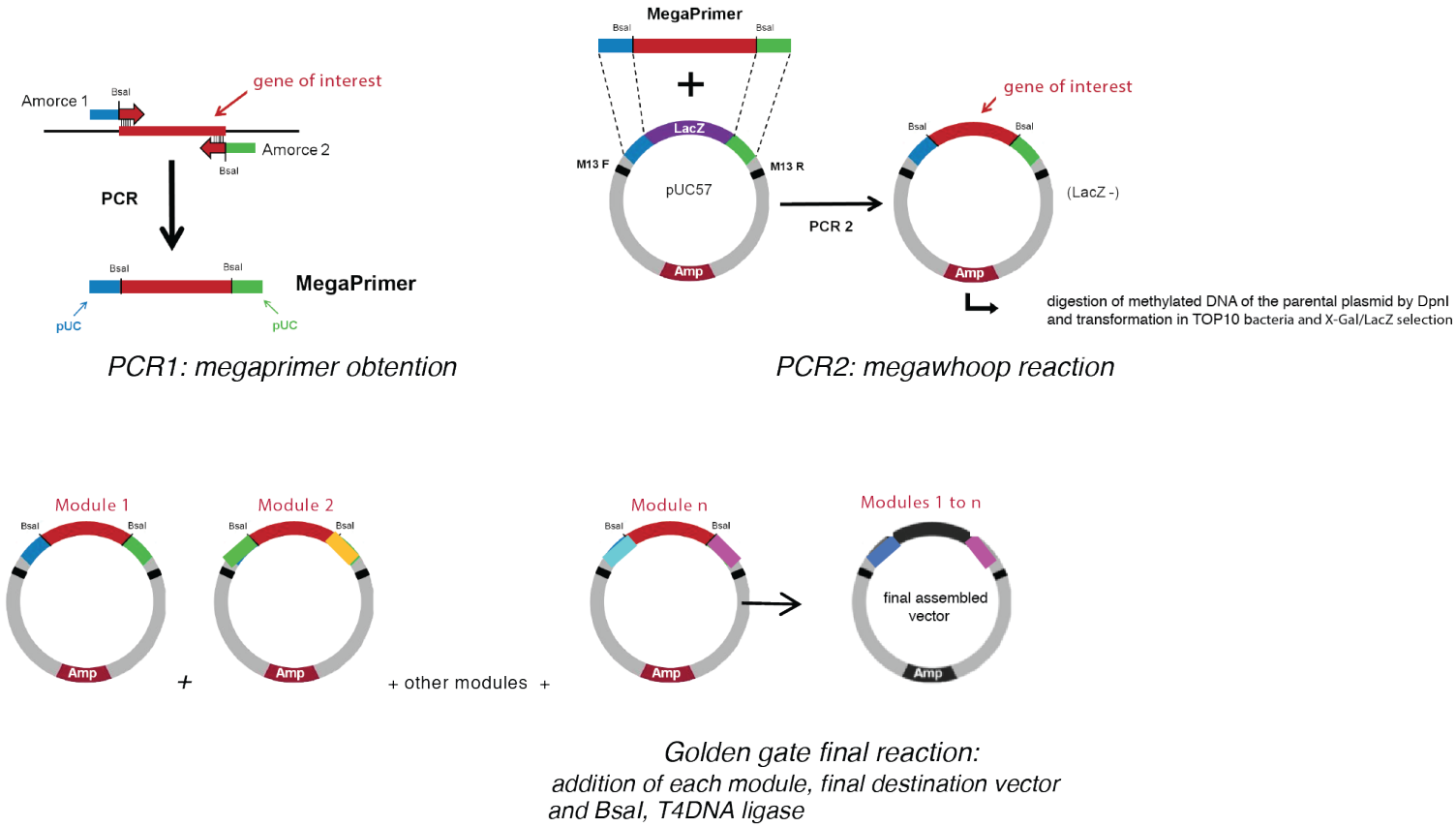


Figure 36: Megawhoop Golden Gate cloning

(figure adapted from Dr Reina San Marin, IGBMC)

1- A first PCR is performed to amplify the fragment of interest with primers. Adapters with BsaI site and sequences for futher cloning into pUC plasmid are contained in the primers.
2- A second PCR using the product of the first PCR (the megaprimer) is performed with the target plasmid (pUC57) as template. This second PCR is digested with Dpn1 enzyme to remove the parental plasmid. Purified plasmid is then transformed in E. coli TOP10.
3- A single tube final reaction allows BsaI digestion of the megaprimers from pUC and of the destination plasmid, and T4 DNA ligase orderly and seamless assembly (BsaI sites are absent of the final construction)

Table 8: description of the module (megaprimers) used to design egr1 and klf2a transgenic constructions using Golden Gate Megawhoop approach.

Megaprimer name	Sequences fused	Origin and information
1: m573	Tol2 L200	pDestol2CG2 gateway vector
	cmlc2 promoter - Kozak	pDestol2CG2 gateway vector
	eCFP	gratefully provided by Dr San Martin (IGBMC) *This will be a control for the transgenesis and efficient transfection. Cardiac cells should be CFP positive.*
2A and 3: m574 and m577	SV40 poly A	from pDesTol2CG2 gateway vector
	5X-UAS	Complete sequence ordered as an ultramer at IDT *Five repetitions were shown to be optimal to reduce the silencing effects observed with increasing generation numbers (Akitake et al. 2011)*
2B *klf2a: m575* *egr1: m576*	Bio-tag/FLAG/(4GS)3	gratefully provided by Dr San Martin (IGBMC)
	Egr1	cDNA transcribed by Jean-Marie Garnier (IGBMC). BsaI sites were mutated by PCR prevent impairment during the final assembly. NotI and EcoRI sites replaced them.
	Klf2a	From pExpress plasmid available in the lab
4	mcherry	gratefully provided by Dr San Martin (IGBMC)
5	P2A	gratefully provided by Dr San Martin (IGBMC) *fusion peptide, 2 proteins produced in the same time : control of Klf2a production visible by fluorescence (mcherry).*
6	BirA - Stop	gratefully provided by Dr San Martin (IGBMC)
7 (m578)	SV40 polyA & Tol2 R150	pDestol2CG2 gateway vector

Table 9: Primers used for the design of the transgenic constructions

Primer name	Sequence
m573-Fwd1	CGAATGCATCTAGATATCGGATCCCGGTCTCGATCCCAGAGGTGTAAAGTACTTGAGTAATTTTACTTG
m573-Fwd2	CAGTCAAAAAGTACTTATTTTTTGGAGATCACTTAAAGCTTAAATCAGTTGTGTTAAATAAGAG
m573-Fwd3	CCAGACCAACAGCAAAGCAGACAGTGACCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC
m573-Rev1	CTCTTATTTAACACAACTGATTTAAGCTTTAAGTGATCTCCAAAAAATAAGTACTTTTTGACTG
m573-Rev2	GGTGAACAGCTCCTCGCCCTTGCTCACCATGGTGGCGGTCACTGTCTGCTTTGCTGTTGGTCTGG
m573-Rev3	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTCATATCACTTGTACAGCTCGTCCATGCCG
m574/577-Fwd	AGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTCATCAATGTATCTTAAC
m574/577-rev	CGTTAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCT
m574-Rev2	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTCACCGTGTGGAGGAGCTCAAAGTGAGGCTGAGACGCGATGG
m575-Fwd1	CGAATGCATCTAGATATCGGATCCCGGTCTCGGGTGGCCACCATGGGCCTGAATGACATCTTTGAGGCCCAGAAGATCG
m575-Fwd2	GACATCTTTGAGGCCCAGAAGATCGAGTGGCATGAGGGAGGAATGGACTACAAGGACGACGATGACAAGGGTGGCGGAGGGAGTGG
m575-Fwd3	GATGACAAGGGTGGCGGAGGGAGTGGAGGCGGTGGCAGCGGTGGCGGAGGGAGTATGGCTTTGAGTGGAACGATTTTACC
m575-Rev	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTGTTCCTACATATGACGTTTCATATGAAGGGC
m576-Fwd1	CGAATGCATCTAGATATCGGATCCCGGTCTCGGGTGGCCACCATGGCTGCAGCCAAGACAGAGATGC
m576-Fwd2	CACGAGCGCCCCTACGCATGCCCTGTGGAAACCTGTGACAGGCGCTTCTCACGCTCAGACG
m576-Fwd3	CCTCTCCCATCACTTCTTACCCCTCTCCGGTTTCCTCTTTCCCGTCTCCAGTCAACTCCTGC
m576-Rev1	CGTCTGAGCGTGAGAAGCGCCTGTCACAGGTTTCCACAGGGCATGCGTAGGGGCGCTCGTG
m576-Rev2	GCAGGAGTTGACTGGAGACGGGAAAGAGGAAACCGGAGAGGGGTAAGAAGTGATGGGAGAGG
m576-Rev3	TTGTAGTCCATACTCCCTCCGCCACCGCTGCCACCGCCTCCACTCCCTCCGCCACCGCAGATGTCGGCTGTCCGAGGGGAAAGCG
m576-Rev4	CGATCTTCTGGGCCTCAAAGATGTCATTCAGGCCCATTCCTCCCTTGTCATCGTCGTCCTTGTAGTCCATACTCCCTCCGCCACC
m576-Rev5	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTGTTCTCACTCATGCCACTCGATCTTCTGGGCCTCAAAGATGTCATTCAGGCC
m577-Rev2	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTTGCGGTGTGGAGGAGCTCAAAGTGAGGCTGAGACGCGATGG
m578-Fwd1	CGAATGCATCTAGATATCGGATCCCGGTCTCGGCTTTGATCATAATCAGCCATACCAC
m578-Fwd2	GTCCAAACTCATCAATGTATCTTAAAATACTCAAGTACAATTTTAATGGAG
m578-Rev1	CTCCATTAAAATTGTACTTGAGTATTTTAAGATACATTGATGAGTTTGGAC
m578-Rev2	GCAGGCCTCTGCAGTCGACGGGCCCGGTCTCTATGGCAGAGGTGTAAAAAGTACTCAAAAATTTTACTC

Ultramer sequence (4X UAS)

ccaaactcatcaatgtatcttaaCGGAGTACTGTCCTCCGgcaaggtCGGAGTACTGTCCTCCGacactagaggtCGGAGTACTGTCCTC
CGacgcaaggCGGAGTACTGTCCTCCGggctggCGGAGTACTGTCCTCCGgcaagggtcgactctagagggtatataatggatcccatcgcgt
ctcagcctcactttgag

Table 10: primers used for sequencing the final pXpA-633/634 constructions

Part to sequence	Primer sequence
first UAS repetitions (cassette 2A)	CMLC2_FP : GTCAGAACCTGCAGTGTTGGC
second UAS repetitions (cassette 3)	P2A_RP : CAGGCTGAAGTTAGTAGCTC
End of *klf2a* genomic sequence	klf2aend_FP : CTGGACGCCAAACCAAAGAG
End of *egr1* genomic sequence	egr1end_FP : CTGACGCCGCTGCACACCAT

IV. Additional information about the ATAC-seq / mRNA-seq combined approach

IV.1. Discussion about quantity of material required, choice of the kit and protocols used for ATAC-seq and mRNAseq

IV.1.a) Optimization of a FACS sorting protocol

FACS-sorting required some improvement to be able to sort a good ratio of positive and viable cells from a given sample in a minimum amount of time. FACSmax buffer (Fisher Scientific) used previously in the lab was abandoned for the following extraction methods:

- for whole embryos: extraction was performed in PBS-2%-Liberase (Roche) for 15min at 37°C, 5%-FBS was then added to stop the reaction. The solution was pipetting up and down thoroughly before being filtered on ice on a 40µm-nylon mesh and finally resuspended into FACS buffer (0.02% FBS; 0,01% Penicillin/streptomycin; 2% EDTA)

- for extracted hearts: after the extractions, hearts were centrifuged rapidly and cells resuspended into FACS buffer (500 hearts in 500µL to ensure an optimal concentration of cells to reduce the time sorting).

IV.1.b) ATAC-seq protocol

The amount of input material was a challenging issue in ChIP. We leveraged the recent development of the assay for transposase-accessible chromatin with high-throughput sequencing (ATAC-seq), which requires only roughly 50,000 cells as input material (Buenrostro et al., 2013). We performed several FACS sorting of endothelial cells from dissected hearts to reach this number of cells. No protocol optimization was required here, we followed the protocol published by Buenrostro et al. (2015).

IV.1.b) Collecting RNA from 48hpf-sorted cells

Preparing the mRNA-seq samples was more challenging for two reasons. First, the number of cells used for an ATAC-seq sample was not enough to obtain a quantitative and qualitative RNA amount after extraction. Second, when studying dynamic biological processes, such as developmental steps, it is critical to work fast. The changes in gene expression are rapidly happening in live cells. In order to recapitulate the whole regulatory network involved at a

precise time point, in the most accurate way, we tried to pay attention to the experiment duration.

FACS-sorting optimized to collect cells in maximum 15 minutes allowed to limit the total technical process of less than one hour. This should ensure to have reproducible data reflecting at most the cellular processes happening at 48hpf in live cells. Total RNA purification followed cell isolation as quickly as possible using Zimo RNA extraction kit.

Unfortunately, the quality of the RNA and subsequent cDNA was not optimal for every replicate. We could also notice the presence of an unexpected peak around 550-600pb on the cDNA quality profile, as illustrated on Figure 37. This was supposed to be a contaminant globin peak. But even after some sorting of endothelial cells from *gata1* / *vlad tepes* mutant which lack blood cells, the peak was still present. From one sample to another its proportion varies but it could still be detected.

We can remark that surprisingly, hemoglobin genes were the most abundant RNA population in the mRNAseq data we obtained. The unexpected peak could be attribute to these genes.

To finally obtain high-quality RNA in amount required for RNA-seq, we decided to use a protocol for small RNA quantity (Clontech SMART-Seq v4 Ultra Low Input RNA Kit). 1000 cells were FACS-sorted and immediately frozen. Once all technical replicates for mutants and respective controls were collected, cells were defrosted and direct cDNA synthesis from intact cells was performed.

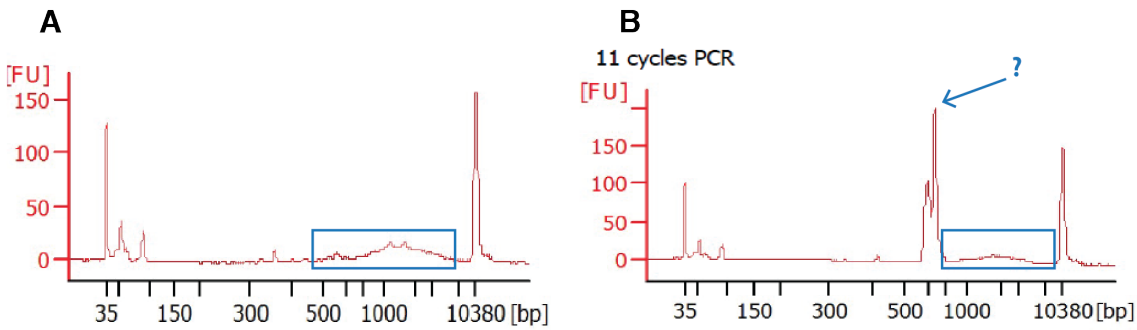


Figure 37: cDNA quality profiles. Analysis performed with Bioanalyser.

Panel A presents a theoretical expected profile and panel B the profile obtained with our sample. The blue arrown points out the unexpected peak obtained around 550-600pb.

IV.2. Complementary information not included in the paper about the mRNA and ATACseq data sets

IV.2.a) Data on the double mutant *klf2a;klf2b*

mRNA seq and ATACseq were also performed on *klf2a-klf2b* double mutants. The data were not included in the paper we would like to submit to editors. They need to be analyzed further down.

We were surprised this mutant do not share many genes in common with the single respective mutants, indeed only 80 genes seem in common between these three mutants. *Has2*, *wnt9b* for example which were both down-regulated in the single mutants were not deregulated in the double mutant analysis. The GO-terms of the deregulated genes in mRNA-seq and of differential peaks found in ATAC-seq were however quite interesting, revealing genes mainly involved in embryogenic morphogenesis, heart development and regulation of transcription (Figure 38, panel C and D).

How these disparities could be explained? it is possible the double knock-out has induced other effects which are not the addition of the effects observed for the single mutants. Some compensation may occur in each single mutant. On the contrary in the double mutant, where *klf2* gene is completely absent, the mechanism of compensation could be different. These data sets will be analyzed more carefully (*on going work*).

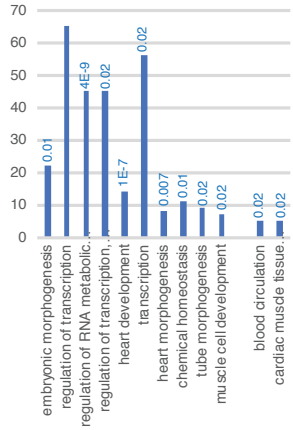
A

Peaks absent in *klf2a*^{ig4}	*Flt1* *wnt11, wnt2ba, wtn2bb* *klf2b, klf2a, klf11,* *notch1a, 1b, 3* *gata2b* *egr3* *bmp3, bmp6 bmp10* *RERG* *Egfl6* *Vcam1a* sox19a, 19b, 6, 4a, 11b, 5	new peaks in *klf2a*^{ig4}	*wnt4b* *egr1* *klf6a, 1, 8, 7b, 3, 11b* *ets2* *ugdh* *sox4b, sox21a*
Peaks absent in *egr1*^{sa64}	*flt1* *klf2a, 4, 17, 6b, 1, 12* *notch1a, 1b, 3* *wnt11, 2ba, 2, 2bb* *egr3* *piezo1* *gata5, gata2b* *tead1a, 1b* *RERG* *egfl6* *v-cam1a* sox19a, 19b, 6, 4a, 11b, 5	new peaks in *egr1*^{sa64}	*klf2b, 11b* *Ets2* *Tead1a*
new peaks in *klf2a;klf2b*	*egr1, egr2b* *klf1, 3, 4, 2a, 2b, 7a, 7b, 8, 13 and 15* *notch1a, 1b, 3* *hes6* *egfl6 and 7* *wnt2ba, 2bb, 8a, 4b, 11* *flt1* *gata5, gata2a, gata2b* *tbx18, 3a, tbx20* *tead1a, 1b*	new peaks in *klf2a;klf2b*	*gata3, 4, 1a,* *gata5, gata2a* *klf2b, 2a, 6, 13, 9, 1, 7a, 3, 17* *sox6, 17, 4b, 4a, 21a* *bmp6, 15* *pecam1* *wnt6a, 6b, 9, 4b, 6b* *tead4, 3*

B

up	*tbx20* *sox17* *sox6, 2* *nFatc1* *wnt5a*	down	*egr3, 4* *gata4* *tbx6* *wnt4b, 16* *hes2.2*

C



D

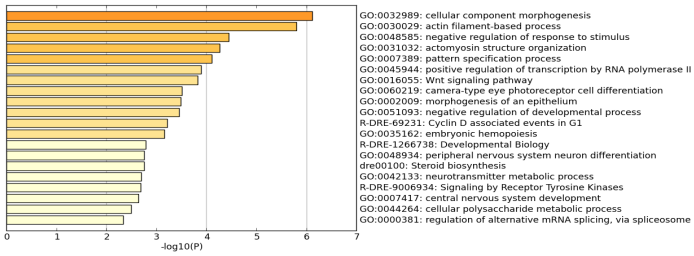


Figure 38: Genome-wide data obtained for the double mutant *klf2a-klf2b*

Panel A: ATAC-seq comparison between mutants, showing interesting deregulated peaks.
Panel B: interesting genes deregulated in the mRNA seq
Panel C: mRNA-seq GO-terms analysis based on cellular compartment terms of upregulated genes highlighted a significant enrichment of ECM-related terms. P values describing the significance of each term enrichment are shown in blue.
Panel D: GO-terms analysis of differential ATAC peaks found in the double mutant compared to control siblings.

IV.2.b) Some disparities with ISH data

An WISH was performed with anti-Has2 probe and contrary to what was expected regarding the mRNA seq and the qPCR data, no change in expression was observed between mutants and their respective controls (Figure 39). Another disparity was observed with *egr1* probe, in *klf2* mutants the down-regulation was total in almost 80% of embryos. However, in mRNAseq and qPCR the down-regulation was not significant for every *klf2* mutant.

WISH is indeed not a quantitative technique. The staining is not always linear. qPCR and mRNA-seq assays would be more accurate to estimate mRNA levels. WISH can detect differences in the distribution of expression and could give an idea of the deregulation. But if the down-regulation is not total, some transcripts are left in cells, and if the staining is too long, it may happen the saturation is reached and the signal is equivalent to the control signal. This does not explain nevertheless the result obtained with *egr1* probe.

Moreover, these results also raise the question of the experiment timing: we tried to perform experiment precisely at 48hpf for WISH, imaging, mRNA-seq, ATAC-seq and qPCR. But as the processing for genome-wide data (embryos processing, heart extraction, FACS-sorting and DNA/RNA processing) is a bit longer we cannot be 100% sure. It would be interesting to perform some qPCR at different close time-points (46, 47, 48, 49 and 50hpf) to study the expression of the gene of interest. The time window were cushions and ECM remodeling happened is highly dynamic and significant changes may occur quickly.

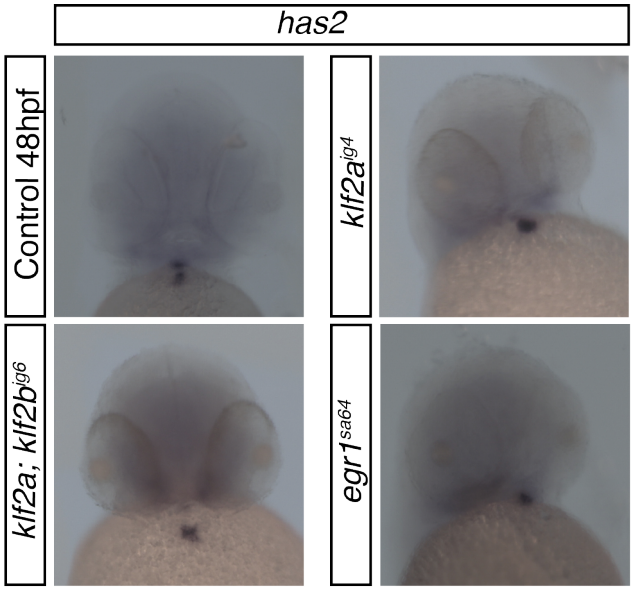


Figure 39: *has2* WISH on 48hpf-zebrafish embryos

has2 expression is restricted to AVC and OFT area at 58hpf. No change in expression was observed between control and klf, egr1 mutants, contrary to the down-regulation observed in mRNAseq and qPCR for this gene.

Chapter 5

Design of CRISPR/cas9 system for generating KO-*fn1b* zebrafish mutant

In the past morpholinos (MO)-based KI were an accepted alternative to genomic lesions, but there were reports of strong phenotypes that are a result of clearly off-target inhibition (Kok et al. 2014; Law and Sargent 2014). Kok et al. (2014) and Stainier et al. (2015) also discovered that 80% of gene knockouts did not recapitulate the morpholino phenotype reported for the same gene. This is a significant setback to researchers interested in studying gene function to find a way to generate proper KO mutants.

To date, experience from the zebrafish community and from the hands of different collaborator labs give advantages to the CRIPSR-Cas9 approach in term of KO efficiency and lack of off targets compared to the use of TALENs system. For example, Chang et al. (2013) targeted *etsrp* and *gata5* loci with CRISPRs and managed to recapitulate the vasculature phenotypes and the cardiac bifida phenotype[10] observed previously with morphants[11] with these "CRISPR-KO" mutants.

For the project of Dr Emily Steed *(results published in Steed et al., 2016 – paper available in Annex)*, we aimed to generate zebrafish *fn1b* mutants using CRISPRs for some validation experiments of results obtained with *fn1b* morphants. When the project started, no viable mutant was available in the worldwide database and only fn1b-Morpholino (MO) was used for studies on this gene. This will be a starting point to optimize a method to generate KO-zebrafish mutant using CRISPRs and could be implement to other genes of interest in the future.

V.1. Presentation of CRISPR-Cas9 system

CRISPR-associated protein 9 (Cas9) is an RNA guided DNA endonuclease enzyme associated with the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) adaptive immunity system in Streptococcus pyogenes. This bacterium uses Cas9 protein to memorize and later to interrogate and cleave any foreign DNA, such as invading bacteriophage or plasmid DNA, complementary to the 20bp spacer region of the guide RNA. This system has some analogy with the RNA interference (RNAi) mechanism in eukaryotes.

[10] Failure of the bilateral myocardial cells to fuse into a single central heart tube resulting in the presence of two independent hearts

[11] Name given to an organism treated with a morpholino antisense oligo to temporarily knock down expression of a targeted gene

CRISPR represents a family of DNA non-contiguous direct repeats separated by variable sequence spacers, they are found in most archaeal (~90%) and bacterial (~40%) genomes. Spacer mostly correspond to segments of captured plasmid or virus. The size of CRISPR repeats and spacers varies between 23 to 47 bp and 21 to 72 bp, respectively. CRISPR are often adjacent to CRISPR-associated genes (cas). Cas genes encode a heterogeneous family of proteins that carry functional domains typical of nucleases, helicases, polymerases, and polynucleotide-binding proteins. (From Horvath et. Barrangou, 2010.)

Three classes of CRISPR/Cas systems have been described so far. Whereas Type I and Type III systems require multiple distinct effectors Cas proteins acting in a complex, type II utilizes a single effector nuclease: Cas9 (Makarova et al., 2011). The CRISPR-Cas9 system is based on the recognizing by a Cas complex o foreign DNA and subsequent integration of exogenous DNA fragments into the repeat-spacer unit at the leader end of the CRISPR loci (Figure 40). The integrated sequence is called a protospacer and is followed by a short stretch of conserved nucleotides, the protospacer adjacent motif or PAM. PAM is a component of the invading virus or plasmid, but is not a component of the bacterial CRISPR locus. The immunity process relies on the subsequent transcription of the CRISPR repeat-spacer array into pre-crRNA and its processing into mature short CRISPR RNAs (crRNAs). This crRNA then forms a RNA duplex with a trans-activating crRNA (tracrRNA) which acts as a guide for the endonuclease Cas9 to direct sequence-specific silencing of the corresponding invading nucleic acid (Figure 40).

This ingenious system has been adapted to create 'molecular scissors' to induce DNA direct double stranded beaks, and are widely used in genome editing. Gene knockout (KO) using CRISPR "molecular scissors" technology is accomplished by Cas9-mediated double-stranded (ds)-DNA breaks. Following a Cas9-cut, the error-prone natural repair mechanism of non-homologous end joining (NHEJ) often leads to the generation of indels (insertion and/or deletions) and thus frameshifts disrupting the protein-coding capacity of a locus.

For editing genes, synthetic and customizable single guideRNAs (gRNAs) are synthesized to perform the function of the tracrRNA:crRNA complex in recognizing gene sequences having a PAM sequence at the 3'-end. The canonical PAM recognized by Cas9 is the sequence 5'-NGG-3' where "N" is any nucleobase followed by two guanines. This site occurs once in every 128 bp of random DNA sequence (Hwang et al. 2012). Potential target sites are

both [5'-20nt-NGG] and [5'-CCN-20nt], as it is equally efficacious to target the coding or non-coding strand of DNA.

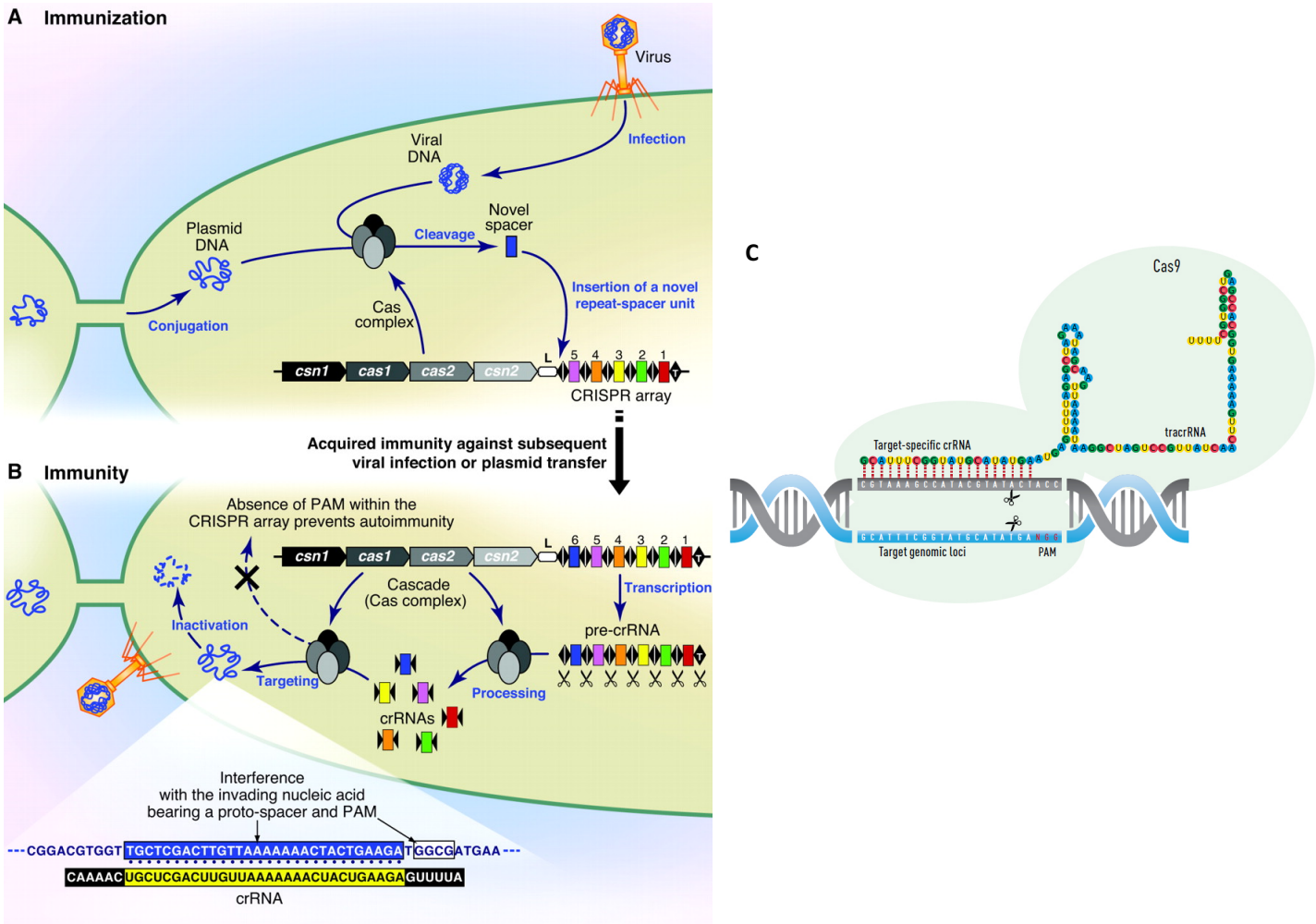


Figure 40: Overview of the CRISPR/Cas system

Mechanism of action in bacteria (*from Horvath and Barrangou, 2010*)

(**A**) Immunization process: After insertion of exogenous viral or plasmid DNA, a complex of Cas proteins recognizes this foreign DNA and integrates it as a novel repeat-spacer unit the protospacer, followed by a PAM sequence, at the leader end of the CRISPR locus.

(**B**) Immunity process: The CRISPR repeat-spacer array is transcribed into a pre-crRNA that is processed into mature crRNAs used as a guide by a Cas complex to interfere with the corresponding invading nucleic acid (for the type II CRISPR/Cas9, crRNA complexes with tracrRNA and this complexe is recognized by Cas9 nuclease).

Repeats are represented as diamonds, spacers as rectangles, and the CRISPR leader is labeled L.

(C) Schematic representation of the customizable complex formed *in vitro* between Cas9 protein, the gRNA (with the tracrRNA scaffold recognized by Cas9) and targeted genome locus (from ThermoFisher Scientific).

V.2. Design of a CRISPR/Cas9 approach to target *fn1b* zebrafish gene

V.2.a) Identification and customization of gRNA

SgRNA:Cas9-targetable sites were generated from a comparison between different web-based targeter CRISPR programs: ZiFiT Targeter program14 (hosted by the Zinc Finger Consortium, Hwang et al., Mali et al. 2013), CRISPOR (Haeussler et al., 2016) and CRISPRdirect (developed by Naito et al., DBCLS, Tokyo), CRISPRScan tool (Yale University, Giraldez lab, Moreno-Mateos et al. 2015). They provide a list of customizable sgRNA bearing 20 nucleotides (nts) of sequence complementary to a target site of interest. Constraints on the range of targetable sequences are due to sequence requirements imposed by the T7 promoter used to make sgRNAs (GG at the 5' end of the transcript[12]) and by the requirement for PAM sequence in genomic DNA just 3' to the target site.

In this project, exon 1 and exon 3 of *fn1b* gene were targeted. Targeting 5'-constitutively expressed early exons is preferred when loss-of-function mutant is required (Hwang et al., 2013 ; Meier et al., 2017). Exons 1 and 3 are also present in all splice variants and different from *fn1a* gene. It should reduce the chances that the targeted region will be removed from the mRNA due to alternative splicing. Exons near the N-terminus are also often targeted since frameshift mutations here will increase the likelihood that a nonfunctional protein product is produced. According to our collaborators experience, every sgRNA do not work efficiently in zebrafish. It is better to design a minimum of four independent sgRNA per gene (some even use up to 10 different gRNA targeting the same region).

gRNAs obtained by correlation of these different program, with the best score were chosen (GC content below 55% ; CRISPRscan score of at least 55, ideally >70 (Moreno-Mateos et al., 2015). A Blast research confirmed the designed sequence does not exist at any other location in the genome was performed. Selected *fn1b*-gRNAs are listed in Table 10.

To ensure no error on the available genome version could affect the targeting, before injecting gRNA in WT AB embryos, *fn1b* WT genomic sequence was verified in five AB adult parent fish by extracting genomic DNA from their tail. PCR was realized using primers reported

[12] "For *Streptococcus pyogenes* Cas9, the PAM sequence is NGG, leading to a consensus CRISPR target site of N21GG. For sgRNAs produced *in vitro* using a T7 promoter the first two transcribed bases are G, making sgRNAs with a GGN19GG consensus" (Hwang et al., 2013)

in table 11. Fragments were sequenced (GATC/Eurofins) and aligned with zebrafish 10 version. No error was reported.

"Off-targets analysis" should also be performed to assess the targeting specificity of the designed gRNA/Cas9 system, and check the nuclease did not generate unwanted cuttings at other loci in the genome. Nevertheless, off-targets events were shown to be rare and of minimal concern with zebrafish compared to the off-target rates observed in cell culture for example, certainly due to the duration of Cas9/sgRNA activity in faster dividing cells or segregation of unlinked loci during meiosis (Varshney et al., 2015).

A first analysis was performed while designing the gRNAs, to avoid the choice of putatively prone-off target sequences. Previous studies have shown "that 8–12 bases ("seed sequence") immediately to the 5' end of the PAM, NGG, are critical in determining target specificity: single mismatches within this region largely abolish the gRNA/Cas9 target activity, whereas single mismatches beyond can be tolerated" (Jiang et al., 2012; Cong et al., 2013). Potential off-target sites were searched in this way in the genome by Blast analysis and using the browser "CRISPRs" track on the ZebrafishGenomics track hub (which computationally predicts 18,367,469 CRISPR 20-mers in the zebrafish reference genome; Varshney et al., 2015) The track allows for rapid selection of guide sequences with embedded information on predicted off-target sites in the genome. The number appearing in the name field of the BED track indicates the number of off-targets for the 12-mer seed region. Then it is also possible to sequence putative off-target sites to look for Cas9-cuttings in the injected embryos.

V.2.b) gRNA production

The single-guide RNA (sgRNA) contains in general a 20-nucleotide sequence complementary to the genomic loci to target, and 80 nts of chimeric guide RNA (crRNA:tracrRNA) (Cong et al. 2013; Mali et al. 2013). Most methods rely on subcloning the 20-nt target sequences between a T7 promoter and the crRNA:tracrRNA sequences and obtaining the sgRNAs from *in vitro* transcription off of the plasmid.

gRNAs were ordered as oligonucleotides containing 20nts homologous to the CRISPR targeted region and overhangs compatible with directional cloning into the BsaI-digested pDR274 vector, following Hwang et al. (2013) protocol (Table 10 summarizes the chosen sequences). Vector pDR274 harbors the T7 required for further *in vitro* transcription. This was

realized using Megascript T7 kit (Ambion) at 37°C overnight (better yield than the two hours recommended by the kit). One step of DNA digestion was performed using Turbo DNase and RNA cleaned using Qiagen RNeasy Minikit. Concentrations obtained were in the range 1.5-3 µg/µL. Another approach using T7-cas9-sgRNA2 vector (kindly provided by Dr Eirini Trompouki) was tested. One step-digestion/ligation using BsmBI, BgIII and SalI to digest the vector and T4 ligase inserted the customized gRNA inside this plasmid.

It is also possible to avoid the cloning step and used a simpler straightforward PCR approach to obtain the gRNA (Gagnon et al., 2014; Varshney et al., 2017). Two partially overlapping synthetic oligonucleotides can be ordered, one containing the target-specific gRNA and the other T7 transcription promoter and crRNA:tracrRNA, which can be used as a generic oligonucleotide for all constructs. Oligonucleotides are annealed and PCR-amplified before being transcribed.

V.2.c) CRISPR injection

Three papers in zebrafish suggest that injecting Cas9 protein is faster and yields more mutagenesis than injecting cas9 mRNA (Gagnon et al., 2014; Kotani et al., 2015; Sung et al., 2014), our collaborators confirmed this from their own experiments.

One-cell WT (AB) embryos were micro-injected with 1pL of CRISPR/cas9 mix (2µM Cas9 protein (NEB), 10x Cas9 buffer (NEB) and 300ng/µL sgRNA) to have roughly 1:1 cas9/sgRNA ratio. The mix was incubated 5 min at RT before injection. Injection was performed in the yolk. As a control, non-injected embryos or embryos injected with Cas9 alone were grown in parallel. WT sequence for *fn1b* gene was verified by tail-clipping the future parents used for injection (Table 11).

In a first assay, only the gRNAs 1 and 2 targeting exon 1 were injected. The plan was to then inject the other alone or in combination of two (*see the discussion part for further details and perspectives*).

V.2.d) Results of the injection, estimation of the mutagenesis rate

Ten embryos were randomly selected after 5 days to quantify the mutagenesis rates using the T7 endonuclease I (T7EI) assay (NEB). Not all assayed embryo mix showed mutagenesis. We

observed targeted indels in >60% of embryo mix injected with gRNA 2 targeting exon 1 (Figure 41).

Their siblings from the same injected clutch were raised to adulthood. 40% of them died during the few days of development, certainly due to the toxicity of the injected construct. However, the survival embryos did not present any phenotypical defects. Adult fish were then fin-clipped and fn1b gene sequenced to identify those who are carrying mutated alleles. Such screening fin clips of G0 individuals is however not predictive of the germline transmission rates or the transmitted alleles. Over 50 fish for each gRNA injected none of them presented mutation inside *fn1b* gene. No F0 founder grew to adulthood. It may be possible the putative founders died after dew days, the other did not get efficient cutting on both alleles.

V.3 Discussion and perspectives

CRISPR/Cas9 system was shown to be an efficient genome editing tool to target, with few off-targets effects, endogenous loci in wild-type zebrafish, where two copies of a given endogenous gene are present. Unlike Zinc Finger Nucleases (ZFNs) and TALENs, gRNA is the only component that needs customization for each genomic target of interest, thus greatly simplifying the design and lowering the cost of gene targeting.

In our study, considering the correct mutagenesis rate observed after few days after injection, surprisingly no CRISPR-injected fish presented a mutation at the adult state. Normally Cas9-induced mutations are heritable, but it seems the mutations did not pass the germ line. To achieve this, it would have been interesting to vary the injected gRNA/Cas9 protein/mRNA concentration and determine the optimal quantity necessary to obtain the highest mean frequency of mutations, and in the same time the best concentration avoiding high levels of toxicity (40% in our case was a limit acceptable level of toxicity). Cas9-induced mutagenesis was indeed shown to be dose-dependent (Jao et al., 2013).

Literature reports some issues to target some loci with CRISPR/Cas9 system in zebrafish. The mutagenesis efficiencies for the different Cas9 applications were shown to vary widely in reported studies, with several groups experiencing 50% or more of sgRNAs being ineffective for mutagenesis (Moreno-Mateos et al., 2015; Shah et al., 2015; Sung et al., 2014; Varshney et al., 2015).

The mutagenesis rates could also be increased by injecting simultaneously several gRNA targeting multiple loci of the same gene, thus maximizing the chance to get a KO and to achieve a germ-line transmission of Cas9-induced mutations. CRISPR/Cas9 system can induce heritable large deletions when at least two sgRNAs were simultaneously injected, deletions which are germline-transmitted mutations (Varshney et al., 2015).

Another point which could help improving the mutagenesis rate is the localization of delivery of the CRISPR/cas9 system. For facility, since the cytoplasm could be difficult to target especially when the first cell is really flat, the injections were performed into the yolk in the first cell and not in the cell itself. Aiming it could nevertheless improve the chance of successfully producing transgenic fish. The efficiency seems to be increased and lethality decreased with this approach (Erickson et al., 2016)

However, as in parallel, we received a *fn1b* mutant from ZIRC database: fn1b^{sa553} (point mutant), the on-going CRISPR project has not been finished, and these points never tested.

Table 10: gRNA and primer sequences to target CRISPR cutting in exons 1 and 3 of zebrafish *Fn1b* gene

Targeted exon	Online program used for the design	gRNA sequences	Primer sequences
Exon 1	Zifit	5'-TGAGTCAGTAAAGAGACTCC-3' PAM: CGG	E1_1F : TAGGAGTCTCTTTACTGACTCA E1_1R: AAACTGAGTCAGTAAAGAGACT
Exon 1	Zifit	3'-GGTCATTTTCACCCGCGTTA-5' PAM: AGG	E1-2F : TAGGTCATTTTCACCCGCGTTA E1_2R : AAACTAACGCGGGTGAAAATGA
Exon 1	Zifit	5'-GGAGTTTAGCCATCCACG-3' PAM: AGG	E1_5F : TAGGGGAGTTTAGCCATCCACG E1_5R : AAACCGTGGATGGCTAAACTCC
Exon 1	Crisprdirect (here T7 binding sites should be added and PAM removed)	5'-AAGCTGCTCTCCATAACG-3' PAM: CGG	E1_3F : TAGGAAGCTGCTCTCCATAACG E1_3R : AAACCGTTATGGAGAGCAGCTT
Exon 3	Crisprdirect (here T7 binding sites should be added and PAM removed)	5'-TGCATGCCACAATCCGCA-3' PAM: GGG	E1_4F : TAGGTGCATGCCACAATCCGCA E1_4R : AAACTGCGGATTGTGGCATGCA
Exon 3	Zifit	5'-CGCACGTTCCTACCGAGT-3' PAM: CGG	E3_1F : TAGGCGCACGTTCCTACCGAGT E3_1R : AAACACTCGGTAGGAACGTGCG
Exon 3	Zifit	5'-ACGAGCGACCAAAGGATA-3' PAM: AGG	E3_2F : TAGGACGAGCGACCAAAGGATA E3_2R : AAACTATCCTTTGGTCGCTCGT
Exon 3	Zifit	5'- AAGGATAACATGATATGG-3' PAM: AGG	E3_3F : TAGGAAGGATAACATGATATGG E3_3R : AAACCCATATCATGTTATCCTT

Table 11: primers to sequence WT genomic *fn1b* exon 1 and exon 3

Primer sequence	Fragment length
FP1_exon1: AGGGTGAGAGAACCTCATAAAGC RP1_exon1: CTCACTTAAACCGCGAACTGTCC	491pb
FP2_exon3: ACAGTTCGCGGTTTAAGTGAG RP2_exon3: aagcatttattaaacgttgtc	290 pb

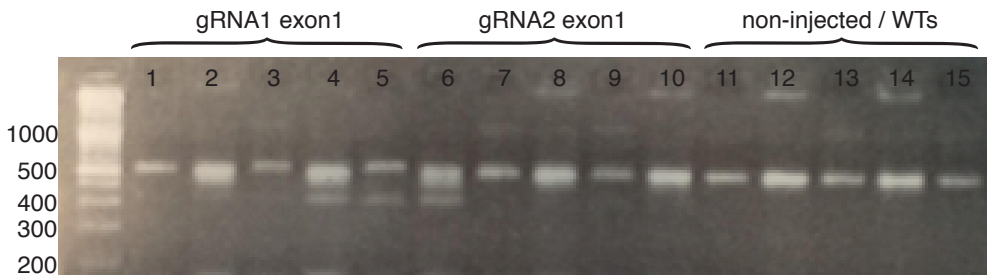


Figure 41: T7 endonuclease I assay to detect Cas9 induced mutations

T7 endonuclease I enzyme is able to detect heteroduplex DNA that results from the annealing of a modified DNA strand after a sgRNA/Cas9 mediated cut to WT DNA strand. In this thesis, this assay was used to get a first estimation of whether our Cas9-targeting approach was successful or not. 5 pools of 10 5dpf-injected embryos were used for each gRNA injected. Non-injected embryos from the same clutch were used as a control. The target sites in fn1b gene was first amplified by PCR, the generated fragments were then digested with T7E1. The digested products were loaded immediately on a 1.5% agarose gel.
Groups 2, 4, 5 and 6 show mismatches at the target sequence.

General conclusion and perspectives

The embryogenic morphogenetic processes shaping the vertebrate heart are complex and still poorly understood. Understanding the molecular and cellular details underlying AVC patterning and valve development is important in order to gain insights towards the etiology of valve diseases, which are a leading cause of birth defects and morbidity. Zebrafish is a powerful model organism to study these processes including AVC patterning and the function of cardiac valves.

This thesis aimed at elucidating the role of the stress-inducible transcription factor Egr1 in zebrafish AVC valvulogenesis. Our data demonstrated *egr1* gene as a key actor of zebrafish AV valve morphogenesis. Deletion of *egr1* results in altered cardiac valve formation and cushion remodeling. From ISH and wide-genomic data, its seems that *egr1* expression is regulated by *klf2a/klf2b* genes at 48hpf. The deletion of both *egr1* and *klf2a* is embryogenic lethal. This argue that *klf2a* and *egr1* are two master regulators of valve development. We analyzed further in details the valve defects of this double mutant *klf2a-klf2b* line and those of the single mutant *klf2b.* They both present a range of valve defects similar to those observed by Steed et al. (2016) for the single mutant *klf2a.* The double knock-out surprisingly does not affect more the phenotype than the single loss of *klf2a* or *klf2b*. This pointed out a putative compensatory role of *egr1* in valvulogenesis in double mutant *klf2a-klf2b*, which manage to survive to adulthood at a homozygous state, compared to the double mutant *egr1-klf2a* which died after few days. As discovered through the mRNAseq data, *klf2a* seems to be more activator rather *klf2b* a repressor, it could also explain the results observed for the double mutants.

The analysis of genome-wide analysis of transcriptomic landscape and chromatin organization at 48hpf in endothelial cells revealed that *klf2a* and *klf2b* do not share so many transcriptional target genes in common. This is in agreement with the hypothesis that they regulate different genes during valvulogenesis, however linked in the same signaling pathways. Together the combined mRNAseq/ATACseq results demonstrated that *egr1, klf2a and klf2b* have multi-faceted roles in this developmental program. They all regulate the development of AV valve by modulating the expression of *flt1*, *wnt9b and has2* genes in zebrafish heart at 48hpf. The role of *flt1* gene in this process was further investigated. *flt1*^{sa1504}

present severe valve defects. These data sets may also prove meaningful in other signaling pathways happening at 48hpf in endocardial cells. This comparative analysis mRNA-seq/ATACseq provide a great resource for analyzing transcriptomics landscape at 48hpf in endocardial cells. It will provide valuable information on different pathways, as well as for mechanotransduction pathways activated at 48hpf. We could perform a similar analysis with sequencing in addition the miRNA. Indeed, some of them were shown as important in development. In particular miR126 was shown to be regulated by *klf2a* during valvulogenesis.

This analysis could also be performed in a small subset of cells, for example only in AVC cells to reduce the heterogeneity and discovered specific genes important at 48hpf.

Valvulopathies are still a challenging issue in our society. This amount of data could help finding new interesting targets. In *klf2b* mutant, numerous genes linked with human valvulopathies were interestingly deregulated.

In addition, different projects were developed in this thesis. First, the generation of monoclonal antibodies against Klf2a, Klf2b and Erg1 zebrafish proteins was performed. It was undertaken trying to validate the KO of these proteins in mutant lines. Klf2a protein was confirmed absent in the KO-mutants used in the lab with one home-made monoclonal antibody (1KLF). This thesis presents an efficient protocol to generate specific and purified monoclonal antibodies against zebrafish proteins. For *klf2b*^{ig5} and *egr1*^{sa1504} mutant lines, the complete disappearance of the proteins still has to be proved. Further work should be done to characterize the generated antibodies and to conclude with certitude. Second, a ChIP assay against Klf2a protein for 48hpf zebrafish cells was developed. It could be repeated for IPs against Egr1 and Klf2a using the designed fish lines over-expressing a biotinylated version of Klf2a and Egr1 in endocardial cells. This assay would allow to discover direct downstream transcriptional targets of *egr1* and *klf2a*. Third, a CRISPR-Cas9 endonuclease system was designed to target *fn1b* gene. Even if the project was not finalized, the general protocol could be adapted in future for the generation of zebrafish KO mutants for a gene of interest.

Material and Methods

Zebrafish accession gene numbers

klf2a: ENSDARG00000042667
klf2b: ENSDARG00000040432
egr1: ENSDARG00000037421
flt1: ENSDARG00000019371
klf4: ENSDARG00000079922
klf17: ENSDARG00000038792

Zebrafish husbandry, fish strains and embryo treatments

The zebrafish (Danio rerio) lines used in the experiments were the following: wild-type AB, Tg(egr1^{sa64}) - from ZIRC, received as an heterozygous line ; Tg(klf2a^{ig4}) - generated in Vermot's lab using TALEN technology, published in (Steed et al., 2016) ; Tg(klf2b^{ig5}) - generated in Vermot's lab using TALEN technology (further information below), Tg(flt1^{sa1504}) - from ZIRC, received as an heterozygous line, Tg(pkd2^{tc321}) or *cup* mutants from ZIRC, Tg(vlt^{m651}) or *vald tepes*, *gata1* mutant, from ZIRC, Tg(tnnt2a^{b109}) or *silent heart* (sih) mutant from Stainier's lab (Sehnert et al., 2002). Tg(fli:gal4FF^{ubs}; UAS:kaede) (Herwig et al., 2011) from Markus Affolter's lab (Biozemtrum, Basel) and Tg(flk1:mCherry) (Bertrand et al., 2010).

Embryos were staged according to hours (hpf) and days postfertilization (dpf). They were incubated at 28.5°C in 0.3% Danieau medium supplemented 5h after birth with 0.003% (wt/vol) 1-phenyl-2-thiourea (PTU) (Sigma Aldrich) to inhibit pigment formation. All zebrafish strains were maintained at the IGBMC fish facility under standard husbandry conditions (14h light/10h dark cycle). The Animal Experimentation Committee of the Institutional Review Board of IGBMC approved all animal experiments performed in this project.

<u>*klf2a^{ig4}* mutant line</u>

Mutant described in (Steed et al., 2016). The INDEL generated using TALEN technology leads to premature stop codon in the protein. A second ATG could be a putative protein-coding start site, this site will be conserved in *klf2a^{ig4}* mutant. Its effective presence and cellular activity have still to be proved.

Klf2b^{ig5} mutant line

Tg(Klf2b) line was generated in the same way as Tg(Klf2a), using TALEN system. A TALEN pair (left and right arms: 5'GGACATGGCTTTACCT-3' and 5'AACGTTTGCAAACCAG-3') were designed to target exon 1 of the klf2b gene and injected into single cell wild-type (AB) first cell. We identified the alleles generated and confirmed that potential targeting events could be transmitted through the germline by out-crossing the F0 fish with AB animals and sequencing genomic DNA from pools of 6 F1 embryos. We focused on a 28pb-deletion mutation (5'-GGACATGGCTTTACCTTGCCTTTTGCCT-3') leading to a premature stop codon in klf2b transcript. Studies were performed from F4 fish and later generations, and on transgenic lines resulting of out-crossings. A PCR-based genotyping strategy was established using the following primers to identify the wild-type and mutant alleles, the length of the deletion allowing their visualization directly on a 3%-agarose DNA gel: forward 5'-GGAAAGCGCGTATATTTGGA-3', reverse 5'- CAAGTAGGAAATGCAAGTGT-3' and sequencing forward primer 5'- AGAGCGCACTGTGCCTTATA-3'.

Klf2a^{ig4}, *klf2b*^{ig5} and *egr1*^{sa64} *mutant* fish were viable and kept as homozygous fish to be able to work with maternal zygotic embryos in this study. Double mutant lines (*klf2a-klf2b*, *egr1-klf2a*) came from crossings between single mutant lines and genotyping.

egr1^{sa64} mutant line

Egr1^{sa64} mutant contains a C>T point mutation in the exon 2 of the gene leading to a premature stop codon in the predicted translation product. Studies on Tg(egr1^{Sa64}) line were realized from F4 fish and later generations, following out-crossing to transgenic lines of interest. Genotyping was performed by sequencing of the PCR product generated with the following primers: forward 5'-ATATCCTACACAGGCCGTTTCAC-3', reverse 5'-CACTGGGATATGTTGATGAGGAG-3', and sequenced with the primer 5'-TGGTTGAGAAGTTAGGGCCAGAC-3'.

flt1^{sa1504} mutant line

Flt1^{sa1504} mutant contains a T>G point mutation in the exon 9 of the gene causing a modified essential splice site. Studies on Tg(flt1^{Sa1504}) line were realized from F4 fish and later generations, following out-crossing to transgenic lines of interest. Genotyping was performed by sequencing of the PCR product generated with the following primers: forward for PCR and genotyping: 5'- TTAGGCTGAAGGATGGGATG -3', reverse 5'- TGGTCCTCTTTGAACAACCA-3'.

Over-expression of Klf2a protein in endothelial cells

Tg(fli:gal4FF; UAS:kaede) line allows the expression of gal4FF gene under the control of Friend Leukemia Integration 1a (fli) promoter, which is specific of endothelial and hematopoietic cells. This line was crossed with Tg(UAS:klf2a LV or SV; cmlc2:eGFP). As explained in a previous paragraph, a second ATG could express a shorter version of Klf2a protein, LV stands for long version (known klf2a protein form) and SV for short version (putative form). Fish were selected under green fluorescence macroscopy: the over-expressing Klf2a fish present fluorescence inside the cardiovascular system and an enhanced staining in the heart, controls fish are fluorescent only for the heart.

Over-expression of a biotinylated version of Klf2a and Egr1 proteins in endothelial cardiac cells

The assembly of the different modules /megaprimers presented in the chapter 4 was realized using Golden Gate method (adapted from Miyazaki protocol, 2011) based on Megawhoop cloning. Adapters containing BsaI restriction sites and pUC-sequences of insertion were added to the extremities to each megaprimer by PCR. Megaprimers were amplified by successive PCR to assemble the different required parts. For the cloning of modules 2A and 3 containing UAS promoter, the complete sequence with 4xUAS sequences was ordered as an ultramer (IDT). For the fragment containing *egr1* genomic sequence, additional PCRs were performed to mutate BsaI restriction site, which will have impaired the module assembly.

PCRs to generate the megaprimers:

- m573 (Tol2-L200-zCmlc2p-Kozak-ECFP): PCR1 fragment tol2-L200 Tm=66°C, m573-fwd1/m573-rev1; PCR2 fragment zCmlc2 Tm=62°C, m573-fw2/m573-rev2; PCR3 fragment ECFP Tm=72°C, m573-fwd3/m573-rev3 and fusion PCR (purified PCR1, PCR2, PCR3) Tm=72°C, m573-fwd1/m573-rev3.

-m574/577 (SV40pA-UAS-TS): PCR1, SV40 pA, Tm=63°C, m574/577-fwd/m574/577-rev; PCR2 UAS Tm=72°C, either m574/577-fwd/m574-rev2 or m577rev2.

-m575 (klf2a-(4GS)3-Flag-bioTag): PCR1 Tm=64°C, m575-fwd3/m575-rev; PCR2 with PCR1 purified fragment Tm=72°C, m575-fwd2/m575-rev and PCR3 with PCR2 purified fragment Tm=70°C, m575-fwd1/m575-rev.

- m576 (egr1-(4GS)3-Flag-bioTag): PCR1 Tm=72°C, m576-Fwd1/m576rev3 with the modified cDNA of *egr1*; PCR2 Tm=72°C m756-Fwd1/m576Rev4 with PCR1 purified fragment; PCR3 Tm=72°C m576-Fwd1/m576-rev5 with PCR2 purified fragment.

-m578 (SV40pA-Tol2-R150): PCR1, SV40 pA, Tm=61°C, m578-fwd1/m578-rev1; PCR2, tol2 R150, Tm=59°C, m578-fwd2/m578-rev2; fusion PCR3 with purified PCR1 and PCR2, Tm=72°C, m578-fwd1/m578-rev2.

PCR mix
Q5-buffer 5X= 5µL
10mM dNTPs: 0.5µL
50µM Fwd primer: 0.25µL
50µM Rev primer: 0.25µL
5ng template: 1µL
Q5 polymerase HF: 0.5µL
H2O qsp: 17.5µL

PCR settings

1. 98°C 30”
2. 98°C 10”
3. *Tm* 10”
4. 72°C extT 20”/kb >> goto2 - loops 30X
5. 72°C 2’
6. 4°C hold

`

PCR fragments were purified using NucleoSpin columns (Gel and PCR clean-up, MN).

Megaprimers were inserted into pUC plasmid in Q5-PCR (megawhoop reaction).

Cloning of Megaprimer in pUC vector
Reaction mix
Q5-buffer 5X: 5µL
dNTPs 10mM: 0.5µL
100ng megaprimer: from 0.5 to 2µL
50ng pYC vector: 1µL
qsp H20 to 25µL final volume

PCR settings

1. 98°C 30”
2. 98°C 10”
3. 68°C 10”
4. 72°C extT 20”/kb (vector + fragment) >> goto2 - loops 30X
5. 72°C 2’
6. 4°C hold

Plasmid were digested one hour 37°C-digestion by Dpn1 (NEB) and transformed into TOP10 competent cells (Sigma) by heat shock (1µl reaction / 50µL bacteria). Bacteria were plated (9/10 and 1/10) on LB supplemented with ampicillin and Xgal for blue/white selection. Mini-prep from white colonies were performed and their sequences verified by sequencing (GATC/Eurofins). The final assembly was realized into pXpA vector (unpublished data, kindly provided by Dr Reina San Marin’s lab, IGBMC) in one reaction (digestion by BsaI and ligation by T4 DNA ligase (NEB)). Final plasmids were checked to be error-prone by sequencing. For each construction one clone was selected: pXpA_633 mini prep n°6 for klf2a construction and pXpA_634 mini prep n°9 for egr1 construction.

Genotyping

Genotyping of adult fish was performed three months after birth on genomic DNA extracted from a small piece of the caudal fin lysed at 55°C in 100µL of SDS-lysis buffer (10mM Tris-HCl pH8, 200mM NaCl, 10mM EDTA, 0.5% SDS and 100 µg/mL Proteinase K), and then purified using isopropanol/70%-ethanol.

Genotyping of embryos was either realized on whole embryos (after live imaging or ISH) or from dissected tails (before immunofluorescence), lysed in 50µL of 50mM NaOH at 95°C for 10min. Lysis reaction was stopped with 10µL Tris-HCl pH8. 1µL used for PCR reaction.

PCR

Reaction mix for one sample:
10µL Phusion High-Fidelity PCR master mix (F531L, Thermo Scientific)
1µL Forward primer at 10µM
1µL Reverse Primer at 10µM
7µL milliQ water
1µL DNA (100 to 500 ng of genomic DNA)

PCR settings

7. 98°C 3'
8. 98°C 10"
9. *Tm* 20"
10. 72°C extT 20"/kb >> goto2 - loops 30X
11. 72°C 7'
12. 4°C hold

Electrophoresis

2 or 3% agarose gel
1 drop of 10%-Ethidium bromide per 100mL
Gel run at 130 V
DNA check under UV lamp

In Situ Hybridization (ISH)

ISH assay was performed as in (Thisse and Thisse 2008) using the following anti-sense probes:

- *notch1b probe*: plasmid pCR-scriptSK+, made by Sonja Chocron, Baekkers's lab. Linearized by BamHI (NEB) Transcription using T3 polymerase.
- *klf2a probe*: obtained by PCR amplification of the plasmid IRBOp991B0734D, provided by RPDZ, Berlin, using forward primer: 5'- CAGGCGACTACAGAATGCA -3' and reverse primer: 5'-TAATACGACTCACTATAGGGAGTGAC. Transcription with T7 polymerase.
- *klf2b probe*: obtained by PCR amplification of the plasmid #1343 pSCB-klf2b, provided by Mrs Cecile Otten, Seyfried group), 640pb fragment amplified with reverse primer 5'- CTACGGTCCGGTGATAGGCATG-3' and forward primer 5'-AGCATTTAGGTGACACTATAGTCACAGGTGTCTCTTCATGTGCAG-3'. Transcription by SP6 polymerase.
- *egr1 probe*: from pBSK plasmid, PCR amplification using forward primer: 5'-ATGACCCGTGAGTCAGTAA-3' and reverse primer: 5'-ATTAACCCTCACTAAAGGGACTTGGTGCCCTGAGTTCTGAT-3'. Transcription using T3 polymerase.
- *flt1 probe*: from a partial cDNA fully sequenced ordered at Science Biosource (IRCYp5023H065D, 9038840 IMAGE ID), pCR4-TOPO plasmid, linearized by NCOI (NEB). Transcription using SP6 polymerase.
- *Bmp4 probe*: Transcription using T7 polymerase
- *Wnt9b probe*: forward primer: 5'-TATTGCCCTCTGCATCCTTC-3' and reverse primer: 5'-TGACATTCAACGTGACAGCA-3'
- *Klf4 probe*: designed from a partial cDNA (EST) ordered on GenomeCube (IMAGE ID 4467522/IMAGp998H1910280Q M13F, sequence ID BI882122.1), 515pb probe amplified using forward primer: 5'-ACTGAGTTTGATAGCATGGCAC-3' and reverse primer 5'-AGCATTTAGGTGACACTATAGGGGAAGGTGTGAGTGTGTAGG-3'. Transcription using SP6 polymerase.

Probe design

A PCR amplification was preferred to plasmid linearization before transcription, this method gave us cleaner ISH with less background. PCR products were purified using Qiagen PCR columns and resuspended in 30µL water solution.

Transcription

Transcription was performed for two hours at 37°C using the following mix:

4µL Buffer transcription 5x
2µL UTPdig ATP/GTP/CTP/UTP mix
0.5µL Rnasine (20-50U/µL)
0.5µL DTT (0.5mM)
2ug PCR product (5uL in general)
2µL RNA pol at 20U/µL T3, T7 or SP6 (second microliter added after one hour)
Up to 20µL H2O sterile

2µL of reaction were kept for control of transcription on an agarose gel. The last 18µL were treated with DNAse I for 20min at 25°C (0,1µL of DNase, 2,5uL of buffer 10x and 3.4µL H20 RNAse free were added to the transcription mix). The efficiency of digestion was controlled on an agarose gel with the non-digested sample.

RNA was purified using 2.5µL LiCl 4M and 75µL 100% cold EtOH and stored 30 min at -80°c or 2H at -20°C. Spin in a cold centrifuge for 20min. The pellet was then dissolved with 22µL H2O, 2.5µL LiCl and 75µL 100% EtOH and stored 2H at -20°C before spinning down 15 min, wash one time with cold 70% EtOH and air dried. RNA was finally resuspended in 100µL Hybridization buffer according to the quantity of the transcript and stored at -20°C.

Revealing solution

After 3 days embryos were washed four in PBS-0.1% Tween (PBST) times 30min at RT under gentle agitation followed by three washes in the revealing solution (100mM pH9.5 Tris-HCl; 50mMgCl2; 100mM NaCL; 0.1% Tween 20 in water), prepared fresh, 10 min at RT under gentle agitation. BM purple (Sigma Aldrich, Roche, 11442074001) was added in a volume sufficient to cover the embryos. The signal was developed in the dark at 37°C. The reaction was stopped with three wash 5min in PBS-T before fixation 20min in 4%-PFA. The stained embryos can be stored in the dark several weeks at 4°C in PBS-T.

Imaging of WISH

To optimize the imaging of the heart after staining, in certain cases the embryos were made more transparent using fructose. A first bath in 45%-fructose solution (D-fructose, Sigma, ref F0127) containing 1/100 1-Thioglycerol (Sigma M6145) was performed for 15 min followed by a second bath in 90%-fructose solution from 30min to 1 hours depending the stage of the embryo. Imaging was performed using a Leica M165 macroscope with a TrueChrome Metrics (Tucsen) with a Leica 1.0X objective (10450028).

High throughput dissection of 48hpf-hearts

Hearts were dissected from *klf2a*^{ig4}; *klf2b*^{ig5} ; egr1^{sa64} and *klf2a;klf2b* lines in background *fli:lifeact-eGFP; flk:nls-mcherry* and respective control embryos at the desired stage (48hfp) using high-throughput extraction technique, optimized from (Lombardo et al., 2015) protocol. Hearts were separated mechanically by pipetting up and down batchs of 200 embryos in culture medium (L-15 Leibovitz, 10% FCS 9150, 1.25mM CaCl2, 800mg/L glucose, 50microg/mL penicillin, 0.05mg/mL streptomycin), medium prepared fresh for each experiment. Hearts were then separated from debris after passages on two different filters: first purification though a 100µm nylon cell stainer (Falcon, 352360), and further collection on a pre-separation filter 30µm (Miltenyi Biotech). After centrifugation 10min at 2.5g, culture medium was replaced by cold FACS medium (PBS with 2% Fetal Calf Serum 9150, 1% Penicilline/streptavidine, 1mM EDTA) and final filtration was performed with tip strainers (40µm) (Scienceware, Bel-Art, flowmi tip strainers).

Fluorescence-activated cell sorting (FACS)

Subsequent separation of endocardial cells from the other cardiac cell types was achieved directly after heart extraction by FACS. Extracted hearts were pipetted up and down to dissociate cells in FACS medium. FACSAria Fusion (BD Biosciences) device was used. Cells were sorted at 4°C on a FITC-A detector and flow cell passed through a 70µm-nozzle.

Sorted cells were collected at 4°C in PBS (9,5µL cold PBS and 0,24µL RNasin (Sigma)). Immediate freezing in dried ice of 1000 cell-samples was either performed for mRNA sequencing assay. Or cells collected in 50µL cold PBS were directly treated after sorting with the transposase Tn5 for ATAC-seq assay.

Chromatin Immunoprecipitation

Samples preparation

48hpf-embryos were dechorionated by Pronase. Crosslinking was performed in 1%-formaldehyde for 5 min at RT with shaking. The reaction was quenched with 1/20 volume 2.5 M glycine for 3 min at RT with shaking. Embryos were washed twice with ice-cold PBS and the pellet was flash-frozen in liquid nitrogen. Embryos were kept at −80°C until the experiments were performed.

Embryos were lysed in three successive buffers. First in 5ml of lysis buffer 1 (50 mM HEPES-KOH, pH 7.5, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP-40, 0.25% Triton X-100, 1× protease inhibitors) for 10 min at 4°C – then in 5mL lysis buffer 2 (10 mM Tris-HCl, pH 8.0, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 1× protease inhibitors) and finally in 1ml lysis buffer 3 *(200µL for FACS-sorted cells)* (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.1% Na-Deoxycholate, 0.5% N-lauroylsarcosine, 1× protease Inhibitors)

Sonication

Sonication was performed in Covaris E220 (Sonolab 7.3. software) in 1 mL glass tubes for 5 min (duty cycle 20%, intensity 8 and 200 cycles per burst). After a spinning down for 15 min at full speed at 4°C to eliminate the debris. 10% of Triton was added to the supernatants. 50ul of sonicated chromatin (called INPUT sample) for futher quantification of ChIP was kept at -20°C for later use.

Incubation of antibodies with the magnetic beads

50µL of magnetic beads coupled with protein A (Dynabeads® Protein A Beads & Microspheres) (Life Technologies) were used for one ChIP sample. After two washed with fresh PBS/0.5%BSA, the beads were incubated with 5µg of antibody in 250µl of PBS/0.5%BSA, for at least 5 hours at 4°C under rocking. After two washes with PBS/0.5%BSA, the sonicated chromatin was added to the beads and Incubatied o/n under rotation at 4°C. The following day, beads were washed for 5min at 4°C with the following three buffers: buffer 1 (20mM Tris-HCl pH8, 150 mM NaCl, 2mM EDTA, 0.1% SDS, 1%Triton X-100), buffer 2 (20mM Tris-HCl pH8, 500mM NaCl, 2mM EDTA, 0.1% SDS, 1%Triton X-100) and buffer 3 (10mM Tris-HCl pH8, 250nM LiCl, 2mM

EDTA, 1% NP40). A final wash was performed in TE. The beads were then resuspended in 200µL of elution buffer (50mM Tris-HCl pH8.0, 10mM EDTA and 1%-SDS) - 150µL was added to the input sample too. DNA was eluted from the beads by shaking the tubes at 65°C for at least 30min. The beads were removed and the reverse-crosslink reaction was continued o/n at 65°C. 200µl of TE and 0,2mg/mL of RNAse were added and the IPs left for 3h at 37°C. 2µg/ml of proteinase K was then added at 55°C for 2h. Chromatin IP was purified with MiniElute kit (Qiagen). The pellet was finally resuspended in 30µL of water.

ATAC-seq

see Materials and Methods of the paper.

mRNA-sequencing

see Materials and Methods of the paper.

Confocal valve imaging and AV photoconversion

See Materials and Methods of the paper.

Additional data about the imaging of heart valves of lines which over-express biotinylated Egr1 or Klf2a protein using spinning disk confocal microscopy

For live imaging, zebrafish embryos were staged, anaesthetized with 0.02% tricaine solution and mounted in 0.7% low melting-point agarose (Sigma Aldrich). Confocal imaging was performed either on a Leica spinning disk (to check for mCherry signal in the fli:gal4FF/UAS:kaede; cmlc2:CFP; UAS:(4GS)3-Flag-BioTag-egr1/klf2a-P2A-mCherry). Fast confocal imaging to image heart structure was performed using a Leica DMi8 combined with a CSU-X1 (Yokogawa) spinning at 10 000 rpm, 2 simultaneous cameras (TuCam Flash4.0, Hamamatsu) and a water immersion objective (Leica 20X, N.A. 0.75 or Leica 40X, N.A. 1.1). 20ms exposure was used for whole heart imaging. 1% of 488 laser power and 100% of 561 laser power were used for validating green (GFP signal) and red (mCherry signal) expressions.

Western Blot

Fifty 48hpf-embryos were dechorionated either manually with forceps or using Pronase (Roche) enzyme (10 min in danieau and 1mg/mL of pronase). Their yolk was removed by pipetting up and down in deyolking buffer (55mM NaCl; 1.8mM; KCL, 1.25mM NaHCO3) before being homogenized into 150µL of cold lysis SDS-ß-mercaptoethanol-buffer (0.25mM Tris-HCl pH6.8; 40% glyceryl, 10% ß-mercaptoethanol, 4% SDS and 0.3% bromophenol blue) for 10 min on ice and then 5min at 100°C to denature the proteins. Lysates were centrifuged at full speed at 4°C for 15min to pellet any remaining debris and embryos. The protein concentration of the lysates was determined using BCA dosage assay/Bradford or Nanodrop measurement. The protein solution was either used directly or snap-frozen and stored at -80°C until use (*not recommended*). Around 400ug / 13µL of the lysate solution were loaded per well on 12%-precast SDS-polyacrylamide gel mini protean TGX gel (Biorad). Proteins were then transferred onto a nitrocellulose membrane (0,45µm) (Amersham GE Healthcare) at 400mA for 1h in transfer buffer. Blots were blocked for one hour at RT in PBST-5% dry milk and then incubated at 4°C with primary antibody o/n. They were washed three times 10min in PBS-0,5% Tween 20 and incubated with either secondary goat anti-rabbit HRP or goat anti-mouse HRP antibodies for 1 hour at RT. Chemiluminescent detection was performed using Clarity system (Bio-rad).

Monoclonal and polyclonal antibody production and purification

Sulfolink coupling gel (Pierce N°20401) was equilibrated with 3mL of TE (50mM Tris, 5mM EDTA, pH8.5) before being coupled with 1mL of peptide (1 to 5mg) in TE during 45min at RT. Wash in TE two times. Non-specific sites were blocked with 1mL of 50mM L-cysteine in TE during 45min at RT. Wash in 3mL 1M NaCl two times. Then 1mL of each serum was incubated with the peptide coupled-Sulfolink o/n at 4°C under agitation. The following day, the column was washed with PBS two times and elution was realized with 0.1M pH2.8 glycine elution buffer. 500 µL fractions were eluted and 25µL 1M pH9.5 Tris was added to neutralize the solution. The concentration of each fraction was determined with Nanodrop. Purified antibodies were stored in 30% glycerol in small aliquots at -20°C.

Immunoprecipitation protocol

(from Mustapha Oulad-Abdelghani, IGBMC)

For 1mL of input material, 1/10 volume of protein G-sepharose beads (Sigma, ref 17-0618-05) was washed twice with milliQ water and twice with IP 100mM KCl buffer (25mM Tris HCl ph7.9 ; 0.1% NP-40 ; 5mM MgCl2 ; 10% glycerol ; 100mM KCl ; 2mM DTT and 1X protease cocktail inhibitor). Beads were first incubated with 1mg of antibody of interest per mL of protein G sepharose and 1 beads volume of IP 100mM KCl buffer, at least 1 hour at RT under gentle agitation. Beads were then washed twice with at least 10 volumes of IP 500mM KCl buffer and three times with IP 100mM KCl buffer. 1mL of protein lysate (*prepared as explained in the Wblot protocol part*) was incubated with the beads o/n at 4°C under gentle agitation. The resin was then washed the following day twice with 1 volume of IP 500mM KCl buffer and three times with 1 volume of IP 100mM KCl buffer. Protein was eluted with 1mL 0.1M ph2.8 glycine. Supernatant was neutralized with Tris pH8.8 buffer and loaded on gel protein. A western blot was performed to confirm the efficiency of the immunoprecipitation.

Immuno-fluorescence assay

Embryos were fixed in PBS/2%-PFA o/n at 4°C protected from light. The following day, they were washed four times 5' in PBS-0.5%Tween 20 and permeabilized in PBS-T + 1% TritonX-100 o/n at 4°C under rocking. An additional step of piercing was performed for some assays, embryos were pierced at the back of the pericardial cavity near the "neck" to facilitate antibody entry, using a pair of forceps. A first incubation in blocking buffer was performed o/n at 4°C into specific buffer (PBS-T; 0.5% TritonX-100; 1% BSA and optional, 10% normal goat serum). The primary antibody was incubated in the specific blocking solution o/n at 4°C. Six washed in PBS-T over 4 hours at RT were done on day 5. The suitable secondary antibody (Alexa Fluor red-fluorescent 594, far-red-fluorescence 633 or green-fluorescence 488 anti-rabbit or anti-mouse depending on the primary antibody used, Invitrogen) in the same blocking solution was added o/n at 4°C. Final six washes in PBS-T were performed before imaging on confocal.

Statistical Analysis

The Chi-squared analysis was performed for statistical analysis of WISH data. P-values were indicated in the figures as *p<0.05, **p<0.01, ***p<0.001 and ****p<0.0001

Student's t-test and Fisher F-test were used for statistical analysis for qPCR. Standard deviation was used to measure deviation from the mean, for all experiments. For all of the statistical qPCR tests, p values #0.05 were considered significant.

Statistical analysis for valve defects was realized using an unpaired two-tailed Student's t test.

References

Ahuja, S., D. Dogra, D. Y. R. Stainier and S. Reischauer, (2016). "Id4 functions downstream of Bmp signaling to restrict TCF function in endocardial cells during atrioventricular valve development." Dev Biol 412(1): 71-82.

Aksoy, I., V. Giudice, E. Delahaye, F. Wianny, M. Aubry, M. Mure, J. Chen, R. Jauch, G. K. Bogu, T. Nolden, H. Himmelbauer, M. Xavier Doss, A. Sachinidis, H. Schulz, O. Hummel, P. Martinelli, N. Hubner, L. W. Stanton, F. X. Real, P. Y. Bourillot and P. Savatier, (2014). "Klf4 and Klf5 differentially inhibit mesoderm and endoderm differentiation in embryonic stem cells." Nat Commun 5: 3719.

Anders, S. and W. Huber, (2010). "Differential expression analysis for sequence count data." Genome Biol 11(10): R106.

Anders, S., P. T. Pyl and W. Huber, (2015). "HTSeq--a Python framework to work with high-throughput sequencing data." Bioinformatics 31(2): 166-169.

Anderson, K. P., C. B. Kern, S. C. Crable and J. B. Lingrel, (1995). "Isolation of a Gene Encoding a Functional Zinc Finger Protein Homologous to Erythroid Kru¨ppel-Like Factor: Identification of a New Multigene Family." Mol Cell Biol: 5957-5965.

Armstrong, E. J. and J. Bischoff, (2004). "Heart valve development: endothelial cell signaling and differentiation." Circ Res 95(5): 459-470.

Arthur, H. M. and S. D. Bamforth, (2011). "TGFbeta signaling and congenital heart disease: Insights from mouse studies." Birth Defects Res A Clin Mol Teratol 91(6): 423-434.

Asakawa, K. and K. Kawakami, (2008). "Targeted gene expression by the Gal4-UAS system in zebrafish." Dev Growth Differ 50(6): 391-399.

Bakkers, J., (2011). "Zebrafish as a model to study cardiac development and human cardiac disease." Cardiovasc Res 91(2): 279-288.

Balciunas, D., K. J. Wangensteen, A. Wilber, J. Bell, A. Geurts, S. Sivasubbu, X. Wang, P. B. Hackett, D. A. Largaespada, R. S. McIvor and S. C. Ekker, (2006). "Harnessing a high cargo-capacity transposon for genetic applications in vertebrates." PLoS Genet 2(11): e169.

Banjo, T., J. Grajcarek, D. Yoshino, H. Osada, K. Y. Miyasaka, Y. S. Kida, Y. Ueki, K. Nagayama, K. Kawakami, T. Matsumoto, M. Sato and T. Ogura, (2013). "Haemodynamically dependent valvulogenesis of zebrafish heart is mediated by flow-dependent expression of miR-21." Nat Commun 4: 1978.

Bartman, T., E. C. Walsh, K. K. Wen, M. McKane, J. Ren, J. Alexander, P. A. Rubenstein and D. Y. Stainier, (2004). "Early myocardial function affects endocardial cushion development in zebrafish." PLoS Biol 2(5): E129.

Bedell, V. M., Y. Wang, J. M. Campbell, T. L. Poshusta, C. G. Starker, R. G. Krug, 2nd, W. Tan, S. G. Penheiter, A. C. Ma, A. Y. Leung, S. C. Fahrenkrug, D. F. Carlson, D. F. Voytas, K. J. Clark, J. J. Essner and S. C. Ekker, (2012). "In vivo genome editing using a high-efficiency TALEN system." Nature 491(7422): 114-118.

Beets, K., M. W. Staring, N. Criem, E. Maas, N. Schellinx, S. M. de Sousa Lopes, L. Umans and A. Zwijsen, (2016). "BMP-SMAD signalling output is highly regionalized in cardiovascular and lymphatic endothelial networks." BMC Dev Biol 16(1): 34.

Beis, D., T. Bartman, S. W. Jin, I. C. Scott, L. A. D'Amico, E. A. Ober, H. Verkade, J. Frantsve, H. A. Field, A. Wehman, H. Baier, A. Tallafuss, L. Bally-Cuif, J. N. Chen, D. Y. Stainier and B. Jungblut, (2005). "Genetic and cellular analyses of zebrafish atrioventricular cushion and valve development." Development 132(18): 4193-4204.

Bennett, J. S., D. M. Stroud, J. R. Becker and D. M. Roden, (2013). "Proliferation of embryonic cardiomyocytes in zebrafish requires the sodium channel scn5Lab." Genesis 51(8): 562-574.
Bertrand, J. Y., N. C. Chi, B. Santoso, S. Teng, D. Y. Stainier and D. Traver, (2010). "Haematopoietic stem cells derive directly from aortic endothelium during development." Nature 464(7285): 108-111.
Bhattacharyya, S., F. Fang, W. Tourtellotte and J. Varga, (2013). "Egr-1: new conductor for the tissue repair orchestra directs harmony (regeneration) or cacophony (fibrosis)." J Pathol 229(2): 286-297.
Bill, B. R., A. M. Petzold, C. J. Clark, L. A. Schimmenti and S. Ekker, (2009). "A primer for Morpholino use in zebrafish." Zebrafish 6(1).
Boselli, F., J. B. Freund and J. Vermot, (2015). "Blood flow mechanics in cardiovascular development." Cell Mol Life Sci 72(13): 2545-2559.
Brite, J., S. K. Laughon, J. Troendle and J. Mills, (2014). "Maternal overweight and obesity and risk of congenital heart defects in offspring." Int J Obes (Lond) 38(6): 878-882.
Bruneau, B. G., (2008). "The developmental genetics of congenital heart disease." Nature 451(7181): 943-948.
Buck, M. J. and J. D. Lieb, (2004). "ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments." Genomics 83(3): 349-360.
Buenrostro, J. D., P. G. Giresi, L. C. Zaba, H. Y. Chang and W. J. Greenleaf, (2013). "Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position." Nat Methods 10(12): 1213-1218.
Buenrostro, J. D., B. Wu, H. Y. Chang and W. J. Greenleaf, (2015). "ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide." Curr Protoc Mol Biol 109: 21 29 21-29.
Burmeister, S. S. and R. D. Fernald, (2005). "Evolutionary conservation of the egr-1 immediate-early gene response in a teleost." J Comp Neurol 481(2): 220-232.
Bussmann, J., J. Bakkers and S. Schulte-Merker, (2007). "Early endocardial morphogenesis requires Scl/Tal1." PLoS Genet 3(8): e140.
Camenisch, T. D., A. P. Spicer, T. Brehm-Gibson, J. Biesterfeldt, M. L. Augustine, A. Calabro, J. S. Kubalak, S. E. Klewer and J. A. McDonald, (2000). "Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme." Journal of clinical investigation 106(3): 349-360.
Chakrabarti, S., G. Streisinger, F. Singer and C. Walker, (1983). "Frequency of Y-RAY induced specific locus and recessive lethal mutations in mature germ cells of the zebrafish, brachydanio rerio " Genetics 103: 109-123.
Chan, A. C., D. Y. Li, M. J. Berg and K. J. Whitehead, (2010). "Recent insights into cerebral cavernous malformations: animal models of CCM and the human phenotype." FEBS J 277(5): 1076-1083.
Chang, C. P., J. R. Neilson, J. H. Bayle, J. E. Gestwicki, A. Kuo, K. Stankunas, I. A. Graef and G. R. Crabtree, (2004). "A field of myocardial-endocardial NFAT signaling underlies heart valve morphogenesis." Cell 118(5): 649-663.
Chappell, J. C., K. P. Mouillesseaux and V. L. Bautch, (2013). "Flt-1 (vascular endothelial growth factor receptor-1) is essential for the vascular endothelial growth factor-Notch feedback loop during angiogenesis." Arterioscler Thromb Vasc Biol 33(8): 1952-1959.
Chen, J. N., F. J. M. van Eeden, K. S. Warren, A. Chin, C. Nüsslein-Volhard, P. Haffter and M. Fishman, (1997). "Left-right pattern of cardiac BMP4 may drive asymmetry of the heart in zebrafish." Development 124: 4373-4382.

Chen, S. J., H. Ning, W. Ishida, S. Sodin-Semrl, S. Takagawa, Y. Mori and J. Varga, (2006). "The early-immediate gene EGR-1 is induced by transforming growth factor-beta and mediates stimulation of collagen gene expression." J Biol Chem 281(30): 21183-21197.
Chen, T. and S. Y. Dent, (2014). "Chromatin modifiers and remodellers: regulators of cellular differentiation." Nat Rev Genet 15(2): 93-106.
Chiplunkar, A. R., T. K. Lung, Y. Alhashem, B. A. Koppenhaver, F. N. Salloum, R. C. Kukreja, J. L. Haar and J. A. Lloyd, (2013). "Kruppel-like factor 2 is required for normal mouse cardiac development." PLoS One 8(2): e54891.
Christy, B. A., L. F. Lau and D. Nathans, (1988). "A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences." Proc Natl Acad Sci U S A 85: 7857-7861.
Close, R., S. Toro, J. A. Martial and M. Muller, (2002). "Expression of the zinc finger Egr1 gene during zebrafish embryonic developmen." Mech Dev 18: 269-272.
Colombo, S., C. de Sena-Tomas, V. George, A. A. Werdich, S. Kapur, C. A. MacRae and K. L. Targoff, (2018). "Nkx genes establish second heart field cardiomyocyte progenitors at the arterial pole and pattern the venous pole through Isl1 repression." Development 145(3).
Cullere, X., E. Plovie, P. M. Bennett, C. A. MacRae and T. N. Mayadas, (2015). "The cerebral cavernous malformation proteins CCM2L and CCM2 prevent the activation of the MAP kinase MEKK3." Proc Natl Acad Sci U S A 112(46): 14284-14289.
Dang, D. T., J. Pevsner and V. Yang, W., (2000). "the biology of mammalian Kruppel-like family of transcription factors." int J biochem Cell Biol 32(11-12): 1103-1121.
Dekker, R. J., R. A. Boon, M. G. Rondaij, A. Kragt, O. L. Volger, Y. W. Elderkamp, J. C. Meijers, J. Voorberg, H. Pannekoek and A. J. Horrevoets, (2006). "KLF2 provokes a gene expression pattern that establishes functional quiescent differentiation of the endothelium." Blood 107(11): 4354-4363.
Dekker, R. J., S. van Soest, R. D. Fontijn, S. Salamanca, P. G. de Groot, E. VanBavel, H. Pannekoek and A. J. Horrevoets, (2002). "Prolonged fluid shear stress induces a distinct set of endothelial cell genes, most specifically lung Kruppel-like factor (KLF2)." Blood 100(5): 1689-1698.
Dietrich, A. C., V. A. Lombardo, J. Veerkamp, F. Priller and S. Abdelilah-Seyfried, (2014). "Blood flow and Bmp signaling control endocardial chamber morphogenesis." Dev Cell 30(4): 367-377.
Dina, C., N. Bouatia-Naji, N. Tucker, F. N. Delling, K. Toomer, R. Durst, M. Perrocheau, L. Fernandez-Friera, J. Solis, P. investigators, T. Le Tourneau, M. H. Chen, V. Probst, Y. Bosse, P. Pibarot, D. Zelenika, M. Lathrop, S. Hercberg, R. Roussel, E. J. Benjamin, F. Bonnet, S. H. Lo, E. Dolmatova, F. Simonet, S. Lecointe, F. Kyndt, R. Redon, H. Le Marec, P. Froguel, P. T. Ellinor, R. S. Vasan, P. Bruneval, R. R. Markwald, R. A. Norris, D. J. Milan, S. A. Slaugenhaupt, R. A. Levine, J. J. Schott, A. A. Hagege, M. V. P. France, X. Jeunemaitre and M. N. Leducq Transatlantic, (2015). "Genetic association analyses highlight biological pathways underlying mitral valve prolapse." Nat Genet 47(10): 1206-1211.
Driever, W., L. Solnica-Krezel, A. F. Schier, S. C. F. Neuhauss, J. Malicki, D. L. Stemple, D. Y. R. Stainier, F. Zwartkruis, S. Abdelilah, Z. Rangini, J. Belak and C. Boggs, (1996). "A genetic screen for mutations affecting embryogenesis in zebrafish." Development 123: 37-46.
Drummond, I. A., P. Rohwer-Nutter and V. P. Sukhatme, (1994). "The Zebrafish egrl Gene Encodes a Highly Conserved, Zinc-Finger Transcriptional Regulator." DNA and cell biology 13(10): 1047-1055.
Dupuis, L. E., D. R. McCulloch, J. D. McGarity, A. Bahan, A. Wessels, D. Weber, A. M. Diminich, C. M. Nelson, S. S. Apte and C. B. Kern, (2011). "Altered versican cleavage in

ADAMTS5 deficient mice; a novel etiology of myxomatous valve disease." Dev Biol 357(1): 152-164.
Duren, Z., X. Chen, R. Jiang, Y. Wang and W. H. Wong, (2017). "Modeling gene regulation from paired expression and chromatin accessibility data." Proc Natl Acad Sci U S A 114(25): E4914-E4923.
Fahmy, R. G., C. R. Dass, L.-Q. Sun, C. N. Chesterman and L. M. Khachigian, (2003). "Transcription factor Egr1 supports FGF-dependent angiogenesis during neovascularization and tumor growth." nature medecine 9(9).
Felker, A., K. D. Prummel, A. M. Merks, M. Mickoleit, E. C. Brombacher, J. Huisken, D. Panakova and C. Mosimann, (2018). "Continuous addition of progenitors forms the cardiac ventricle in zebrafish." Nat Commun 9(1): 2001.
Francois, M., P. Koopman and M. Beltrame, (2010). "SoxF genes: Key players in the development of the cardio-vascular system." The International Journal of Biochemistry & Cell Biology 42(3): 445-448.
Garcia-Caballero, M., A. R. Quesada, M. A. Medina and M. Mari-Beffa, (2018). "Fishing anti(lymph)angiogenic drugs with zebrafish." Drug Discov Today 23(2): 366-374.
Gardiner, M. R., M. M. Gongora, S. M. Grimmond and A. C. Perkins, (2007). "A global role for zebrafish klf4 in embryonic erythropoiesis." Mech Dev 124(9-10): 762-774.
Garside, V. C., A. C. Chang, A. Karsan and P. A. Hoodless, (2013). "Co-ordinating Notch, BMP, and TGF-beta signaling during heart valve development." Cell Mol Life Sci 70(16): 2899-2917.
Gassann, M., F. Casagranda, D. Orioli, H. Simon, C. Lai, R. Klein and G. Lemke, (1995). "Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor." Letters to nature 378.
Gaut, L., N. Robert, A. Delalande, M. A. Bonnin, C. Pichon and D. Duprez, (2016). "EGR1 Regulates Transcription Downstream of Mechanical Signals during Tendon Formation and Healing." PLoS One 11(11): e0166237.
Ghazvini-Boroujerdi, M., J. Clark, N. Narula, E. Palmatory, J.-M. Connolly, S. DeFelice, J. Xu, B. Jian, S. Hazelwood and R.-J. Levy, (2004). "Transcription factor Egr-1 in calcific aortic valve disease." J heart Valve Dis. 13(6): 894-903.
Giniger, E., S. Varnum and M. Ptashne, (1895). "Specific DNA Binding of GAL4, A Positive Regulatory Protein of Yeast." Cell 40: 767-774.
Goddard, L. M., A. L. Duchemin, H. Ramalingan, B. Wu, M. Chen, S. Bamezai, J. Yang, L. Li, M. P. Morley, T. Wang, M. Scherrer-Crosbie, D. B. Frank, K. A. Engleka, S. C. Jameson, E. E. Morrisey, T. J. Carroll, B. Zhou, J. Vermot and M. L. Kahn, (2017). "Hemodynamic Forces Sculpt Developing Heart Valves through a KLF2-WNT9B Paracrine Signaling Axis." Dev Cell 43(3): 274-289 e275.
Goishi, K., P. Lee, A. J. Davidson, E. Nishi, L. I. Zon and M. Klagsbrun, (2003). "Inhibition of zebrafish epidermal growth factor receptor activity results in cardiovascular defects." Mechanisms of Development 120(7): 811-822.
Guan, Y., Q. Zhu, D. Huang, S. Zhao, L. Jan Lo and J. Peng, (2015). "An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide." Sci Rep 5: 13370.
Guerra, A., R. F. Germano, O. Stone, R. Arnaout, S. Guenther, S. Ahuja, V. Uribe, B. Vanhollebeke, D. Y. Stainier and S. Reischauer, (2018). "Distinct myocardial lineages break atrial symmetry during cardiogenesis in zebrafish." Elife 7.
Guner-Ataman, B., N. Paffett-Lugassy, M. S. Adams, K. R. Nevis, L. Jahangiri, P. Obregon, K. Kikuchi, K. D. Poss, C. E. Burns and C. G. Burns, (2013). "Zebrafish second

heart field development relies on progenitor specification in anterior lateral plate mesoderm and nkx2.5 function." Development 140(6): 1353-1363.
Guo, S., Y. Zhang, T. Zhou, D. Wang, Y. Weng, L. Wang and J. Ma, (2017). "Role of GATA binding protein 4 (GATA4) in the regulation of tooth development via GNAI3." Sci Rep 7(1): 1534.
Haffter, P., M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt, D. A. Kane, J. Odenthal, F. J. M. van Eeden, Y.-J. Jiang, C.-P. Heisenberg, R. N. Kelsh, M. Furutani-Seiki, E. Vogelsang, D. Beuchle, U. Schach, C. Fabian and C. Nüsslein-Volhard, (1996). "The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio." Development 123: 1-36.
Harja, E., L. G. Bucciarelli, Y. Lu, D. M. Stern, Y. S. Zou, A. M. Schmidt and S. F. Yan, (2004). "Early growth response-1 promotes atherogenesis: mice deficient in early growth response-1 and apolipoprotein E display decreased atherosclerosis and vascular inflammation." Circ Res 94(3): 333-339.
Hassan, B. A., E. A. Albanna, S. M. Morsy, A. G. Siam, M. M. Al Shafie, H. F. Elsaadany, H. S. Sherbiny, M. Shehab and O. Grollmuss, (2015). "Nutritional Status in Children with Un-Operated Congenital Heart Disease: An Egyptian Center Experience." Front Pediatr 3: 53.
Heckel, E., F. Boselli, S. Roth, A. Krudewig, H. G. Belting, G. Charvin and J. Vermot, (2015). "Oscillatory Flow Modulates Mechanosensitive klf2a Expression through trpv4 and trpp2 during Heart Valve Development." Curr Biol 25(10): 1354-1361.
Heinz, S., C. Benner, N. Spann, E. Bertolino, Y. C. Lin, P. Laslo, J. X. Cheng, C. Murre, H. Singh and C. K. Glass, (2010). "Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities." Mol Cell 38(4): 576-589.
Helmlinger, G., R. V. Geiger, S. Schreck and R. M. Nerem, (1991). "Effects of Pulsatile Flow on Cultured Vascular Endothelial Cell Morphology." journal of biomechanical engineering 113: 123.
Hergenreider, E., S. Heydt, K. Treguer, T. Boettger, A. J. Horrevoets, A. M. Zeiher, M. P. Scheffer, A. S. Frangakis, X. Yin, M. Mayr, T. Braun, C. Urbich, R. A. Boon and S. Dimmeler, (2012). "Atheroprotective communication between endothelial cells and smooth muscle cells through miRNAs." Nat Cell Biol 14(3): 249-256.
Herwig, L., Y. Blum, A. Krudewig, E. Ellertsdottir, A. Lenard, H. G. Belting and M. Affolter, (2011). "Distinct cellular mechanisms of blood vessel fusion in the zebrafish embryo." Curr Biol 21(22): 1942-1948.
Hoffmann, J. and S. Kaplan, (2002). "The Incidence of Congenital Heart Disease." Journal of the American College of Cardiology 39(12): 1890-1900.
Hogan, B. M., R. Herpers, M. Witte, H. Helotera, K. Alitalo, H. J. Duckers and S. Schulte-Merker, (2009). "Vegfc/Flt4 signalling is suppressed by Dll4 in developing zebrafish intersegmental arteries." Development 136(23): 4001-4009.
Holler, K. L., T. J. Hendershot, S. E. Troy, J. W. Vincentz, A. B. Firulli and M. J. Howard, (2010). "Targeted deletion of Hand2 in cardiac neural crest-derived cells influences cardiac gene expression and outflow tract development." Dev Biol 341(1): 291-304.
Hove, J. R., R. W. Köster, A. S. Forouhar, G. Acevedo-Bolton, S. E. Fraser and M. Gharib, (2003). "Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis." Nature 421(6919): 172-175.
Howe, K., M. D. Clark, C. F. Torroja, J. Torrance, C. Berthelot, M. Muffato, J. E. Collins, S. Humphray, K. McLaren, L. Matthews, S. McLaren, I. Sealy, M. Caccamo, C. Churcher, C. Scott, J. C. Barrett, R. Koch, G. J. Rauch, S. White, W. Chow, B. Kilian, L. T. Quintais, J. A. Guerra-Assuncao, Y. Zhou, Y. Gu, J. Yen, J. H. Vogel, T. Eyre, S.

Redmond, R. Banerjee, J. Chi, B. Fu, E. Langley, S. F. Maguire, G. K. Laird, D. Lloyd, E. Kenyon, S. Donaldson, H. Sehra, J. Almeida-King, J. Loveland, S. Trevanion, M. Jones, M. Quail, D. Willey, A. Hunt, J. Burton, S. Sims, K. McLay, B. Plumb, J. Davis, C. Clee, K. Oliver, R. Clark, C. Riddle, D. Elliot, G. Threadgold, G. Harden, D. Ware, S. Begum, B. Mortimore, G. Kerry, P. Heath, B. Phillimore, A. Tracey, N. Corby, M. Dunn, C. Johnson, J. Wood, S. Clark, S. Pelan, G. Griffiths, M. Smith, R. Glithero, P. Howden, N. Barker, C. Lloyd, C. Stevens, J. Harley, K. Holt, G. Panagiotidis, J. Lovell, H. Beasley, C. Henderson, D. Gordon, K. Auger, D. Wright, J. Collins, C. Raisen, L. Dyer, K. Leung, L. Robertson, K. Ambridge, D. Leongamornlert, S. McGuire, R. Gilderthorp, C. Griffiths, D. Manthravadi, S. Nichol, G. Barker, S. Whitehead, M. Kay, J. Brown, C. Murnane, E. Gray, M. Humphries, N. Sycamore, D. Barker, D. Saunders, J. Wallis, A. Babbage, S. Hammond, M. Mashreghi-Mohammadi, L. Barr, S. Martin, P. Wray, A. Ellington, N. Matthews, M. Ellwood, R. Woodmansey, G. Clark, J. Cooper, A. Tromans, D. Grafham, C. Skuce, R. Pandian, R. Andrews, E. Harrison, A. Kimberley, J. Garnett, N. Fosker, R. Hall, P. Garner, D. Kelly, C. Bird, S. Palmer, I. Gehring, A. Berger, C. M. Dooley, Z. Ersan-Urun, C. Eser, H. Geiger, M. Geisler, L. Karotki, A. Kirn, J. Konantz, M. Konantz, M. Oberlander, S. Rudolph-Geiger, M. Teucke, C. Lanz, G. Raddatz, K. Osoegawa, B. Zhu, A. Rapp, S. Widaa, C. Langford, F. Yang, S. C. Schuster, N. P. Carter, J. Harrow, Z. Ning, J. Herrero, S. M. Searle, A. Enright, R. Geisler, R. H. Plasterk, C. Lee, M. Westerfield, P. J. de Jong, L. I. Zon, J. H. Postlethwait, C. Nusslein-Volhard, T. J. Hubbard, H. Roest Crollius, J. Rogers and D. L. Stemple, (2013). "The zebrafish reference genome sequence and its relationship to the human genome." Nature 496(7446): 498-503.

Hu, C.-Y., C. H. Yang, W. Y. Chen, C.-J. Huang, H.-Y. Huang, M. S. Chen and H.-J. Tsai, (2006). "Egr1 gene knockdown affects embryonic ocular development in zebrafish." Molecular vision 12: 1250-1258.

Huang, C. J., C. T. Tu, C. D. Hsiao, F. J. Hsieh and H. J. Tsai, (2003). "Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish." Dev Dyn 228(1): 30-40.

Huang da, W., B. T. Sherman and R. A. Lempicki, (2009). "Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists." Nucleic Acids Res 37(1): 1-13.

Huang da, W., B. T. Sherman and R. A. Lempicki, (2009). "Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources." Nat Protoc 4(1): 44-57.

Huddelson, J. P., S. Srinivasan, N. Ahmad and J. B. Lingrel, (2004). "Fluid shear stress induces endothelial KLF2 gene expression through a defined promoter region." Biol chem 385: 723-729.

Hurlstone, A. F. L., A.-P. G. Haramis, E. Wienholds, H. Begthel, J. Korving, F. van Eeden, E. Cuppen, D. Zivkovic, R. H. A. Plasterk and H. Clevers, (2003). "The Wnt/b-catenin pathway regulates cardiac valve formation." Nature 425(6958): 628-633.

Hwang, W. Y., Y. Fu, D. Reyon, M. L. Maeder, S. Q. Tsai, J. D. Sander, R. T. Peterson, J. R. Yeh and J. K. Joung, (2013). "Efficient genome editing in zebrafish using a CRISPR-Cas system." Nat Biotechnol 31(3): 227-229.

Isogai, S., M. Horiguchi and B. M. Weinstein, (2001). "The vascular anatomy of the developing zebrafish: an atlas of embryonic and early larval development." Dev Biol 230(2): 278-301.

Jiang, J., Y. S. Chan, Y. H. Loh, J. Cai, G. Q. Tong, C. A. Lim, P. Robson, S. Zhong and H. H. Ng, (2008). "A core Klf circuitry regulates self-renewal of embryonic stem cells." Nat Cell Biol 10(3): 353-360.

Jiao, K., H. Kulessa, K. Tompkins, Y. Zhou, L. Batts, H. S. Baldwin and B. L. Hogan, (2003). "An essential role of Bmp4 in the atrioventricular septation of the mouse heart." Genes Dev 17(19): 2362-2367.
Just, S., I. M. Berger, B. Meder, J. Backs, A. Keller, S. Marquart, K. Frese, E. Patzel, G. J. Rauch, C. Tubingen Screen, H. A. Katus and W. Rottbauer, (2011). "Protein kinase D2 controls cardiac valve formation in zebrafish by regulating histone deacetylase 5 activity." Circulation 124(3): 324-334.
Just, S., S. Hirth, I. M. Berger, M. C. Fishman and W. Rottbauer, (2016). "The mediator complex subunit Med10 regulates heart valve formation in zebrafish by controlling Tbx2b-mediated Has2 expression and cardiac jelly formation." Biochem Biophys Res Commun 477(4): 581-588.
Kaczynski, J., T. Cook and R. Urrutia, (2003). "Sp1- and Kruppel-like transcription factors." Genome Biol 4(2).
Kappas, N. C., G. Zeng, J. C. Chappell, J. B. Kearney, S. Hazarika, K. G. Kallianos, C. Patterson, B. H. Annex and V. L. Bautch, (2008). "The VEGF receptor Flt-1 spatially modulates Flk-1 signaling and blood vessel branching." J Cell Biol 181(5): 847-858.
Kawahara, A. and I. B. Dawid, (2001). "Critical role of biklf in erythroid cell differentiation in zebrafish." Current Biology 11: 1353-1357.
Kawakami, K., (2007). "Tol2: a versatile gene transfer vector in vertebrates." Genome Biol 8 (suppl1)(7).
Kawakami, K., K. Imanaka, M. Itoh and M. Taira, (2004). "Excision of the Tol2 transposable element of the medaka fish Oryzias latipes in Xenopus laevis and Xenopus tropicalis." Gene 338(1): 93-98.
Kawakami, K. and A. Shima, (1999). "Identification of the Tol2 transposase of the medaka fish Oryzias
latipes that catalyzes excision of a nonautonomous Tol2
element in zebrafish Danio rerio." Gene 240(1): 239-244.
Kawakami, K., A. Shima and N. Kawakami, (2000). "Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage." PNAS 97(21): 11403-11408.
Keegan, B. R., J. Feldman, G. Begemann, P. W. Ingham and D. Yelon, (2004). "Retinoic Acid Signaling Restricts the Cardiac Progenitor Pool." Sci Rep 307: 247-249.
Kerstjens-Frederikse, W. S., I. M. van de Laar, Y. J. Vos, J. M. Verhagen, R. M. Berger, K. D. Lichtenbelt, J. S. Klein Wassink-Ruiter, P. A. van der Zwaag, G. J. du Marchie Sarvaas, K. A. Bergman, C. M. Bilardo, J. W. Roos-Hesselink, J. H. Janssen, I. M. Frohn-Mulder, K. Y. van Spaendonck-Zwarts, J. P. van Melle, R. M. Hofstra and M. W. Wessels, (2016). "Cardiovascular malformations caused by NOTCH1 mutations do not keep left: data on 428 probands with left-sided CHD and their families." Genet Med 18(9): 914-923.
Kettleborough, R. N., E. M. Busch-Nentwich, S. A. Harvey, C. M. Dooley, E. de Bruijn, F. van Eeden, I. Sealy, R. J. White, C. Herd, I. J. Nijman, F. Fenyes, S. Mehroke, C. Scahill, R. Gibbons, N. Wali, S. Carruthers, A. Hall, J. Yen, E. Cuppen and D. L. Stemple, (2013). "A systematic genome-wide analysis of zebrafish protein-coding gene function." Nature 496(7446): 494-497.
Khachigian, L. M., (2006). "Early growth response-1 in cardiovascular pathobiology." Circ Res 98(2): 186-191.
Khachigian, L. M., K. R. Anderson, N. J. Halnon, M. A. J. Gimbrone, N. Resnick and T. Collins, (1997). "Egr-1 is activated in endothelial cells exposed to fluid shear stress and interacts with a novel shear-stress-response element in the PDGF A-chain promoter." Arterioscler Thromb Vasc Biol 10: 2280-2286.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley and S. L. Salzberg, (2013). "TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions." Genome Biol 14(4): R36.
Kim, I., T. L. Saunders and S. J. Morrison, (2007). "Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells." Cell 130(3): 470-483.
Kim, J. D., H. J. Kim, S. Koun, H. J. Ham, M. J. Kim, M. Rhee and T. L. Huh, (2014). "Zebrafish Crip2 plays a critical role in atrioventricular valve development by downregulating the expression of ECM genes in the endocardial cushion." Mol Cells 37(5): 406-411.
Knight, H. G. and D. Yelon (2016). Utilizing Zebrafish to Understand Second Heart Field Development. Etiology and Morphogenesis of Congenital Heart Disease: From Gene Function and Cellular Interaction to Morphology. T. Nakanishi, R. R. Markwald, H. S. Baldwin et al. Tokyo**:** 193-199.
Kok, F. O., M. Shin, C. W. Ni, A. Gupta, A. S. Grosse, A. van Impel, B. C. Kirchmaier, J. Peterson-Maduro, G. Kourkoulis, I. Male, D. F. DeSantis, S. Sheppard-Tindell, L. Ebarasi, C. Betsholtz, S. Schulte-Merker, S. A. Wolfe and N. D. Lawson, (2015). "Reverse genetic screening reveals poor correlation between morpholino-induced and mutant phenotypes in zebrafish." Dev Cell 32(1): 97-108.
Kormish, J. D., D. Sinner and A. M. Zorn, (2010). "Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease." Dev Dyn 239(1): 56-68.
Kotkamp, K., R. Mossner, A. Allen, D. Onichtchouk and W. Driever, (2014). "A Pou5f1/Oct4 dependent Klf2a, Klf2b, and Klf17 regulatory sub-network contributes to EVL and ectoderm development during zebrafish embryogenesis." Dev Biol 385(2): 433-447.
Kouzarides, T., (2007). "Chromatin modifications and their function." Cell 128(4): 693-705.
Kozera, B. and M. Rapacz, (2013). "Reference genes in real-time PCR." J Appl Genet 54(4): 391-406.
Kozyreva, S. V. H., L. L.; Poltarausa, A. B.; Domninskya, D. A.; Kisseleva L. L., (1999). "Structure of the human CpG-island-containing lung Kruppel-like factor
(LKLF) gene and its location in chromosome 19p13.11^13 locus." FEBS Letters 448(21816): 149-152.
Krueger, J., D. Liu, K. Scholz, A. Zimmer, Y. Shi, C. Klein, A. Siekmann, S. Schulte-Merker, M. Cudmore, A. Ahmed and F. le Noble, (2011). "Flt1 acts as a negative regulator of tip cell formation and branching morphogenesis in the zebrafish embryo." Development 138(10): 2111-2120.
Kumar, A., Z. Lin, S. SenBanerjee and M. K. Jain, (2005). "Tumor necrosis factor alpha-mediated reduction of KLF2 is due to inhibition of MEF2 by NF-kappaB and histone deacetylases." Mol Cell Biol 25(14): 5893-5903.
Kuo, C., E. Morrisey, R. Anandappa, K. Sigrist, M. Lu, M. Parmacek, C. Soudais and J. Leiden, (1997). "GATA4 transcription factor is required for ventral morphogenesis and heart tube formation." Genes Dev 11(8): 1048-1060.
Lander, A. D. and S. B. Selleck, (2000). "The Elusive Functions of Proteoglycans: In Vivo Veritas
." Journal of cell biology 148(2): 227-232.
Langmead, B. and S. L. Salzberg, (2012). "Fast gapped-read alignment with Bowtie 2." Nat Methods 9(4): 357-359.
Laurent, F., A. Girdziusaite, J. Gamart, I. Barozzi, M. Osterwalder, J. A. Akiyama, J. Lincoln, J. Lopez-Rios, A. Visel, A. Zuniga and R. Zeller, (2017). "HAND2 Target Gene Regulatory Networks Control Atrioventricular Canal and Cardiac Valve Development." Cell Rep 19(8): 1602-1613.

Lawson, N. D. and B. M. Weinstein, (2002). "In Vivo Imaging of Embryonic Vascular Development Using Transgenic Zebrafish." Developmental Biology 248(2): 307-318.
Lazic, S. and I. C. Scott, (2011). "Mef2cb regulates late myocardial cell addition from a second heart field-like population of progenitors in zebrafish." Dev Biol 354(1): 123-133.
Lee, T. I., S. E. Johnstone and R. A. Young, (2006). "Chromatin immunoprecipitation and microarray-based analysis of protein location." Nat Protoc 1(2): 729-748.
Lee, Y. M., J. J. Cope, G. E. Ackermann, K. Goishi, E. J. Armstrong, B. H. Paw and J. Bischoff, (2006). "Vascular endothelial growth factor receptor signaling is required for cardiac valve formation in zebrafish." Dev Dyn 235(1): 29-37.
Lemaire, P., O. Revelant, R. Bravo and P. Charnay, (1988). "Two mouse genes encoding potential transcription factors with identical DNA-binding domains are activated by growth factors in cultured cells." Proc Natl Acad Sci U S A 85: 4691-4695.
Lescroart, F., R. G. Kelly, J. F. Le Garrec, J. F. Nicolas, S. M. Meilhac and M. Buckingham, (2010). "Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo." Development 137(19): 3269-3279.
Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin and S. Genome Project Data Processing, (2009). "The Sequence Alignment/Map format and SAMtools." Bioinformatics 25(16): 2078-2079.
Liebling, M., A. S. Forouhar, R. Wolleschensky, B. Zimmermann, R. Ankerhold, S. E. Fraser, M. Gharib and M. E. Dickinson, (2006). "Rapid three-dimensional imaging and analysis of the beating embryonic heart reveals functional changes during development." Dev Dyn 235(11): 2940-2948.
Lin, C., C. Y. Lin, B. Zhou and C. Chen, (2012). "Partitioning the heart: Mechanisms of cardiac septation and valve development." Development 139: 3277-3299.
Lin, L., L. Bu, C. L. Cai, X. Zhang and S. Evans, (2006). "Isl1 is upstream of sonic hedgehog in a pathway required for cardiac morphogenesis." Dev Biol 295(2): 756-763.
Lindeman, L. C., A. H. Reiner, S. Mathavan, P. Alestrom and P. Collas, (2010). "Tiling histone H3 lysine 4 and 27 methylation in zebrafish using high-density microarrays." PLoS One 5(12): e15651.
Liu, H., E. J. Leslie, Z. Jia, T. Smith, M. Eshete, A. Butali, M. Dunnwald, J. Murray and R. A. Cornell, (2016). "Irf6 directly regulates Klf17 in zebrafish periderm and Klf4 in murine oral epithelium, and dominant-negative KLF4 variants are present in patients with cleft lip and palate." Hum Mol Genet 25(4): 766-776.
Lombardo, V. A., C. Otten and S. Abdelilah-Seyfried, (2015). "Large-scale zebrafish embryonic heart dissection for transcriptional analysis." J Vis Exp(95): 52087.
Long, Q., A. Meng, J. Wang, J. R. Jessen, M. J. Farrell and S. Lin, (1997). "GATA-1 expression pattern can be recapitulated in living transgenic zebrafish using GFP reporter gene." Development 124: 4105-4111.
Lyons, S. E., N. D. Lawson, L. Lei, P. E. Bennett, B. M. Weinstein and P. P. Liu, (2002). "A nonsense mutation in zebrafish gata1 causes the bloodless phenotype in vlad tepes." PNAS 99(8): 5454-5459.
Mably, J. D., C. G. Burns, J.-N. Chen, M. C. Fishman and M.-A. P. K. Mohideen, (2003). "heart of glass Regulates the Concentric Growth of the Heart in Zebrafish." Current Biology 13(24): 2138-2147.
Mably, J. D., L. P. Chuang, F. C. Serluca, M. A. Mohideen, J. N. Chen and M. C. Fishman, (2006). "santa and valentine pattern concentric growth of cardiac myocardium in the zebrafish." Development 133(16): 3139-3146.

MacGrogan, D., G. Luxan, A. Driessen-Mol, C. Bouten, F. Baaijens and J. L. de la Pompa, (2014). "How to make a heart valve: from embryonic development to bioengineering of living valve substitutes." Cold Spring Harb Perspect Med 4(11): a013912.
Mardis, E. R., (2007). "ChIP-seq: welcome to the new frontier." nature methods 4(8): 613.
McCormick, S. S., S. G. Eskin, L. V. McIntire, C. V. Teng, C.-M. LU, C. G. Russell and K. K. Chittu, (2001). "DNA microarray reveals changes in gene expression of shear stressed human umbilical vein endothelial cells." PNAS 98(16): 8955-8960.
McMahon, A. P., J. E. Champion, J. A. mcMahon and V. P. Sukhatme, (1990). "Developmental expression of the putative transcription factor Egr-1 suggests that Egr-1 and c-fos are coregulated in some tissues." Development 108: 281-287.
Meyer, D. and C. Birchmeier, (1995). "Multiple essential functions of neuregulin in development." Letters to nature 378: 386.
Milan, D. J., A. C. Giokas, F. C. Serluca, R. T. Peterson and C. A. MacRae, (2006). "Notch1b and neuregulin are required for specification of central cardiac conduction tissue." Development 133(6): 1125-1132.
Misfeldt, A. M., S. C. Boyle, K. L. Tompkins, V. L. Bautch, P. A. Labosky and H. S. Baldwin, (2009). "Endocardial cells are a distinct endothelial lineage derived from Flk1+ multipotent cardiovascular progenitors." Dev Biol 333(1): 78-89.
Mito, Y., J. G. Henikoff and S. Henikoff, (2005). "Genome-scale profiling of histone H3.3 replacement patterns." Nat Genet 37(10): 1090-1097.
Miyasaka, K. Y., Y. S. Kida, T. Banjo, Y. Ueki, K. Nagayama, T. Matsumoto, M. Sato and T. Ogura, (2011). "Heartbeat regulates cardiogenesis by suppressing retinoic acid signaling via expression of miR-143." Mech Dev 128(1-2): 18-28.
Morton, S. U., P. J. Scherz, K. R. Cordes, K. N. Ivey, D. Y. R. Stainier and D. Srivastava, (2008). "microRNA-138 modulates cardiac patterning during embryonic development." PNAS 105(46): 17830–17835.
Motoike, T., S. Loughna, E. Perens, B. L. Roman, W. Liao, T. C. Chau, C. D. Richardson, T. Kawate, J. Kuno, B. M. Weinstein, D. Y. Stainier and T. M. Sato, (2000). "Universal GFP Reporter for the Study of Vascular
Development." Wiley Liss 28: 75-81.
Nemer, G. N., M., (2002). "cooperative interaction between GATA5 and NF-ATc regulates endothelial-endocardial differenciation of cardiogenic cells." Development 129(17): 4045-4055.
Nicoli, S., C. Standley, P. Walker, A. Hurlstone, K. E. Fogarty and N. D. Lawson, (2010). "MicroRNA-mediated integration of haemodynamics and Vegf signalling during angiogenesis." Nature 464(7292): 1196-1200.
Novodvorsky, P., O. Watson, C. Gray, R. N. Wilkinson, S. Reeve, C. Smythe, R. Beniston, K. Plant, R. Maguire, M. K. R. A, S. Elworthy, F. J. van Eeden and T. J. Chico, (2015). "klf2ash317 Mutant Zebrafish Do Not Recapitulate Morpholino-Induced Vascular and Haematopoietic Phenotypes." PLoS One 10(10): e0141611.
Oates, A. C., J. C. Pratt, B. Vail, Y.-L. Yan, R. K. Ho, S. L. Johnson, J. H. Postlethwait and L. I. Zon, (2001). "The zebrafish klf gene family." Blood 98(6): 1792-1801.
Odelin, G., E. Faure, F. Coulpier, M. Di Bonito, F. Bajolle, M. Studer, J. F. Avierinos, P. Charnay, P. Topilko and S. Zaffran, (2018). "Krox20 defines a subpopulation of cardiac neural crest cells contributing to arterial valves and bicuspid aortic valve." Development 145(1).
Olave, I. A., S. L. Reck-Peterson and G. R. Crabtree, (2002). "Nuclear actin and actin-related proteins in chromatin remodeling." Annu Rev Biochem 71: 755-781.
Ortega, N., F.-E. L'Faqihi and J. Plouet, (1998). "Control of vascular endothelial growth factor angiogenic activity by the extracellular matrix." biology of the cell 90: 381-390.

Pagnozzi, L. A. and J. T. Butcher, (2017). "Mechanotransduction Mechanisms in Mitral Valve Physiology and Disease Pathogenesis." Front Cardiovasc Med 4: 83.
Palencia-Desai, S., M. S. Rost, J. A. Schumacher, Q. V. Ton, M. P. Craig, K. Baltrunaite, A. L. Koenig, J. Wang, K. D. Poss, N. C. Chi, D. Y. Stainier and S. Sumanas, (2015). "Myocardium and BMP signaling are required for endocardial differentiation." Development 142(13): 2304-2315.
Paranya, G., S. Vineberg, E. Dvorin, S. Kaushal, S. J. Roth, E. Rabkin, F. J. Schoen and J. Bischoff, (2001). "Aortic Valve Endothelial Cells Undergo Transforming
Growth Factor- -Mediated and Non-Transforming
Growth Factor- -Mediated Transdifferentiation
in Vitro." American Journal of Pathology 159(4): 1335.
Parmar, K. M., H. B. Larman, G. Dai, Y. Zhang, E. T. Wang, S. N. Moorthy, J. R. Kratz, Z. Lin, M. K. Jain, M. A. Gimbrone, Jr. and G. Garcia-Cardena, (2006). "Integration of flow-dependent endothelial phenotypes by Kruppel-like factor 2." J Clin Invest 116(1): 49-58.
Peal, D. S., S. N. Lynch and D. J. Milan, (2011). "Patterning and development of the atrioventricular canal in zebrafish." J Cardiovasc Transl Res 4(6): 720-726.
Pestel, J., R. Ramadass, S. Gauvrit, C. Helker, W. Herzog and D. Y. Stainier, (2016). "Real-time 3D visualization of cellular rearrangements during cardiac valve formation." Development 143(12): 2217-2227.
Piper, J., M. C. Elze, P. Cauchy, P. N. Cockerill, C. Bonifer and S. Ott, (2013). "Wellington: a novel method for the accurate identification of digital genomic footprints from DNase-seq data." Nucleic Acids Res 41(21): e201.
Postlethwait, J., I.-G. Woods, P. N. Hazelett, Y.-L. Yan, P. Kelly, F. Chu, H. Huang, A. Hill-Force and W. Talbot, (2000). "Zebrafish Comparative Genomics and the Origins of Vertebrate Chromosomes." Genome Res 10(1): 1890-1902.
Quinlan, A. R. and I. M. Hall, (2010). "BEDTools: a flexible suite of utilities for comparing genomic features." Bioinformatics 26(6): 841-842.
Raj, A. and G. McVicker, (2014). "The genome shows its sensitive side." Nat Methods 11(1): 39-40.
Rath, A., M. Glibowick, V. G. Nadeau, G. Chen and C. M. Debera, (2009). "Detergent binding explains anomalous SDS-PAGE migration of membrane proteins." PNAS 106(6): 1760-1765.
Rayner, B. S., G. A. Figtree, T. Sabaretnam, P. Shang, J. Mazhar, J. C. Weaver, W. N. Lay, P. K. Witting, S. N. Hunyor, S. M. Grieve, L. M. Khachigian and R. Bhindi, (2013). "Selective inhibition of the master regulator transcription factor Egr-1 with catalytic oligonucleotides reduces myocardial injury and improves left ventricular systolic function in a preclinical model of myocardial infarction." J Am Heart Assoc 2(4): e000023.
Riant, F., F. Bergametti, X. Ayrignac, G. Boulday and E. Tournier-Lasserve, (2010). "Recent insights into cerebral cavernous malformations: the molecular genetics of CCM." FEBS J 277(5): 1070-1075.
Rinn, J. L. and H. Y. Chang, (2012). "Genome regulation by long noncoding RNAs." Annu Rev Biochem 81: 145-166.
Rivera-Feliciano, J. and C. J. Tabin, (2006). "Bmp2 instructs cardiac progenitors to form the heart-valve-inducing field." Dev Biol 295(2): 580-588.
Romero-Calvo, I., B. Ocon, P. Martinez-Moya, M. D. Suarez, A. Zarzuelo, O. Martinez-Augustin and F. S. de Medina, (2010). "Reversible Ponceau staining as a loading control alternative to actin in Western blots." Anal Biochem 401(2): 318-320.

Rottbauer, W., S. Just, G. Wessels, N. Trano, P. Most, H. A. Katus and M. C. Fishman, (2005). "VEGF–PLC gamma1 pathway controls cardiac contractility in the embryonic heart." Genes Dev 19(1): 1624-1634.
Sangwung, P., G. Zhou, L. Nayak, E. R. Chan, S. Kumar, D. W. Kang, R. Zhang, X. Liao, Y. Lu, K. Sugi, H. Fujioka, H. Shi, S. D. Lapping, C. C. Ghosh, S. J. Higgins, S. M. Parikh, H. Jo and M. K. Jain, (2017). "KLF2 and KLF4 control endothelial identity and vascular integrity." JCI Insight 2(4): e91700.
Sauls, K., A. de Vlaming, B. S. Harris, K. Williams, A. Wessels, R. A. Levine, S. A. Slaugenhaupt, R. L. Goodwin, L. M. Pavone, J. Merot, J. J. Schott, T. Le Tourneau, T. Dix, S. Jesinkey, Y. Feng, C. Walsh, B. Zhou, S. Baldwin, R. R. Markwald and R. A. Norris, (2012). "Developmental basis for filamin-A-associated myxomatous mitral valve disease." Cardiovasc Res 96(1): 109-119.
Scherz, P. J., J. Huisken, P. Sahai-Hernandez and D. Y. Stainier, (2008). "High-speed imaging of developing heart valves reveals interplay of morphogenesis and function." Development 135(6): 1179-1187.
Schroeder, J. A., L. F. Jackson, D. C. Lee and T. D. Camenisch, (2003). "Form and function of developing heart valves: coordination by extracellular matrix and growth factor signaling." J Mol Med (Berl) 81(7): 392-403.
Scott, E. K., L. Mason, A. B. Arrenberg, L. Ziv, N. J. Gosse, T. Xiao, N. C. Chi, K. Asakawa, K. Kawakami and H. Baier, (2007). "Targeting neural circuitry in zebrafish using GAL4 enhancer trapping." Nat Methods 4(4): 323-326.
Sehnert, A. J., A. Huq, B. M. Weinstein, C. Walker, M. Fishman and D. Y. Stainier, (2002). "Cardiac troponin T is essential in sarcomere assembly and cardiac contractility." Nat Genet 31(1): 106-110.
Serbedzija, G. N., J.-N. Chen and M. C. Fishman, (1996). "Regulation in the heart field of zebrafish." Development 125: 1095-1101.
Shannon, P., A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, (2002). "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks." Genome Res 13(1): 2498-2504.
Sheer, N. and J. A. Campos-Ortega, (1999). "Use of the Gal4-UAS technique for targeted gene expression in the zebrafish." Mech Dev 80: 153–158.
Shiraia, M., K. Imanaka-Yoshidab, M. D. Schneiderc, R. J. Schwartzd and T. Morisakie, (2009). "T-box 2, a mediator of Bmp-Smad signaling, induced hyaluronan synthase 2 and Tgf 2 expression and endocardial cushion formation." PNAS 106(44): 18604–18609.
Smith, K. A., A. K. Lagendijk, A. D. Courtney, H. Chen, S. Paterson, B. M. Hogan, C. Wicking and J. Bakkers, (2011). "Transmembrane protein 2 (Tmem2) is required to regionally restrict atrioventricular canal boundary and endocardial cushion development." Development 138(19): 4193-4198.
Stainier, D. Y., Z. Kontarakis and A. Rossi, (2015). "Making sense of anti-sense data." Dev Cell 32(1): 7-8.
Stainier, D. Y., R. K. Lee and M. C. Fishman, (1993). "Cardiovascular development in the zebrafish
I. Myocardial fate map and heart tube formation." Development 119.
Stainier, D. Y. R., (2002). "Contribution du poisson zèbre à l'étude moléculaire du développement du cœur des vertébrés." médecine/sciences 18(4): 448-456.
Stainier, D. Y. R., B. Fouquet, J.-N. Chen, K. S. Warren, B. M. Weinstein, S. E. Meiler, M.-A. P. K. Mohideen, S. C. F. Neuhauss, L. Solnica-Krezel, A. F. Schier, F. Zwartkruis, D. L. Stemple, J. Malicki, W. Driever and M. C. Fishman, (1996). "Mutations affecting the formation and function of the cardiovascular system

in the zebrafish embryo." Development 123: 285-292.
Stankunas, K., C. T. Hang, Z. Y. Tsun, H. Chen, N. V. Lee, J. I. Wu, C. Shang, J. H. Bayle, W. Shou, M. L. Iruela-Arispe and C. P. Chang, (2008). "Endocardial Brg1 represses ADAMTS1 to maintain the microenvironment for myocardial morphogenesis." Dev Cell 14(2): 298-311.
Stanley, E. G., C. Biben, A. Elephanty, L. Barnett, F. Koentgen, L. Robb and R. P. Harvey, (2002). "Efficient cre-mediated deletion in cardiac progenitor cells conferred by a 3'UTR-ires-Cre allele of the homeobox gene Nkx2-5." Int. J. dev. Biol. 46: 431-439.
Steed, E., N. Faggianelli, S. Roth, C. Ramspacher, J. P. Concordet and J. Vermot, (2016). "klf2a couples mechanotransduction and zebrafish valve morphogenesis through fibronectin synthesis." Nat Commun 7: 11646.
Strubbe, G., C. Popp, A. Schmidt, A. Pauli, L. Ringrose, C. Beisel and R. Paro, (2011). "Polycomb purification by in vivo biotinylation tagging reveals cohesin and Trithorax group proteins as interaction partners." Proc Natl Acad Sci U S A 108(14): 5572-5577.
Stuckley, D., (1956). "Congenital heart defects following maternal rubella during pregnancy " Br Heart J. 4: 519-522.
Summerton, J., (1999). "Morpholino antisense oligomers: the case for an RNase H-independent structural type." Biochimica et Biophysica Acta 1489: 141-158.
Taylor, J. S., I. Braasch, T. Frickey, A. Meyer and Y. Van de Peer, (2002). "Genome Duplication, a Trait Shared by 22,000 Species of Ray-Finned Fish." Genome Res 13(1): 382-390.
Theodoris, C. V., M. Li, M. P. White, L. Liu, D. He, K. S. Pollard, B. G. Bruneau and D. Srivastava, (2015). "Human disease modeling reveals integrated transcriptional and epigenetic mechanisms of NOTCH1 haploinsufficiency." Cell 160(6): 1072-1086.
Thiel, G. and G. Cibelli, (2002). "Regulation of life and death by the zinc finger transcription factor Egr-1." J Cell Physiol 193(3): 287-292.
Thisse, C. and B. Thisse, (2008). "High-resolution in situ hybridization to whole-mount zebrafish embryos. ." Nat Protoc 3: 59-69.
Timmerman, L. A., J. Grego-Bessa, A. Raya, E. Bertran, J. M. Perez-Pomares, J. Diez, S. Aranda, S. Palomo, F. McCormick, J. C. Izpisua-Belmonte and J. L. de la Pompa, (2004). "Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation." Genes Dev 18(1): 99-115.
Trinh, S., (2004). "FN and myocardial migration."
Tripathi, S., M. O. Pohl, Y. Zhou, A. Rodriguez-Frandsen, G. Wang, D. A. Stein, H. M. Moulton, P. DeJesus, J. Che, L. C. Mulder, E. Yanguez, D. Andenmatten, L. Pache, B. Manicassamy, R. A. Albrecht, M. G. Gonzalez, Q. Nguyen, A. Brass, S. Elledge, M. White, S. Shapira, N. Hacohen, A. Karlas, T. F. Meyer, M. Shales, A. Gatorano, J. R. Johnson, G. Jang, T. Johnson, E. Verschueren, D. Sanders, N. Krogan, M. Shaw, R. Konig, S. Stertz, A. Garcia-Sastre and S. K. Chanda, (2015). "Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding." Cell Host Microbe 18(6): 723-735.
Trompouki, E., T. V. Bowman, L. N. Lawton, Z. P. Fan, D. C. Wu, A. DiBiase, C. S. Martin, J. N. Cech, A. K. Sessa, J. L. Leblanc, P. Li, E. M. Durand, C. Mosimann, G. C. Heffner, G. Q. Daley, R. F. Paulson, R. A. Young and L. I. Zon, (2011). "Lineage regulators direct BMP and Wnt pathways to cell-specific programs during differentiation and regeneration." Cell 147(3): 577-589.
Turner, J. and M. Crossley, (1999). "Mammalian Kruppel-like transcription factors: more than just a pretty finger." trends in biological science 24(6): 236.

Urasaki, A., G. Morvan and K. Kawakami, (2006). "Functional dissection of the Tol2 transposable element identified the minimal cis-sequence and a highly repetitive sequence in the subterminal region essential for transposition." Genetics 174(2): 639-649.
van Gelder, M. M., N. Roeleveld and H. Nordeng, (2011). "Exposure to non-steroidal anti-inflammatory drugs during pregnancy and the risk of selected birth defects: a prospective cohort study." PLoS One 6(7): e22174.
Verhoeven, M. C., C. Haase, V. M. Christoffels, G. Weidinger and J. Bakkers, (2011). "Wnt signaling regulates atrioventricular canal formation upstream of BMP and Tbx2." Birth Defects Res A Clin Mol Teratol 91(6): 435-440.
Vermot, J., A. S. Forouhar, M. Liebling, D. Wu, D. Plummer, M. Gharib and S. E. Fraser, (2009). "Reversing blood flows act through klf2a to ensure normal valvulogenesis in the developing heart." PLoS Biol 7(11): e1000246.
Walsh, E. C. and D. Y. Stainier, (2001). "UDP-glucose dehydrogenase required for cardiac valve formation in zebrafish." Science 293(5535): 1670-1673.
Walter, E. M., T. M. M. H. de By, R. Meyer and R. Hetzer, (2012). "The future of heart valve banking
and of homografts: perspective from
the Deutsches Herzzentrum Berlin." HSR Proceedings in Intensive Care and Cardiovascular Anesthesia 4(2): 97-108.
Wang, L., P. Zhang, Y. Wei, Y. Gao, R. Patient and F. Liu, (2011). "A blood flow-dependent klf2a-NO signaling cascade is required for stabilization of hematopoietic stem cell programming in zebrafish embryos." Blood 118(15): 4102-4110.
Wang, W., C. H. Ha, B. S. Jhun, C. Wong, M. K. Jain and Z. G. Jin, (2010). "Fluid shear stress stimulates phosphorylation-dependent nuclear export of HDAC5 and mediates expression of KLF2 and eNOS." Blood 115(14): 2971-2979.
Wang, Y., B. Wu, A. A. Chamberlain, W. Lui, P. Koirala, K. Susztak, D. Klein, V. Taylor and B. Zhou, (2013). "Endocardial to myocardial notch-wnt-bmp axis regulates early heart valve development." PLoS One 8(4): e60244.
Wani, M. A., S. E. Wert and J. B. Lingrel, (1999). "Lung Kruppel-like Factor, a Zinc Finger Transcription." journal of biological chemistry 274(30): 21180-21185.
White, R. M., A. Sessa, C. Burke, T. Bowman, J. LeBlanc, C. Ceol, C. Bourque, M. Dovey, W. Goessling, C. E. Burns and L. I. Zon, (2008). "Transparent adult zebrafish as a tool for in vivo transplantation analysis." Cell Stem Cell 2(2): 183-189.
Wild, R., A. Klems, M. Takamiya, Y. Hayashi, U. Strahle, K. Ando, N. Mochizuki, A. van Impel, S. Schulte-Merker, J. Krueger, L. Preau and F. le Noble, (2017). "Neuronal sFlt1 and Vegfaa determine venous sprouting and spinal cord vascularization." Nat Commun 8: 13991.
Wong, K. S., K. Rehn, S. Palencia-Desai, V. Kohli, W. Hunter, J. D. Uhl, M. S. Rost and S. Sumanas, (2012). "Hedgehog signaling is required for differentiation of endocardial progenitors in zebrafish." Dev Biol 361(2): 377-391.
Xue, Y., S. Gao and F. Liu, (2015). "Genome-wide analysis of the zebrafish Klf family identifies two genes important for erythroid maturation." Dev Biol 403(2): 115-127.
Yelon, D., (2001). "Cardiac Patterning and Morphogenesis in Zebrafish." Dev Dyn 222(3): 552–563.
Yuan, S., L. Zhao, Z. Sun and W. F. Marschlall, (2003). "Dissecting the functional interplay between the TOR pathway and the cilium in zebrafish)." Methods in enzymology 525, Cilia, Part B(Chapter 9): 166.
Zhang, L., J. Cho, D. Ptak and Y. F. Leung, (2013). "The role of egr1 in early zebrafish retinogenesis." PLoS One 8(2): e56108.

Zhang, W. and J. J. Bieker, (1998). "Acetylation and modulation of erythroid Kru¨ ppel-like factor (EKLF) activity by interaction with histone acetyltransferases." Proc Natl Acad Sci U S A 95: 9855-9860.
Zhang, Y., T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, C. Nusbaum, R. M. Myers, M. Brown, W. Li and X. S. Liu, (2008). "Model-based Analysis of ChIP-seq (MACS)." Genome Biol 9(9).
Zheng, X., C. Xu, A. Di Lorenzo, B. Kleaveland, Z. Zou, C. Seiler, M. Chen, L. Cheng, J. Xiao, J. He, M. A. Pack, W. C. Sessa and M. L. Kahn, (2010). "CCM3 signaling through sterile 20-like kinases plays an essential role during zebrafish cardiovascular development and cerebral cavernous malformations." J Clin Invest 120(8): 2795-2804.
Zhou, Z., D. R. Rawnsley, L. M. Goddard, W. Pan, X. J. Cao, Z. Jakus, H. Zheng, J. Yang, J. S. Arthur, K. J. Whitehead, D. Li, B. Zhou, B. A. Garcia, X. Zheng and M. L. Kahn, (2015). "The cerebral cavernous malformation pathway controls cardiac development via regulation of endocardial MEKK3 signaling and KLF expression." Dev Cell 32(2): 168-180.
Zhou, Z., A. T. Tang, W. Y. Wong, S. Bamezai, L. M. Goddard, R. Shenkar, S. Zhou, J. Yang, A. C. Wright, M. Foley, J. S. Arthur, K. J. Whitehead, I. A. Awad, D. Y. Li, X. Zheng and M. L. Kahn, (2016). "Cerebral cavernous malformations arise from endothelial gain of MEKK3-KLF2/4 signalling." Nature 532(7597): 122-126.
Zhu, D., Y. Fang, K. Gao, J. Shen, T. P. Zhong and F. Li, (2017). "Vegfa Impacts Early Myocardium Development in Zebrafish." Int J Mol Sci 18(2).

Résumé de thèse détaillé

Nathalie Faggianelli-Conrozier
Soutenance le 14 décembre 2018, IGBMC, Illkirch
Laboratoire Dr Julien VERMOT, CNRS UMR 7104 – Inserm U 964

Etude du rôle de trois facteurs de transcription dans le développement des valves cardiaques, en utilisant le poisson zèbre comme modèle animal.

Les valves cardiaques

Parce qu'elles permettent un pompage du sang à sens unique, les valves cardiaques sont indispensables au bon fonctionnement du cœur chez les Vertébrés. Ces fines membranes séparent les différentes chambres cardiaques (oreillette, "atrium" en anglais et ventricule, "ventricle" en anglais). Leur formation est très contrôlée au cours du développement embryonnaire. Cependant, il arrive que celle-ci soit défectueuse, et donc à l'origine de maladies cardiaques congénitales. Ces maladies représentent une des causes majeures de décès à la naissance.

L'étude de la formation des valves cardiaques constitue donc un champ de recherche majeur. En effet les mécanismes mis en jeu lors de leur formation sont encore mal connus.

Le poisson zèbre (*Danio rerio,* ou zebrafish*)*

Ce petit poisson d'eau douce, originaire d'Inde, est devenu un modèle d'étude très populaire en biologie du développement et en génétique. Sa petite taille, le développement externe et rapide de ses embryons, la transparence optique de ces derniers et la possibilité d'arrêter le cœur durant les premiers jours sans affecter le développement font de lui un organisme de choix. La figure 1 présente quelques stades de développement.

Le cœur du poisson zèbre est une « version simplifiée » du cœur humain. Le schéma de la figure 2 compare les deux anatomies cardiaques.

Son génome a été complètement séquencé et annoté. 69% des gènes codant pour des protéines ont un homologue[13] dans le génome humain. Le génome du zebrafish

[13] Homologue : on parle de gènes/protéines homologue ou d'homologie pour décrire une relation entre des gènes/protéines qui résultent d'une évolution divergente, sur la base de leur similarité/identité de séquences. On peut subdiviser ensuite en différentes catégories dont les paralogues qui désignent des gènes qui ont été dupliqués et sont présents en deux exemplaires chez une espèce par rapport à une autre.

comme celui de beaucoup de poissons a subi un évènement de duplication au cours de l'évolution et environ 15% des gènes humains ont plus qu'un paralogue chez le poisson.

Dans cette étude par exemple, le gène *Klf2* humain existe en deux exemplaires chez le poisson : *klf2a* et *klf2b*. Il est possible qu'au cours du développement les deux gènes aient évolué différemment et acquis des rôles différents.

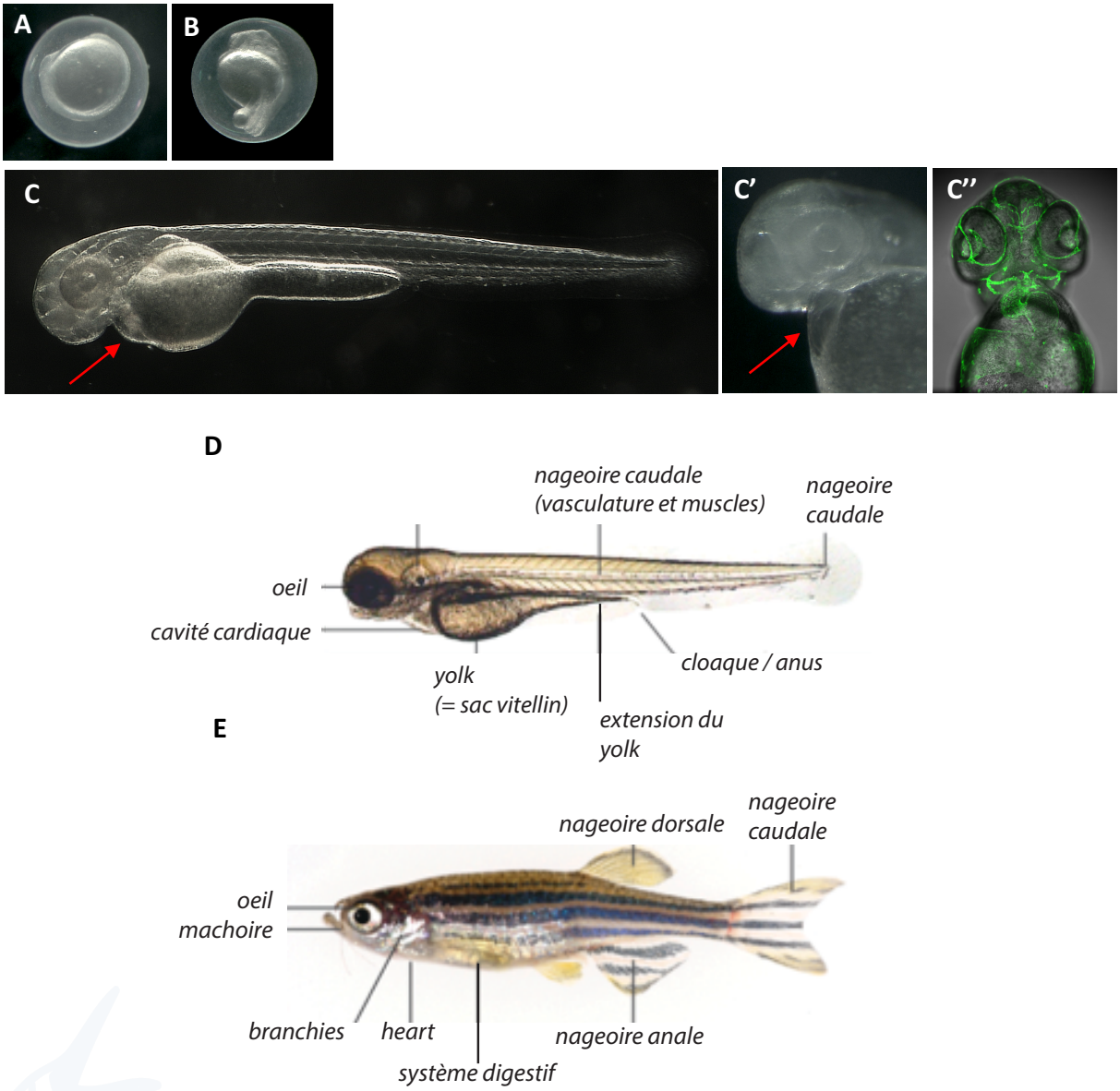


Figure 1 : Différents stades de développement du poisson zèbre (*Danio rerio*).
A: 12hpf, B: 24hfp, C, C' et C'': 48hpf, D: 96hpf, E': adulte. Les embryons sont observables sous un microscope classique (A, B, C, E x50 ; D: x250, E': 1x). Les flèches pointent la cavité cardiaque. La transparence optique des embryons est permise par l'ajout d'un composé chimique non-invasif, appelé PTU, au milieu de culture.
Des poissons transgéniques (= dont le génome a été modifié par l'insertion de séquences d'ADN exogènes) peuvent être créés afin de faciliter l'étude. C'': lignée transgénique exprimant la protéine GFP (Green fluorescent Protein, protéine isolée à l'origine de la méduse, largement utilisée en biologie) dans les cellules endothéliales du cœur et des vaisseaux sanguins.

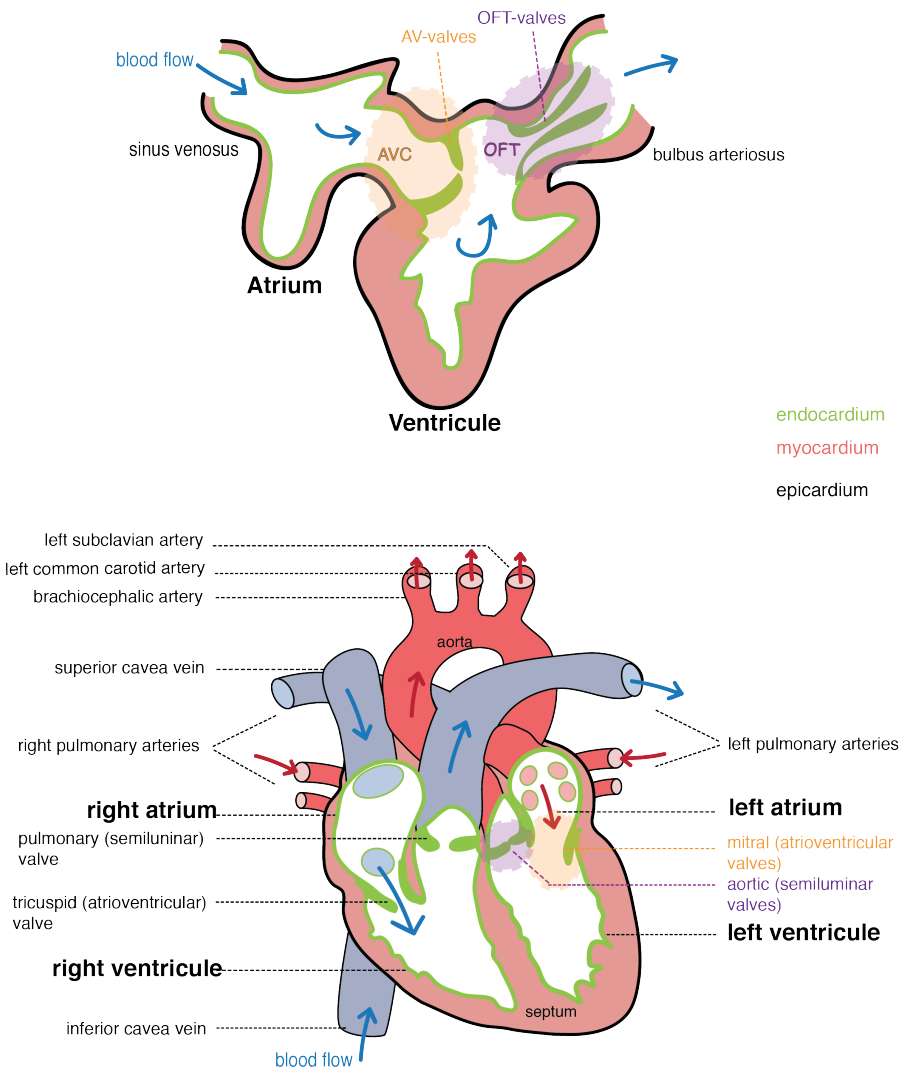


Figure 2: cœur de poisson zèbre à 96hpf (heures post fertilisation) en haut, et cœur humain adulte en bas. Le cœur du poisson ne possède que deux chambres : une oreillette et un ventricule contrairement au cœur humain.
Cette étude s'est focalisée sur la formation des valves dites « atrio-ventriculaires » (AV) entre l'oreillette et le ventricule du poisson, qui correspondent aux valves mitrales chez l'homme.

Le poisson zèbre est un organisme facilement manipulable génétiquement. De nombreux mutants (KO : Knock-Out, extinction totale du gène ou KI : Knock-In, extinction temporaire) peuvent être créés afin d'étudier les fonctions de ces gènes. La création de tels mutants peut se faire de manière aléatoire (en utilisant des agents mutagènes, ENU - N-ethyl-N-nitrosourea - ou UV) ou dirigée (en utilisant des technologies d'édition du génome appelées TALENS ou CRISPRs, de véritables ciseaux moléculaires insérant une coupure dans une séquence désirée).

Il est également possible de générer des lignées transgéniques permettant d'exprimer un fluorophore/fluorochrome (molécule chimique ou biologique capable d'émettre de la lumière par fluorescence après excitation) dans une population cellulaire spécifique et/ou pour suivre l'expression d'une protéine. Dans cette étude, nous avons notamment utilisé des lignées où les cellules endothéliales du cœur ou du système cardiovasculaire sont marquées en rouge (mCherry) ou vert (GFP) (panel C'' de la figure 1).

Mon projet de thèse

Grâce à des techniques d'imagerie *in vivo* à haute résolution et des techniques de biologie moléculaire, j'ai pu étudier le rôle des trois facteurs de transcription dans la formation des valves atrio-ventriculaires.

Ces trois facteurs sont Klf2a/Klf2b et Egr1. Dans de précédentes études, ils avaient déjà été démontrés comme actifs à 48hpf, dans la zone appelée AVC (cf figure 2) où vont se former les valves cardiaques. Chez l'homme, une mutation dans le gène *Egr1* entraine notamment des défauts dans les valves mitrales (valves calcifiées). En ce qui concerne *Klf2*, il n'a pas encore été trouvé impliqué dans une pathologie cardiaque.

Qu'est-ce qu'un facteur de transcription ?

C'est une protéine importante dans la cellule, nécessaire à l'initiation ou à la régulation de la transcription d'un gène (c'est-à-dire à l'étape permettant de passer de la lecture du gène dit exprimé, en intermédiaire messager ou ARNm, qui sera ensuite traduit en protéine). Ce sont en quelque sorte les architectes d'un programme génétique qui va permettre d'activer ou de réprimer (réduire leur expression) les gènes acteurs / ouvriers intervenant au cours de la formation de la valve. Tout un programme spatio-temporel d'expression génétique se met en place.

Notion de chromatine, d'épigénétique et présentation de techniques d'analyse de l'expression des gènes

L'ADN - support de l'information génétique - est compacté dans le noyau des cellules. Une compaction basale est permise par des complexes protéiques appelés nucléosomes. L'édifice formé ressemble à un collier de perles (les perles = nucléosomes). La chromatine est ensuite plus ou moins sur-enroulée selon l'état d'activation des gènes. On parle d'hétérochromatine (état super-enroulé, gènes non transcrits) et d'euchromatine (état plus lâche correspondant à des gènes / régions géniques actives ou transcrites) (voir figure 3).

Une méthode appelée ATAC-seq[14] permet de détecter les régions d'ADN actives à un moment donné dans une cellule. Une enzyme (Tn5) coupe l'ADN accessible entre les nucléosomes. Le séquençage au débit des fragments générés permet d'avoir une information sur ces gènes / régions activés.

[14] seq signifie « sequencing » ou séquençage : avec le progrès des techniques en génétique, il est aujourd'hui possible de séquencer rapidement de nombreux fragments d'ADN en même temps.

Il est aussi possible de découvrir les gènes régulés directement par des facteurs de transcription. On lie de manière covalente les protéines à l'ADN dans une première étape, puis l'ADN est fragmenté. Seuls les fragments d'ADN liés à la protéine d'intérêt, en utilisant un anticorps qui reconnait spécifiquement la protéine, sont récupérés, (on parle d'immunoprécipitation) avant d'être séquencés. Cette technique appelée Chromatin Immunoprecipitation (ChIP) requiert des anticorps spécifiques à chaque facteur étudié, ce qui n'est pas toujours évident à obtenir.

Le transcriptome, c'est-à-dire l'ensemble des intermédiaires ARN produits à un temps donné dans une cellule, peut aussi être étudié. Il donne une information sur les gènes activés ou réprimés et donc potentiellement en amont des gènes étudiés (si on compare le transcriptome d'un mutant par rapport au contrôle). On parle de RNAseq.

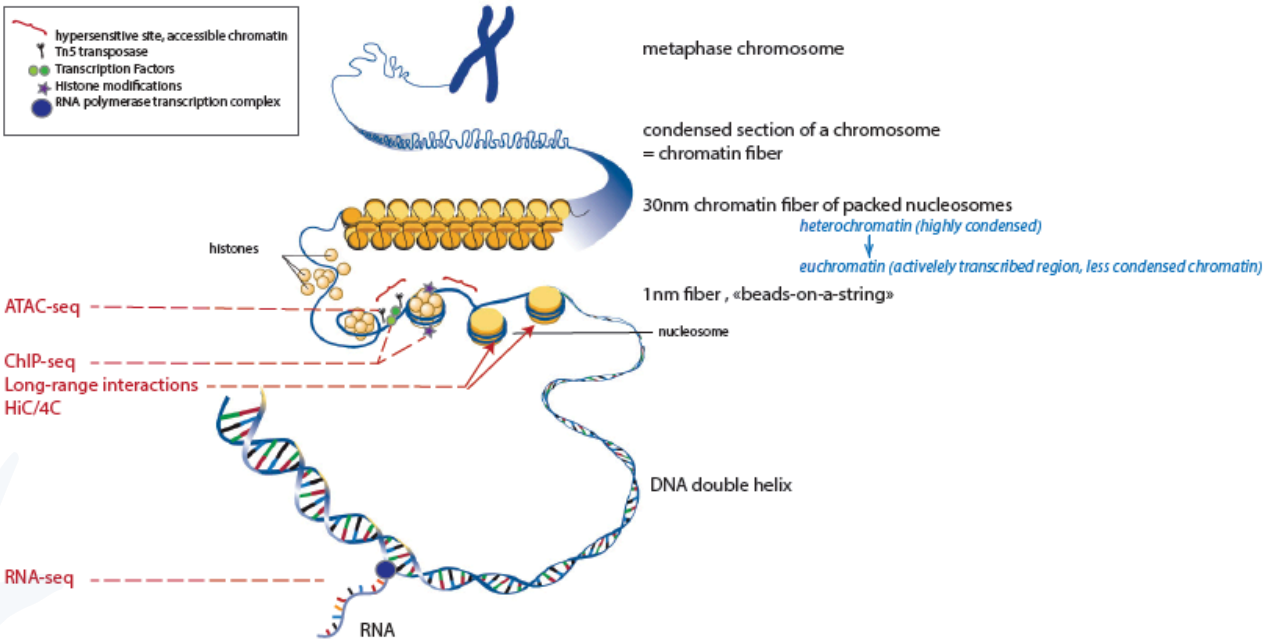


Figure 3: représentation schématique de la compaction de l'ADN depuis la double-hélice jusqu'au chromosome dit métaphasique. La transcription de l'information génétique en intermédiaire messager (ARN) est également représentée. En rouge sont citées quelques techniques utilisées en biologie moléculaire pour avoir accès à différents types d'informations (sur les protéines liées à l'ADN, les régions actives du génome, le transcriptome = l'ensemble des ARNs produits à un instant donné, information reliée aux gènes activés correspondants)

Notion de méchano-transduction

Le flux sanguin est un paramètre important dans la formation des valves. Les forces de frottement générées sur les parois cellulaires au niveau de l'AVC ("shear stress" ; flux oscillatoire sanguin entraînant des forces de friction) et le flux réserve ("reverse flow") présent à 48hpf avant l'apparition des valves, constituent un signal majeur. Toute une cascade d'évènements est ainsi activée afin de traduire ce message mécanique (d'où le terme de méchano-transduction) et d'activer au final des gènes dits « méchano-sensitifs » ou "blood-flow sensitive/responsive" en anglais) dont *klf2a*.

Dans cette étude, j'ai pu montrer que les gènes *egr1* et *klf2b* sont exprimés en réponse au flux sanguin à 48hpf dans l'AVC, comme il avait déjà été montré pour *klf2a*.

Les expériences mises en place

- Des mutants KO de ces gènes ont été générés et étudiés afin d'analyser les effets de la perte de ces gènes sur les valves -> observations des valves par microscopie confocale à 96hpf. La figure 4 présente les résultats obtenus pour les lignées mutantes pour *egr1*, *klf2a*-b.

Résultats: Klf2a, *klf2b* et *egr1* orchestrent l'activation d'un programme génétique spécifique contrôlant les mouvements cellulaires conduisant à la formation des valves et l'environnement moléculaire de la matrice extra-cellulaire (espace entre les deux couches de cellules du cœur le myocarde et l'endocarde, appelé "cardiac jelly" chez le zebrafish) dans lesquels les cellules formant les valves vont migrer/se déplacer pour se réorganiser.

- Etude de l'expression spatio-temporelle de ces trois facteurs -> In Situ Hybridization (ISH) Une sonde synthétique complémentaire à une partie de l'ARNm du gène d'intérêt est incubée sur des poissons préalablement fixés (par paraformaldéhyde, PFA). L'hybridation entraine une réaction chimique générant un signal lumineux visuel. La figure 5 présente un exemple de cette technique : l'expression du gène *egr1* dans le cœur d'embryons fixés à 48hpf.

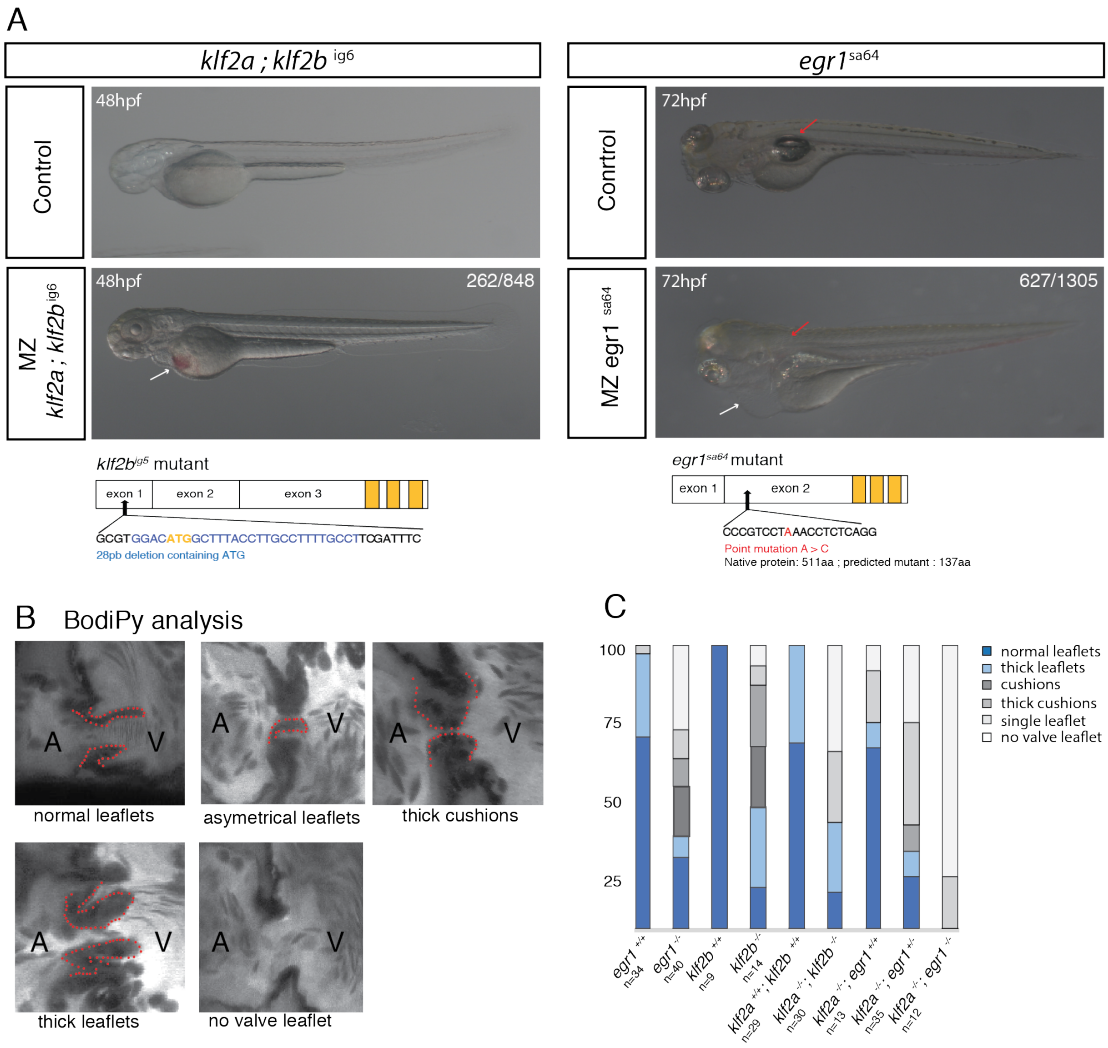


Figure 4 :

A : Phénotype des mutants étudiés. On peut noter la présence d'œdèmes cardiaques pour les deux mutants dans 31% des embryons mutés pour *klf2a* et *klf2b* et 50% des mutants pour *egr1*. De plus, une partie de ces derniers ne développe pas de vessie natatoire[15] (petite flèche rouge). La mutation dans les deux cas entraîne l'apparition d'un codon stop dans l'ADN ne permettant pas de produire la protéine native.

B : Observation de la structure des valves par microscopie confocale.

Cette analyse nécessite l'ajout de BodiPy, un colorant fluorescent utilisé pour marquer les membranes par ses propriétés hydrophobes, quelques heures avant l'expérience. (A : atrium/oreillette et V : ventricle/ventricule). On observe chez les mutants des valves[16] plus épaisses que la normales ("thick"), des valves asymétriques ("asymetrical leaflets") : l'une n'est pas formée et reste à l'état primitif de « coussin » ("cushion"). Dans le cas le plus sévère, aucune valve ne se forme ("no valve leaflet").

[15] C'est un organe qui permet au poisson de déterminer la profondeur à laquelle il flotte dans l'eau et lui permet ainsi de se mouvoir à la profondeur qu'il veut en ajustant sa densité à celle de l'eau dans laquelle il vit.

[16] On parle aussi de « leaflets » : une « valve » supérieure et une « valve » inférieure pouvant être distinguées

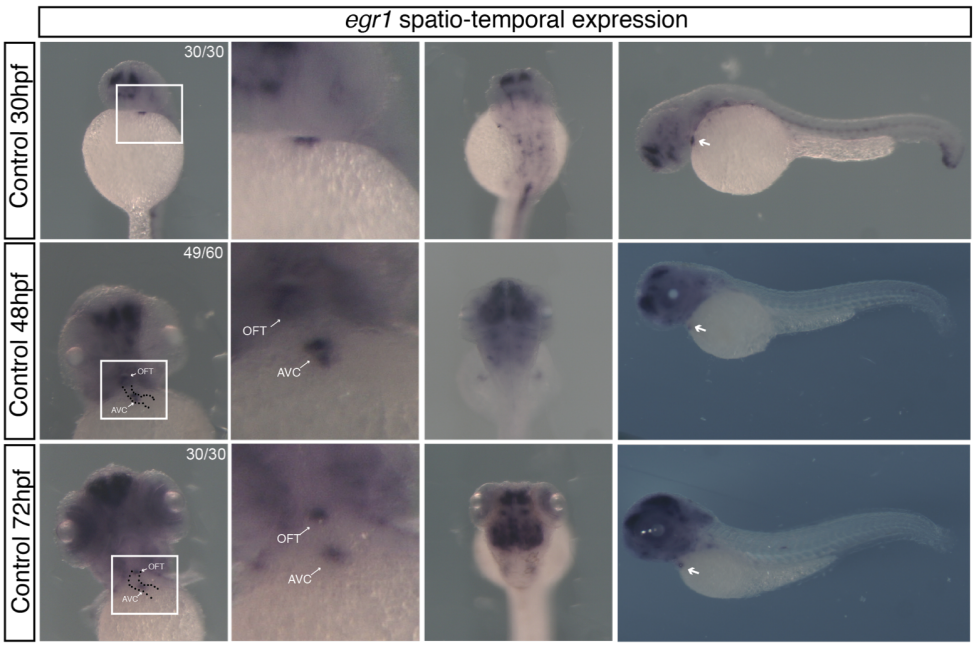


Figure 5: résultats d'une expérience d'In Situ Hybridization (ISH). Une molécule d'ADN simple brin complémentaire à l'ARN du gène d'intérêt (appelée sonde) est incubée sur des embryons préalablement fixée. Une réaction chemi-luminescence (signal visible de couleur violette) permet de détecter la localisation du complexe formé correspondant à la zone d'expression du gène étudié.
Une tâche violette révèle une expression du gène *egr1* sur la figure. On peut notamment noter une expression dans le cœur (plus particulièrement les zones AVC / OFT où se forment les valves. Egr1 est un gène majeur du développement, il s'exprimer aussi dans le cerveau et la rétine.

- Outils d'analyses de l'expression des gènes (analyses dites « Genome-wide »)
J'ai ensuite cherché à découvrir quels étaient les gènes situés en amont (ou "downstream" en anglais) de mes trois facteurs de transcription ; en d'autres termes découvrir quels étaient les gènes régulés (activés ou réprimés) à 48hpf dans l'AVC. Une approche combinant deux techniques moléculaires (mRNAseq & ATACseq) a été mise en place. Elles ne permettent pas de dire si les gènes trouvés sont des cibles directes de *klf2a*-b, *egr1* cependant (pour cela une expérience de ChIP par exemple serait nécessaire) simplement que ces gènes sont en amont dans le programme génétique ("genetic regulation pathway" en anglais).

Conclusion

Cette étude a permis de découvrir le rôle majeur de *klf2a*, *klf2b* et *egr1* dans la formation des valves de l'AVC. S'ils sont absents, cela conduit à des défauts sévères des valves (elles ne sont pas ou mal formées). Il est possible que ces gènes interagissent (l'un régulant l'autre et vice-versa) dans cette formation, mais d'autres recherches doivent être entreprises pour le prouver. L'approche couplée mRNA/ATAC-seq a permis de réaliser une image précise des réseaux géniques en jeu à 48hpf et des gènes qui se situent en amont des facteurs Klf2a-2b et Egr1 dans ces réseaux. Parmi ceux-ci, un gène important: *flt1*. Pour valider l'approche utilisée, une étude plus poussée d'un mutant pour le gène *flt1* a été réalisée.

Annex 1

Steed, E., N. Faggianelli, S. Roth, C. Ramspacher, J. P. Concordet and J. Vermot. "klf2a couples mechanotransduction and zebrafish valve morphogenesis through fibronectin synthesis." Nat Commun 7: 11646.

Author contribution on this paper: generation of anti-klf2a antibody to validate the KO-state of the klf2a^{ig4} line used in this study. Help for the *fn1b* Immuno-Fluorescence assays. Design o CRISPR tools to generate a *fn1b* zebrafish mutant (not published as we received a ZIRC mutant to finish the study before publication).

Annex 2

Publication of the following manuscript is ongoing:

Duchemin, A.-L., H. Vignes, N. Faggianelli and J. Vermot. "Piezo channels control mechanosensitive outflow tract valve development through the Hippo pathway effector Yap1 and the Klf2-Notch signaling axis".

Author contribution on this paper: generation of stable klf2b^{ig5} TALEN-mutant line and preliminary characterization of this line.

nature COMMUNICATIONS

ARTICLE

Received 15 Jul 2015 | Accepted 15 Apr 2016 | Published 25 May 2016

DOI: 10.1038/ncomms11646 OPEN

klf2a couples mechanotransduction and zebrafish valve morphogenesis through fibronectin synthesis

Emily Steed[1,2,3,4], Nathalie Faggianelli[1,2,3,4], Stéphane Roth[1,2,3,4], Caroline Ramspacher[1,2,3,4], Jean-Paul Concordet[5,6,7] & Julien Vermot[1,2,3,4]

The heartbeat and blood flow signal to endocardial cell progenitors through mechanosensitive proteins that modulate the genetic program controlling heart valve morphogenesis. To date, the mechanism by which mechanical forces coordinate tissue morphogenesis is poorly understood. Here we use high-resolution imaging to uncover the coordinated cell behaviours leading to heart valve formation. We find that heart valves originate from progenitors located in the ventricle and atrium that generate the valve leaflets through a coordinated set of endocardial tissue movements. Gene profiling analyses and live imaging reveal that this reorganization is dependent on extracellular matrix proteins, in particular on the expression of *fibronectin1b*. We show that blood flow and *klf2a*, a major endocardial flow-responsive gene, control these cell behaviours and *fibronectin1b* synthesis. Our results uncover a unique multicellular layering process leading to leaflet formation and demonstrate that endocardial mechanotransduction and valve morphogenesis are coupled via cellular rearrangements mediated by fibronectin synthesis.

[1] Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch 67404, France. [2] Centre National de la Recherche Scientifique, UMR7104, Illkirch 67404, France. [3] Institut National de la Santé et de la Recherche Médicale, U964, Illkirch 67404, France. [4] Université de Strasbourg, Illkirch 67404, France. [5] Muséum National d'Histoire Naturelle, 75231 Paris Cedex 05, France. [6] CNRS UMR 7196, 75231 Paris Cedex 05, France. [7] INSERM U1154, 75231 Paris Cedex 05, France. Correspondence and requests for materials should be addressed to J.V. (julien.vermot@igbmc.fr).

Tissue morphogenesis and organ formation depend upon the cooperative behaviour of groups of cells as well as the integration of chemical inputs generated in growing tissues. In addition, cells experience environmental mechanical stresses, such as pressure, strain and shear stress, due to tissue deformation and biological flows[1,2], which subsequently participate in driving morphogenetic movements[1,3–5]. Due to the early initiation of heart contraction, the formation of the cardiovascular system is intricately linked to its function. Indeed, flow forces are necessary for cardiac ballooning, trabeculation and epicardium formation with flow forces being necessary for cardiac ballooning[6], trabeculation[7,8] and epicardium formation[9]. In both the lymphatic and cardiac systems, valves serve to maintain unidirectional fluid flow and, pertinently, depend on their respective flows to form[10,11].

Congenital heart valve malformations constitute an important medical issue challenging our society. In recent years, it has become clear that most valve disease has its origin during embryogenesis, either as signs of abnormal developmental processes or the aberrant re-expression of fetal gene programs normally quiescent in adulthood[12,13]. These include mutations in genes encoding signalling factors (Notch1 and TGFβ)[14] for the aortic valves, and actin-binding proteins (Filamin A)[15] for the mitral valves. Diseased valves often also display defects in extracellular matrix (ECM) deposition[16], which plays an essential function in valve architecture[17,18]. Interestingly, studies of lymphatic valve formation have shown that the ECM proteins fibronectin and laminin are deposited during the initial stages of valve development[11,19], implicating ECM deposition in the earliest stages of the valve-forming process. The complex three-dimensional (3D) shape and constant motion of the heart, however, make imaging the morphogenetic events during cardiac valve development particularly challenging, although live imaging approaches are being continuously pioneered to observe endothelial cell behaviours in their mechanically active context[20–23].

In the heart, the atrioventricular (AV) valve emanates from the endocardial wall and is composed of endocardial cells (EdCs) and ECM components[12]. While blood flow has a broad influence on the shape and growth of EdCs[6], the oscillatory flow profile specific to the early AV canal (AVC) directs AV valve (AVV) formation by specifically increasing Krüppel-like factor 2a (*klf2a*) expression in the AVC[24,25]. As a transcription factor, *klf2a* expression likely allows EdCs to couple mechanotransduction to valve morphogenesis by activating a range of downstream target genes. The identity of such Klf2a target genes in valve-forming EdCs and the subsequent cellular behaviours induced, however, are unknown.

In this study, we investigated the cellular events taking place during valve formation and addressed their regulation by the flow-responsive transcription factor Klf2a. We show that valve formation proceeds via an initial stage of cell clustering followed by the appearance of cellular extensions towards the cardiac jelly. Subsequent global tissue remodelling events result in the appearance of ventricular and AVC-derived EdCs in the cardiac jelly overlying atrial-derived EdCs exposed to the lumen. Using transcriptomic analyses to highlight the transcriptional changes accompanying these temporally coordinated cell-movement events, we identified *fibronectin1b* as a key Klf2a- and flow-dependent factor necessary for the correct coordination of valvulogenesis. These data describe cell behaviour that is coordinated by the mechanical environment and mechanotransduction via Klf2a and ECM deposition.

Results

Endocardial cell contributions to the atrioventricular valve. AVV morphogenesis begins ~48 hours post fertilization (hpf). By 5 days post fertilization (dpf) a set of functional valve leaflets, extend into the AVC, occluding the passage of reversing blood flow[26–28]. To uncover the origins of the EdCs contributing to the AVV, we performed photoconversion experiments using the *Tg(fli1a:Gal4FF^{ubs}, UAS:kaede)* transgenic line, in which the photoconvertible protein kaede is expressed in the endothelial cells, including the endocardium. The exposure of kaede to 405 nm light results in an irreversible fluorescence conversion from fluorescent green to fluorescent red, enabling the development of cells labelled with the red form to be followed with respect to their green neighbours during AVV formation. As EdCs of the AVC can be identified by their positivity for Alcama[26], we used our knowledge of this staining pattern (Fig. 1a) to specifically photoconvert green kaede to its red form in the atrium and ventricle at 48 hpf. We then focused on the subsequent development of the superior AVC as it undergoes valve morphogenesis earlier than the inferior AVC[26]. Heart contraction was temporarily blocked using 2,3-butanedione-2-monoxime (BDM) to enable the photoconversion to be performed. Following photoconversion, heart contraction was resumed and embryos were allowed to develop under standard conditions until imaging at 80 hpf, enabling us to assess the contribution of EdCs from each region to the forming superior AVV leaflet (Fig. 1b). Atrial cells photoconverted at 48 hpf were seen lining the AVC lumen at 80 hpf (Fig. 1c), and were never present inside the cardiac jelly following the formation of multiple cell layers ($n = 12/12$). In contrast, photoconversion of cells in the ventricular region of the superior AVC at 48 hpf resulted in photoconverted cells in the cardiac jelly at 80 hpf ($n = 7/7$ Fig. 1d). When ventricular photoconversion was performed away from the ventricular inner curvature, no photoconverted cells were observed inside the cardiac jelly ($n = 3/3$; Fig. 1e). Thus, in addition to the cells of the AVC, EdCs from the atrium and ventricular inner curvature make significant and distinct contributions to the forming valve leaflets at 80 hpf (Fig. 1f). Furthermore, these distinct contributions are maintained at 120 and 168 hpf, suggesting mixing of cells from atrial and ventricular origins does not occur in the AVV at later stages (Supplementary Fig. 1). These findings implicate the ventricle and atrium as important sources of valve progenitors and suggest that stereotyped and coordinated cellular behaviours guide valve morphogenesis.

Cell density and protrusive activity in early valvulogenesis. To elucidate how cells of the atrium and ventricle reorganize to contribute to the developing valve leaflets, we characterized the organization of EdCs in the AVC in the moments preceding the appearance of multiple cell layers, beginning at 36 hpf. Using *Tg(fli:nlsmCherry)* embryos, in which the EdC nuclei are labelled, we observed nuclei to be relatively evenly spaced around the AVC at 36 hpf, before undergoing regional increases in cell density at 48 hpf (Fig. 2a–d and Supplementary Movie 1). Quantification of total cell numbers in the AVC showed a doubling of EdCs in the AVC between 36 and 48 hpf, (Fig. 2b). Photoconversion experiments suggest that cells move towards the AVC from the atrium between 36 and 48 hpf, while ventricular cells maintain their position at the ventricular inner curvature/exit of the AVC (Supplementary Fig. 2). In the absence of multi-layering, at this stage, this results in an increased density of cells within the AVC, particularly on the ventricular side. Incubation of embryos in BrdU between these stages demonstrated that ~60% of the atrial cells and ~40% of the cells from both the ventricle and the AVC proliferate during this time (Supplementary Fig. 2E,F), suggesting cell proliferation throughout the heart could be an important contributing factor to the increased cell density observed in the AVC at 48 hpf. Visualization of the superior leaflet alone enabled

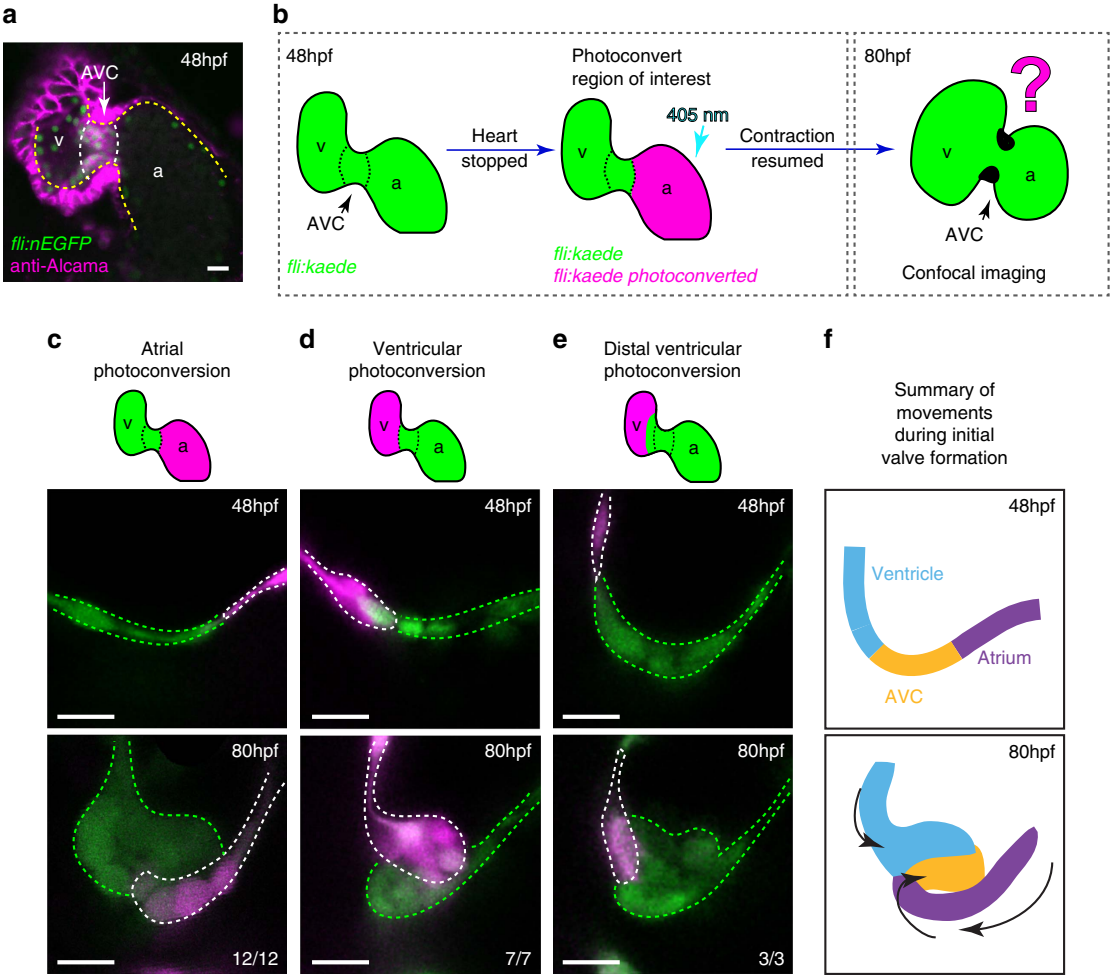


Figure 1 | The cellular contribution of heart chambers to emerging valve leaflets. (**a**) anti-Alcama immunofluorescence analysis shows Alcama-positive EdCs in the AVC of *fli:nEGFP* hearts (white dotted line), used to define the cardiac chambers at 48 hpf. Yellow dashed line delineates the endocardium from the myocardium, which also stains positively for Alcama. v = ventricle, a = atrium, AVC = atrioventricular canal. (**b**) The experimental set-up for the photoconversion studies. Heart contraction was stopped in *fli:kaede* embryos at 48 hpf, the region of interest exposed to 405 nm light to convert kaede from a green to red (shown here to be the atrium, in magenta) fluorescent form and heart contraction was then resumed until 80 hpf. Stopped hearts were imaged at 80 hpf by confocal microscopy. (**c**) EdCs present in the atrium and Alcama-negative at 48 hpf lined the lumen of the AVC at 80 hpf ($n = 12/12$, 3 experiments), while those in the ventricle contribute cells to the cardiac jelly (**d**; $n = 7/7$, 3 experiments). Ventricular cells outside the ventricular inner curvature (distal ventricle) do not enter the cardiac jelly at 80 hpf (**e**; $n = 3/3$, 2 experiments). In all cases only the superior valve leaflet is shown. (**f**) Schematic representation of the cellular contributions of each chamber to the emerging valve leaflet at 80 hpf. Atrial cells (purple) line the lumen of the AVC, while EdCs originating in the ventricular inner curvature (blue) and AVC (yellow) contribute cells to the cardiac jelly. Black arrows highlight the coordinated movements of the groups of cells. Scale bars, 10 μm.

the changes in cell density between 36 and 48 hpf to be seen more clearly (Fig. 2c,d). A region of increased cell density was particularly apparent on the ventricular side of the superior AVC at 48 hpf, containing 11 ± 3 cells ($n = 5$). The numbers of cells in this clustered region stayed the same at 56 hpf (11 ± 2 cells; $n = 5$; Fig. 2d and Supplementary Movie 2). Quantification of distances between neighbouring nuclei at these stages confirmed cells in the clustered region were indeed more closely packed together than those around the rest of the AVC (Fig. 2e and see Supplementary Movie 3) suggesting regional increases in cell density accompany the increases in cell numbers observed.

We recently demonstrated the sensitivity of EdCs in detecting flow forces and inducing the expression of the flow-responsive transcription factor *klf2a* (ref. 24). In light of this, we investigated the distribution of *klf2a* expressing cells, more specifically, within the AVC. Using a *klf2a* reporter line (*Tg(klf2a:H2BEGFP)*[24], we observed higher levels of GFP expression in those nuclei closest to the ventricle, compared with those on the atrial side of the AVC (Fig. 2f,g), corresponding to the region of cell clustering. Interestingly, protrusions were observed emanating from EdCs in this region towards the cardiac jelly at 48 hpf (Fig. 2h and Supplementary Movie 4). By 56 hpf, cells could be observed extending further into the cardiac jelly and by 72 hpf multiple layers of cells were present. To confirm the relevance of these observations in the beating heart, we performed fast confocal imaging of *Tg(kdrl:EGFP)* embryos and observed the same arrangement of cells at 72 hpf ($n = 3$; Supplementary Movie 5). Indeed when the heart is contracting, the connection of the EdCs within the cardiac jelly to the region of the heart wall from where they originate is clear (Supplementary Movie 5). These observations enable us to describe, for the first time, a cluster of EdCs close to the ventricular inner curvature at 48 hpf, from which cellular protrusions and movement of cells into the cardiac jelly originates.

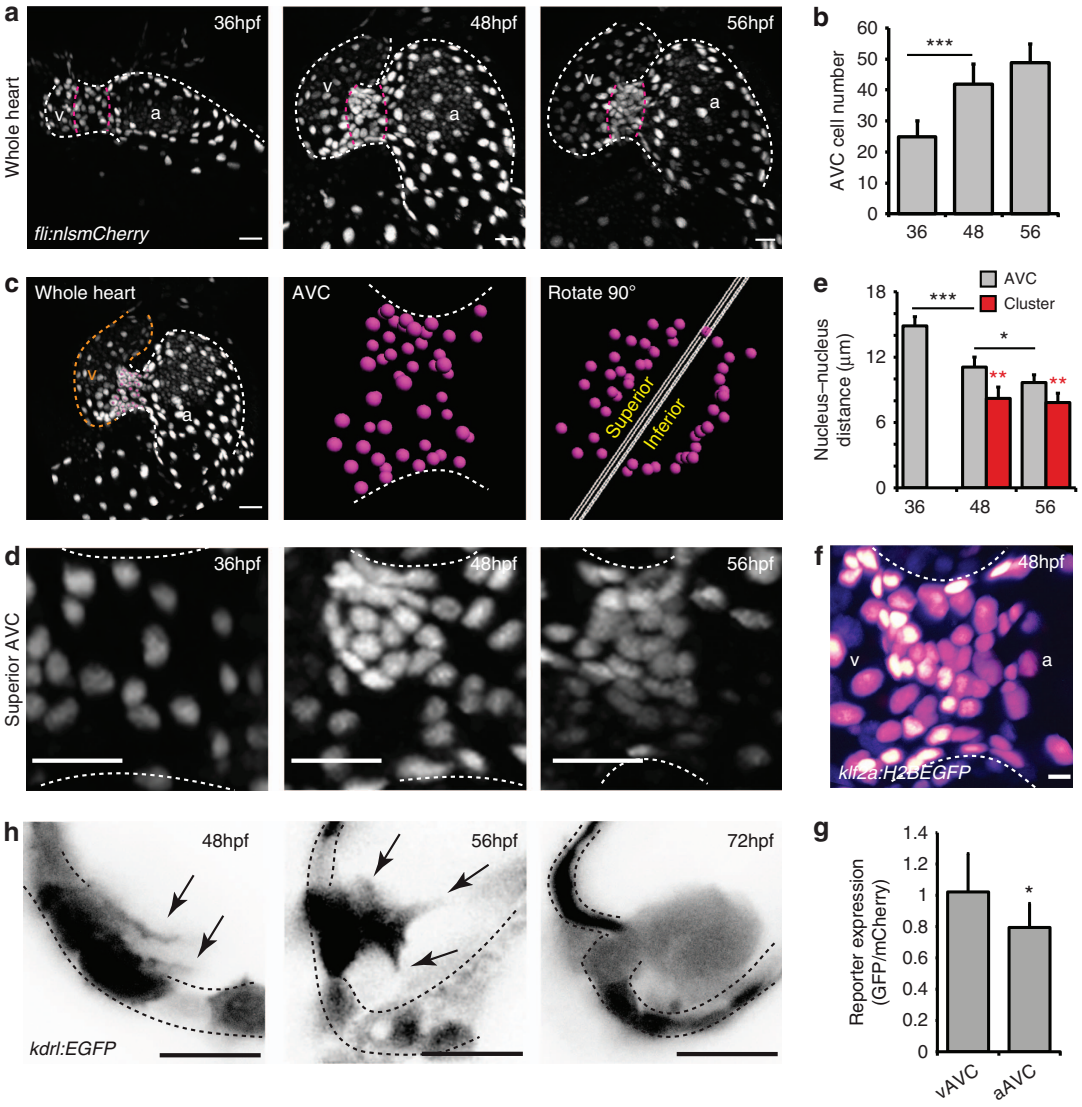


Figure 2 | EdCs move into the cardiac jelly from a region of increased cell density. (**a**) *fli:nlsmCherry* hearts imaged at early stages of heart development show the organization of EdCs between 36 and 56 hpf. White dotted lines outline the endocardium. Magenta dotted lines mark the AVC. (**b**) Total AVC cell number quantification at 36, 48 and 56 hpf. (**c**) Demonstration of how the superior AVC is defined and presented in **d**: The nuclei of AVC EdCs are defined (magenta spots) from images of the whole heart (orange/v = ventricle; white/a = atrium), rotated 90 °C using Imaris software and then divided into two parts, the superior AVC (to be the upper leaflet) and the inferior AVC (to be the lower leaflet). Orthogonal views of the superior AVC are then shown to display nuclei organization in **d**. (**d**) Regional increases in cell density can be seen on the ventricular side of the AVC from 48 hpf ($n = 5$ at each developmental stage). (**e**) Quantification of nucleus-nucleus distances within the AVC shows enhanced proximity of nuclei within the cluster region. No red bar at 36 hpf signifies the absence of a cluster at this stage. (**f**) *klf2a* expression pattern in the whole AVC at 48 hpf in the *klf2a:H2BEGFP* transgenic line shows enrichment of GFP signal on the ventricular side. The GFP signal is shown as FireLUT to aid visualization (white = highest intensity, black = no signal). (**g**) Relative reporter expression level in the ventricular (vAVC) and atrial (aAVC) regions of the AVC ($n = 10$). (**h**) Imaging of *kdrl:EGFP* hearts demonstrates the presence of protrusions (black arrows) extending from the region of clustered cells towards the cardiac jelly at 48 hpf. Cells emanating into the cardiac jelly can be seen from 56 hpf. Protrusions are also still visible (black arrows). Groups of cells are present in the cardiac jelly by 72 hpf. Inverted images of the *kdrl:EGFP* signal are shown to aid visualization of the protrusive structures. Black dotted lines mark the EdCs layer. Error bars in all graphs represent the s.d. Student's *t*-test $^{***}P < 0.005$, $^{**}P < 0.01$, $^{*}P < 0.05$. Scale bars, 10 μm; except in **f**: 2 μm.

Cardiac jelly cells present distinct characteristics. Analysis of the *Tg(kdrl:EGFP)* line at 80 hpf showed cells in the cardiac jelly to have a distinct mesenchymal-like morphology in comparison with those exposed to the lumen (Fig. 3a). Closer inspection of AVC morphology highlighted a deformation in the endocardial wall towards the cardiac jelly, which was wider towards the centre of the AVC (Fig. 3a–d). This morphology suggests a bending of the endocardial wall in response to the localized movement of EdCs into the cardiac jelly. In keeping with their mesenchymal-like morphology, junctional Cdh5 (VE-Cadherin) was lost between the neighbouring cells in the cardiac jelly at 72 hpf (Fig. 3e). Interestingly, BrdU-incorporation assays revealed the presence of BrdU-positive cells in the cardiac jelly, but not in the AVC wall between 56–72 hpf and 72–80 hpf (Fig. 3f and g). This suggests EdCs proliferate in the cardiac jelly, but do not enter it as a result of asymmetric cell division. Finally, analysis of *klf2a* reporter activity at 80 hpf, after multi-layering, showed EdCs exposed to the lumen expressed higher levels of *klf2a* activity than those in the cardiac jelly (Fig. 3h,i). This suggests *klf2a* is not continuously expressed in the EdCs that undergo multi-layering,

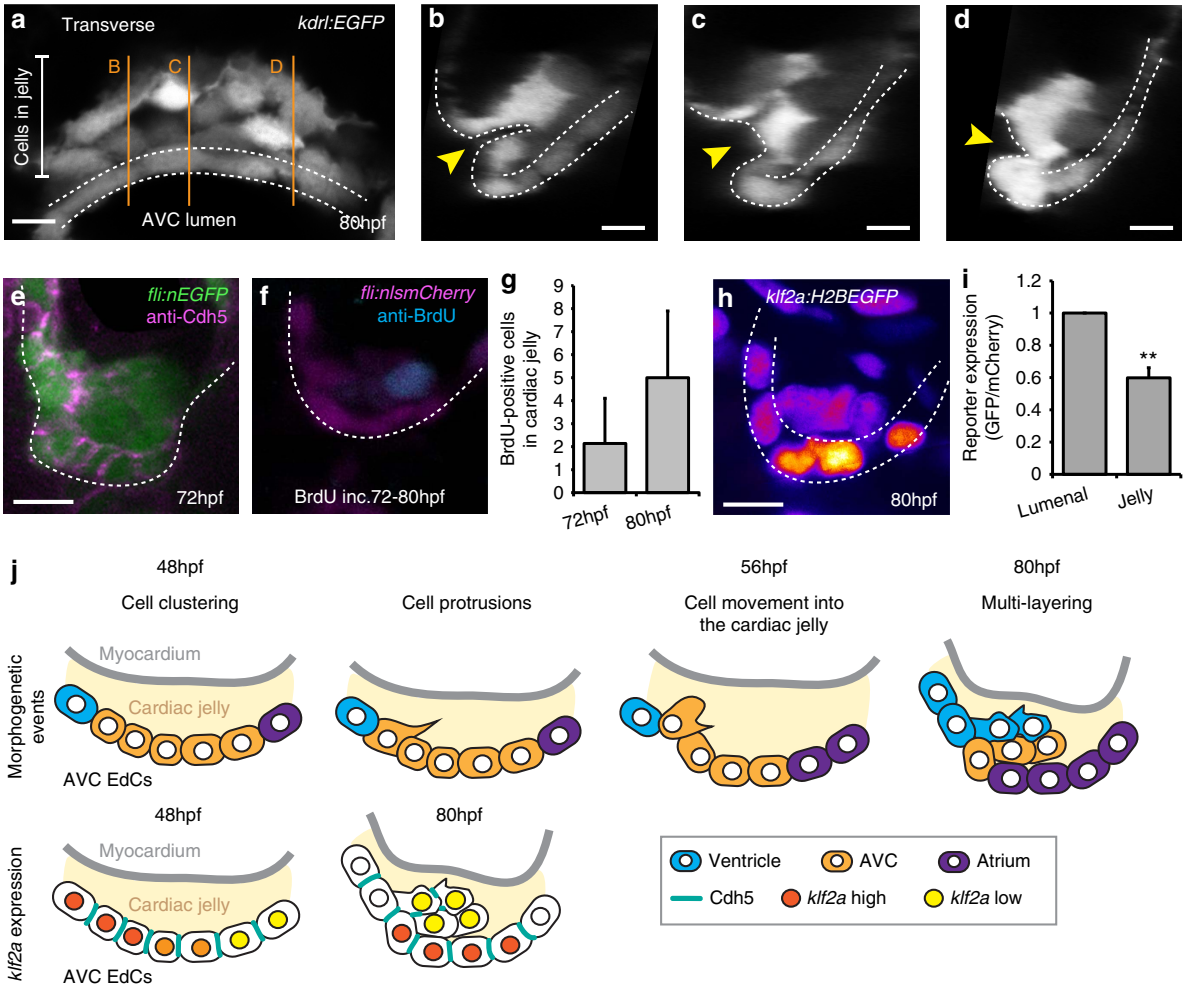


Figure 3 | EdCs display distinct characteristics inside the cardiac jelly. (**a**) Transverse section of a *kdrl:EGFP* AVC at 80 hpf shows distinctive cell shapes within the AVC. Cells in the cardiac jelly have a mesenchymal morphology compared with those lining the lumen (enclosed in white dotted lines). Orange lines mark the positions of images i–iii. (**b**–**d**) white dotted line outlines the luminal EdCs and the yellow arrowheads highlight the deformation of the ventricular wall towards the cardiac jelly. The protrusive morphology of the cells within the cardiac jelly can also be observed. (**e**) Immunofluorescence analysis shows downregulation of Cdh5 between EdCs in the cardiac jelly ($n=9$). (**f**) anti-BrdU immunofluorescence following BrdU incubation between 72–80 hpf shows a BrdU-positive cell in the cardiac jelly. Shown is a single *z* slice. (**g**) Quantification of total numbers of BrdU-positive cells in the cardiac jelly at 72 hpf (BrdU incubation = 56–72 hpf; $n=7$) and 80 hpf (BrdU incubation = 72–80 hpf; $n=5$) (**h**) Imaging of the *klf2a:H2BEGFP* reporter line shows the *klf2a* expression pattern in the AVC at 80 hpf (GFP signal shown as FireLUT). (**i**) Normalized relative reporter expression levels in the EdCs exposed to the AVC lumen (lumenal) and those inside the cardiac jelly (jelly) (all cells in the superior AVC analysed from three embryos). Error bar represents the standard deviation. $^{**}P<0.01$. (**j**) Model of the early stages of AVV formation. In the first series, cells are colour-coded to describe the morphogenetic events occurring during valve formation. Cells originating in the atrium (purple), AVC (yellow) and ventricle (blue) at 48 hpf and their relative positions at 80 hpf are shown. In the second series, Cdh5 and *klf2a* levels within the AVC are represented for the time points assessed. The myocardium is shown in grey and the cardiac jelly in pale yellow. Descriptions of the key stages at the relevant developmental time points are noted. Student's *t*-test $^{***}P<0.005$, $^{**}P<0.01$, $^{*}P<0.05$. Scale bars, 10 μm.

as they originate in the region of the AVC where *klf2a* expression is high and then enter the cardiac jelly where *klf2a* expression is low. Furthermore, cells originally expressing low levels of *klf2a* in the atrium[24] appear to initiate *klf2a* expression upon entering the AVC. These observations define two subsets of cells within the AVC following multi-layering; one *klf2a*^{low}, Cdh5^{low}, proliferative population within the cardiac jelly overlying a second *klf2a*^{high}, Cdh5^{high} population exposed to the blood flow. On the basis of these observations we propose a model to describe the early cellular events involved in AVV formation (Fig. 3j). EdCs cluster in a region of the AVC close to the ventricular inner curvature at 48 hpf, corresponding to the region of increased *klf2a* expression. EdCs in this clustered region extend protrusions and emanate into the cardiac jelly, initiating the coordinated morphogenetic movements that result in multiple layers of EdCs within the cardiac jelly by 80 hpf. Once inside the cardiac jelly, EdCs display a mesenchymal-like phenotype with reduced levels of Cdh5 and *klf2a*, and increased proliferation.

Transcriptional changes in early valvulogenesis. To elucidate how these early events are regulated, we sought to determine the gene expression profile activated at these early stages of valve formation. To do so, we extracted RNA from hearts dissected from 48 and 56 hpf *myl7:EGFP* embryos and performed transcriptome analysis using an Illumina sequencing platform (Fig. 4a). Between 31 and 65 million reads were generated for each RNA sample, of which, on an average, 67% could be mapped

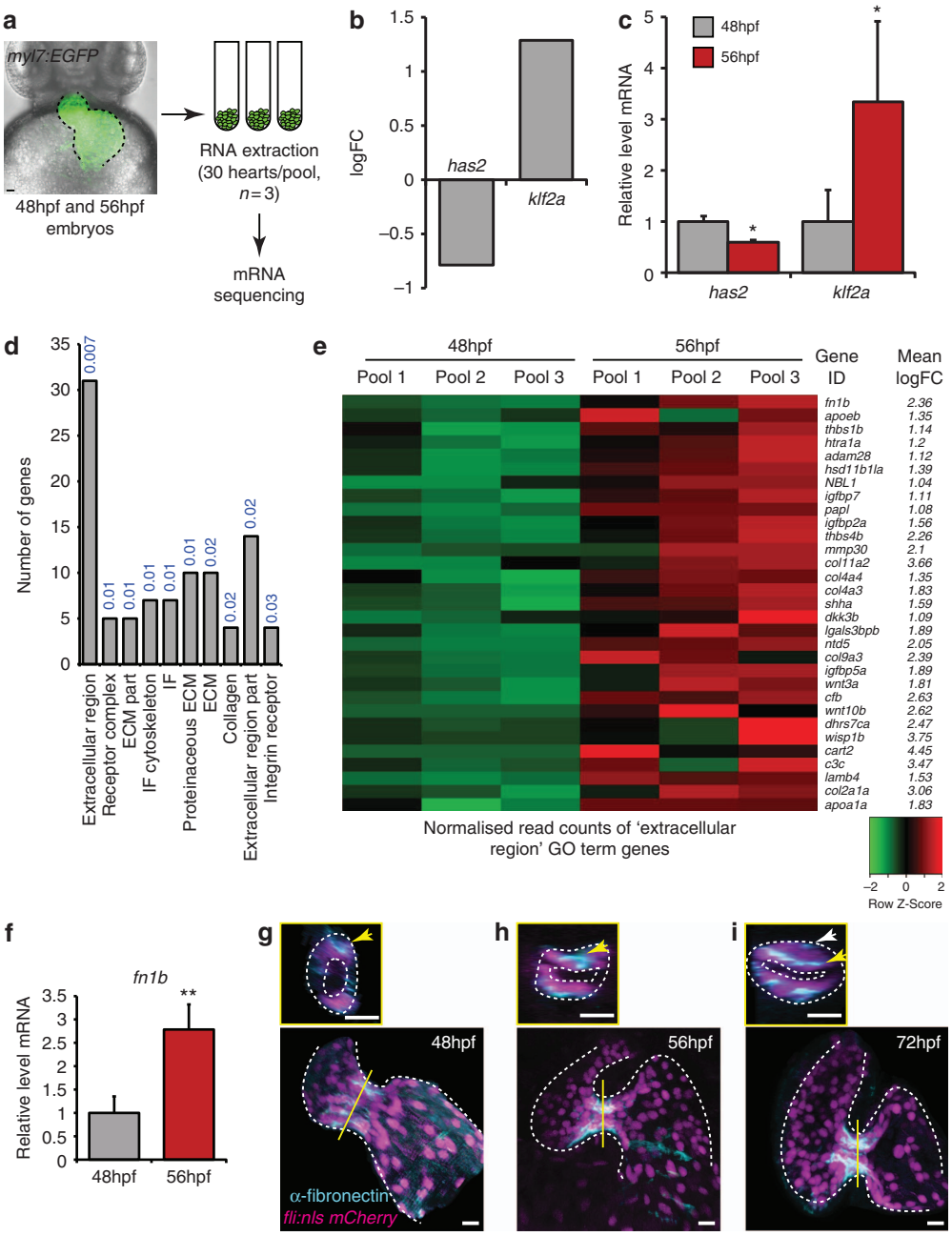


Figure 4 | Transcriptome analysis identifies increased ECM protein gene expression during initial stages of valve formation. (**a**) Experimental set-up for mRNA sequencing. RNA was extracted from hearts of *myl7:EGFP* embryos dissected at 48 and 56 hpf, in pools of 30 ($n=3$). (**b**) Downregulation of *has2* and upregulation of *klf2a* transcripts, confirmed by qPCR (**c**), validated our approach. (**d**) Gene Ontology (GO) analysis based on cellular compartment terms of upregulated genes highlighted a significant enrichment of ECM-related terms. *P* values describing the significance of each term enrichment are shown in blue. (**e**) Specific analysis of the most significantly enriched group (the 'extracellular region' term), highlighting, among others, elevated fibronectin1b expression between 48 and 56 hpf. (**f**) qPCR analysis of heart RNA confirms an increase in fn1b transcript levels between 48 and 56 hpf. (**g–i**) Immunofluorescence analysis shows fibronectin synthesis is localized to the AVC during early stages of valve formation ($n=8$ (48 hpf), $n=11$ (56 hpf) and $n=7$ (72 hpf)). Yellow lines mark the position of the transverse sections shown in small panels. Yellow arrows highlight enriched fibronectin-positive staining in the superior AVC. White arrow in small panel (**i**) points to fibronectin deposition on multiple cell layers. Student's *t*-test $^{***}P<0.005$, $^{**}P<0.01$, $^{*}P<0.05$. Scale bars, 10 μm.

onto the Zv9 assembly of the zebrafish genome. We identified 1,628 genes that were significantly, differentially expressed in the heart between 48 and 56 hpf (FDR < 0.05). Importantly, we saw a downregulation of *has2* and an upregulation of *klf2a*, confirming the reliability of our approach (Fig. 4b). Real-time quantitative PCR (qPCR) analysis of heart RNA confirmed the respective up and downregulation of these genes with time (Fig. 4c). Considering only those genes with a logFC > 1, we found 1,076 genes to be upregulated during these early stages of valve development. To assess the biological significance of these genes, we then performed Gene Ontology analysis using DAVID software[29]. By clustering genes based on cellular compartment annotations we saw a strong enrichment for ECM protein terms (Fig. 4d). Closer analysis of the 31 genes found in the 'extra-

cellular region' term demonstrated that, among others, *fibronectin1b* was significantly upregulated in the heart between 48 and 56 hpf (Fig. 4e). Given the importance of fibronectin deposition in lymph valve formation[19], we questioned whether it could also play a role in AVV formation. qPCR analysis of heart-derived RNA showed fibronectin1b is indeed expressed in the heart during this period of its development and confirmed the increase of transcript levels between 48 and 56 hpf (Fig. 4f). We next addressed the temporal and spatial expression of fibronectin within the heart at the protein level through immunofluorescence analysis. Interestingly, fibronectin protein is expressed specifically in the AVC at 48 hpf (Fig. 4g) with further enrichment by 56 hpf, in keeping with the mRNA sequencing data (Fig. 4h). More specifically, fibronectin is seen on the basal side of cells at both time points and is particularly enriched in the region of cell clustering (Fig. 4g,h). Fibronectin can also be observed on and between the multiple layers of cells that are present by 72 hpf (Fig. 4i), as has previously been shown in 105 hpf hearts[26]. This spatial and temporal expression pattern in the heart suggests a potentially relevant role of *fibronectin1b*/fibronectin in cardiac valve development.

Blood flow and *klf2a* alter fibronectin synthesis in the AVC. As blood flow is an important regulator of EdC behaviour and cardiac valve formation, we next wanted to assess whether changes in flow properties impacted fibronectin synthesis in the AVC. To do so, we first analysed the fibronectin staining pattern in silent heart ($sih^{-/-}$) mutant embryos, which completely lack heart contraction and blood flow[30], and saw that fibronectin was no longer detectable in the AVC at 48 hpf (Fig. 5a). As $sih^{-/-}$ mutants fail to form an AVV, we performed photoconversion experiments following injection of a morpholino specific for *troponin T2a* (*tnnt2a*), which is necessary for heart contraction and reliably mimics the $sih^{-/-}$ mutants[30], to determine whether the cell-movement events described above were impacted in the absence of heart contraction. Indeed, in *tnnt2a*MO hearts at 80 hpf, the photoconverted cells were found on the inner curvature of the ventricle and had failed to enter the cardiac jelly, as observed in age-matched controls (Supplementary Fig. 3A,B). To address the role of flow forces more specifically, we then altered blood viscosity and shear stress by lowering haematocrit content by injecting *gata1* and *gata2* morpholinos, as previously described[25]. In *gata1* morphants, where the fraction of reversing flow in the AVC at 48 hpf is increased and *klf2a* expression is high[24], strong fibronectin staining was observed in the AVC (Fig. 5b), while it was much reduced or absent in *gata2* morphants where the fraction of reversing flow in the AVC, and *klf2a* expression, is reduced (Fig. 5b). When atrial contraction was affected in *myh6* morphants (atrial specific myosin heavy chain, previously *amhc*), which also results in reduced *klf2a* expression[31] (Supplementary Fig. 3E,F), fibronectin deposition was also impaired ($n = 6/7$; Fig. 5b). Quantification of the proportion of the AVC positive for fibronectin confirmed these observations (Fig. 5c). Furthermore, fibronectin synthesis was significantly reduced, compared with controls, when 0.1% tricaine was used to stop heart contraction between 48 and 52 hpf and between 48 and 56 hpf (Fig. 5d). Interestingly, when the 0.1% tricaine was removed and heart contraction resumed at 52 hpf, fibronectin staining in the AVC at 56 hpf was restored (Fig. 5d,e). Taken together these observations suggest that the synthesis of fibronectin in the AVC is flow-dependent. To ascertain how conditions of altered flow may impact cellular organization during valve formation, we repeated our photoconversion experiments in the flow morphants described above. We observed cells in the cardiac jelly of *gata1* MO embryos, but this was greatly reduced in *gata2* MO and absent in *myh6* MO embryos, when compared with controls (Supplementary Fig. 3C,D). Furthermore, in *gata1* MOs, cells in the cardiac jelly appeared to be more disorganized than in controls (Supplementary Fig. 3C,D) suggesting AVC-specific fibronectin synthesis may be necessary for the correct organization of multiple cell layers before leaflet emanation.

AVC-specific fibronectin synthesis is necessary for valve formation. Considering the apparent flow-dependent nature of fibronectin synthesis in the AVC and its increased levels on the ventricular side of the superior AVC, where *klf2a* is most highly expressed, we reasoned that *fibronectin1b* could be a downstream target of Klf2a. As *klf2a* is elevated in response to the specific flow regime found within the AVC[24], such a target would enable *klf2a* to both respond to the mechanical environment of the AVC and impact the local environment of EdCs in a manner necessary for valve formation. Indeed, fibronectin staining was reduced in $trpp2^{-/-}$ mutants (Fig. 6a), which present defects in valvulogenesis and *klf2a* induction despite a normal flow regime in the AVC[24]. To address the effect of loss of *klf2a* and validate previous observations performed using morpholino-based approaches, we generated a mutant of *klf2a* using a TALEN approach targeting a sequence in the first exon of the *klf2a* gene (Supplementary Fig. 4A). Observations of valve morphology at 96 hpf, when valve leaflets can be seen extending into the lumen of the AVC in controls, demonstrated a range of valvular defects in *klf2a* mutants (Fig. 6b,c; $n = 25$ wild-type, $n = 46$ $klf2a^{-/-}$) despite there being no change in overall cell numbers (Supplementary Fig. 4B) or flow properties (Supplementary Fig. 4E) at 48 hpf. Approximately 10% of $klf2a^{-/-}$ embryos were missing any kind of valve structure. Identical analyses performed in *klf2a* morpholino-injected embryos showed a similar, but more severe, phenotype in the knock down (Fig. 6c) and all subsequent studies were performed with the *klf2a* mutant. To investigate origins of the valve defects observed in the $klf2a^{-/-}$ mutants, we examined these embryos during the early stages of valve formation described above. Analysis of cell organization showed $klf2a^{-/-}$ embryos had fewer cells clustering together in the superior AVC at 48 hpf than controls (Supplementary Fig. 4C). Quantification of cells within the cardiac jelly suggested multi-layering was impaired in ~40% of $klf2a^{-/-}$ hearts while the remaining $klf2a^{-/-}$ hearts presented elevated numbers of EdCs in the cardiac jelly at 72 hpf when compared with controls (Fig. 6d and Supplementary Fig. 4D). In those mutants where multi-layering occurred, however, the cells in the cardiac jelly appeared disorganized, with intracellular spaces between neighbouring EdCs, compared with the compact nature of the cells in this area in controls (Fig. 6d,e and Supplementary Movie 6). These data suggest that, in the absence of *klf2a* function, the cellular processes underlying the initiation of valve formation are perturbed and support observations made with *klf2a* morpholinos that *klf2a* expression is necessary for efficient valvulogenesis[25].

In situ hybridization and immunofluorescence analysis demonstrated that fibronectin is downregulated at both the mRNA and protein level in the majority of klf2a mutants (Fig. 7a,b). Furthermore, when we forced the overexpression of *klf2a* in all endothelial cells, we saw a spread of fibronectin synthesis outside the AVC, into the atrium and ventricle, in comparison with the AVC-specific localization observed in controls (Fig. 7c). This suggests that Klf2a, the expression of which is normally restricted to the AVC, is capable of driving the expression of fibronectin in the heart. Indeed, forced expression of *klf2a* was sufficient to rescue fibronectin synthesis in *gata2* and

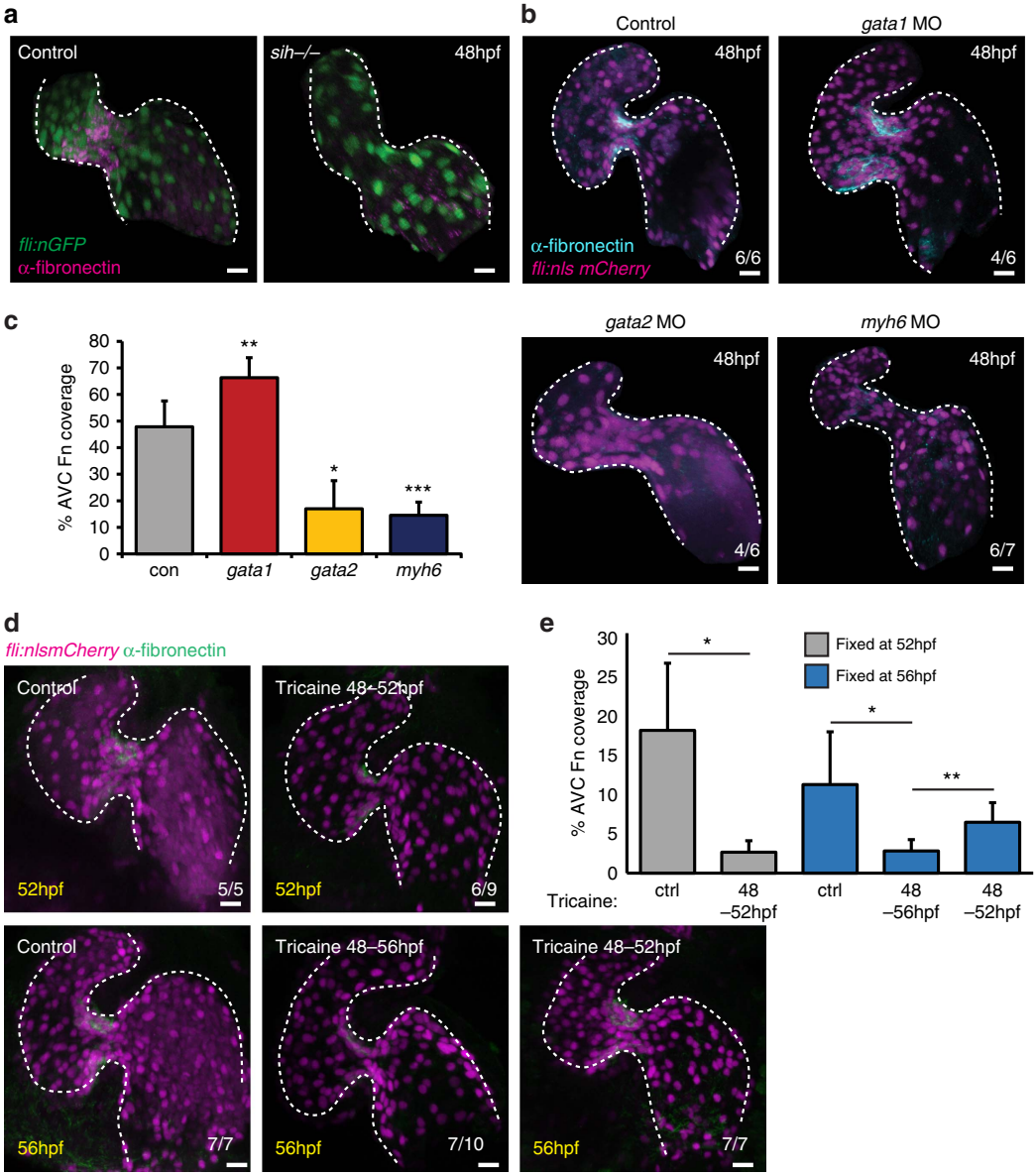


Figure 5 | AVC-specific fibronectin synthesis is dependent on blood flow forces. (**a**) Fibronectin staining (magenta) is lost in the absence of heart contraction (*fli:nEGFP*, *sih*$^{-/-}$ embryos) ($n=6/6$). (**b**) *fli:nlsmCherry* embryos injected with *gata1*, *gata2* or *myh6* morpholinos to alter blood flow forces were fixed at 48 hpf and anti-fibronectin immunofluorescence analysis was performed. Alterations in flow forces impacted fibronectin synthesis (cyan) in the AVC of *gata1*, *gata2* or *myh6* morphants. (**c**) Quantification of extent of fibronectin (Fn)-positive staining confirms an increase in *gata1* morphants and a reduction in both *gata2* and *myh6* morphants. (**d**) *fli:nlsmCherry* (magenta) embryos were incubated with 0.1% tricaine between 48–52 and 48–56 hpf to stop heart contraction. When necessary, heart contraction was restored at 52 hpf by washing out tricaine and embryos were left to develop under normal conditions until 56 hpf. Yellow font shows embryonic age at fixation. Anti-fibronectin immunofluorescence (green) shows the flow-responsive nature of fibronectin in the AVC. (**e**) Quantification of extent of fibronectin-positive staining confirms the restoration of fibronectin synthesis following restoration of heart contraction. Control (ctrl) samples were not incubated in 0.1% tricaine. Student's *t*-test $^{***}P<0.005$, $^{**}P<0.01$, $^{*}P<0.05$. Scale bars, 10 μm.

myh6 morphant embryos (Supplementary Fig. 5G). To quantify the efficacy of valve progenitors in undergoing multi-layering in the absence of fibronectin, a vital process in the formation of the heart valve, we used a *fn1b*-specific morpholino to deplete fibronectin in the AVC (Supplementary Fig. 5A). Care was taken to select embryos presenting no morphological defects following *fn1b*MO injection (Supplementary Fig. 5B), ensuring that *klf2a* reporter activity, AVC cell number and the flow velocity profile at 48 hpf (Supplementary Fig. 5D–F), as well as levels of p53 mRNA (Supplementary Fig. 5C), were not significantly changed in the morphants studied. Cell clustering and multi-layering were dramatically impaired in the absence of *fn1b*, with few or no cells in the cardiac jelly in the majority of cases (Fig. 7d,e) indicating the importance of fibronectin deposition in the early stages of valve development. At 96 hpf, large groups of EdCs and reversing blood flow were visible in the AVC of *fn1b*MO embryos in contrast to the efficient valve leaflets present in controls (Fig. 7f). To confirm these observations, we analysed the valve shape of *fn1b*$^{sa553-/-}$ embryos at 96 hpf. Similar to *fn1b* knock down, we found that all the fn1b mutants had abnormal valves (9/9). The majority of these embryos displayed large blocks of cells occluding the AVC or thick leaflets as sometimes also described in the *klf2a* mutants (Fig. 7g). These data suggest *klf2a* and *fn1b* expression are necessary for the movement of EdCs into

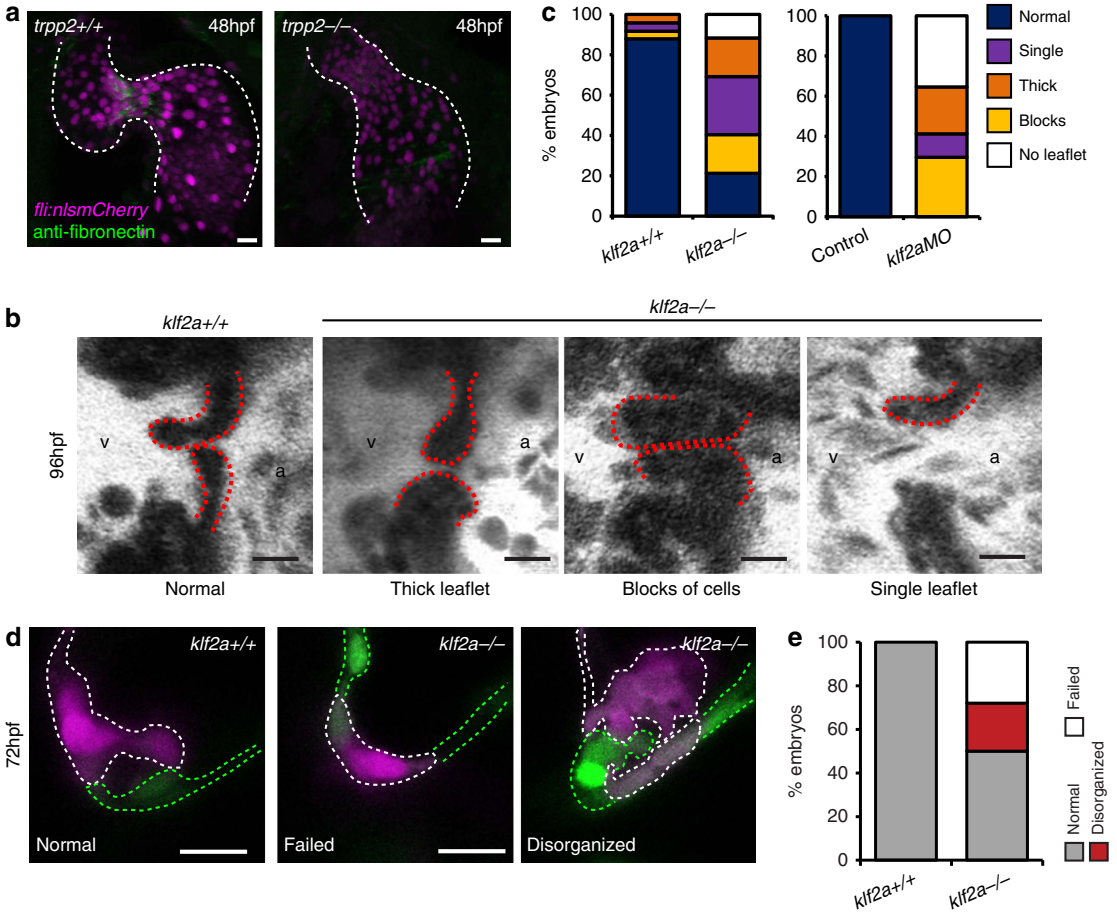


Figure 6 | Characterization of *klf2a*$^{-/-}$ mutants. (**a**) anti-fibronectin staining (green) is reduced in the AVC of *trpp2*$^{-/-}$ mutants (12/12 *trpp2*$^{+/+}$, 9/19 *trpp2*$^{-/-}$). (**b**,**c**) A range of valvular defects were observed in *klf2a*$^{-/-}$ mutant embryos at 96 hpf ($n=25$ *klf2a*$^{+/+}$, $n=48$ *klf2a*$^{-/-}$), similar to those observed in *klf2a*MO-injected embryos (**e**; $n=20$ control, $n=17$ *klf2a*MO). (**d**) Photoconversion of AVC EdCs at 48 hpf shows the organization of cells within the cardiac jelly to be affected by loss of *klf2a* (**d**,**e**; $n=21$ *klf2*$^{+/+}$, $n=32$ *klf2a*$^{-/-}$) at 72 hpf. Scale bars, 10 μm.

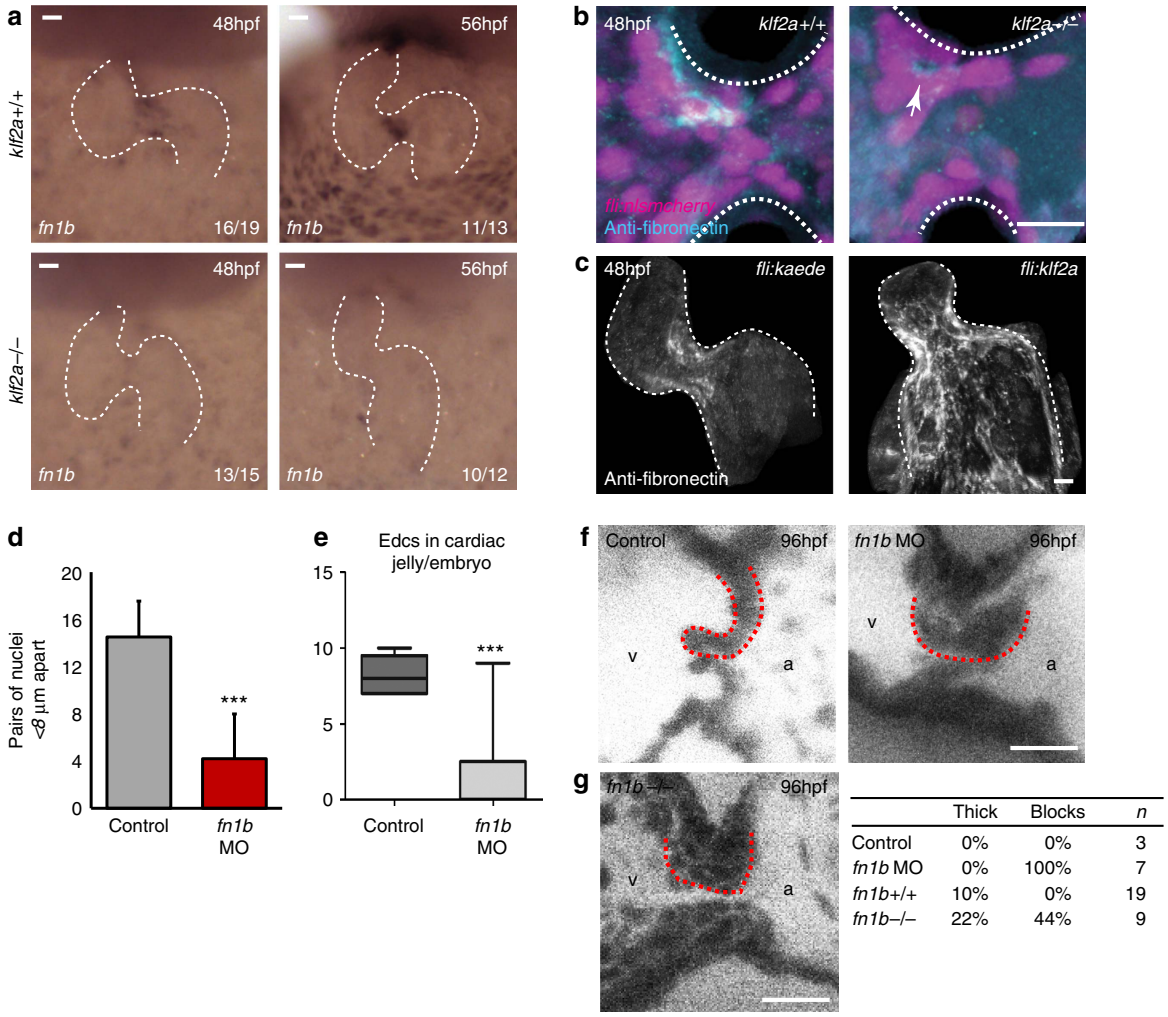
the cardiac jelly and their correct organization within. Thus mechanically-induced *klf2a* expression coordinates the morphogenetic events necessary for valve formation via the regulation of *fibronectin1b* expression and localized fibronectin synthesis within the AVC.

Discussion

Using *in vivo* imaging technologies and cellular scale 3D analysis, we have identified three successive steps highlighting the cellular processes associated with heart valve morphogenesis: (1) the first signs of valve formation correspond to a regional increase in cell density in the superior AVC and localized fibrillogenesis; (2) the first sign of cellular invasion towards the cardiac jelly is highlighted by cell protrusions specifically in this area of cell clustering; and (3) a highly stereotyped multi-layering process within the cardiac jelly leading to the formation of a functional leaflet. These observations enable us to confirm the previously reported role of mechanical forces in valve morphogenesis[10,25] and suggest a refined model in which the origins of the valve progenitors, the behaviour of particular groups of cells and the impact of the mechanotransduction cascade are identified.

Clustering of EdCs in a region of the AVC close to the ventricular inner curvature and localized fibronectin synthesis in the same area at 48 hpf is followed by the appearance of cells protruding into the cardiac jelly by 56 hpf. In the absence of *klf2a*, an important component of the mechanotransduction pathway downstream of blood flow[24], cell clustering and fibrillogenesis are both impaired. Ultimately, the coordinated morphogenetic movements subsequently observed in control embryos are perturbed in *klf2a* mutants highlighting the importance of an intact mechanotransduction pathway in orchestrating the cellular events involved in valve formation. Interestingly, we observed a striking change in cell properties following migration into the cardiac jelly, with cells demonstrating a mesenchymal morphology, downregulating junctional Cdh5 and some of them proliferating. While we cannot rule out a contribution of cell proliferation in other parts of the heart driving cells towards the AVC, within the AVC itself we only observed cell proliferation within the cardiac jelly and not within the cell layer that is exposed to the blood flow. We demonstrated that flow forces primarily influence cell behaviour and ultimately valve shape, but not cell number, via the asymmetric activation of *klf2a* expression in the ventricular region of the AVC. Interestingly, *klf2a*$^{-/-}$ mutants presented an array of valvular phenotypes similar to those observed under a range of altered flow regimes, implicating Klf2a as an integrator of the flow response, which, when absent, impacts valvulogenesis from the earliest stages. It will be interesting to see if fibronectin also relays Klf2a function in other contexts, such as haematopoietic stem cells (HSC) formation[32], development of the branchial arches[33] and endocardial chamber ballooning[6].

During valve formation the cells lining a unicellular tube are required to undergo extensive rearrangements in order to form a



	Thick	Blocks	n
Control	0%	0%	3
fn1b MO	0%	100%	7
fn1b+/+	10%	0%	19
fn1b–/–	22%	44%	9

Figure 7 | Klf2a expression regulates fibronectin synthesis in the AVC to drive valve formation. (**a**) *in situ* hybridization analysis shows the *fn1b* mRNA expression in the AVC of *klf2a*$^{+/+}$ embryos, at 48 and 56 hpf, is lost in the majority of *klf2a*$^{-/-}$ embryos. (**b**) Immunofluorescence analysis shows reduced fibronectin-positive staining (cyan) in the AVC of *klf2a*$^{-/-}$ mutants ($n=12$ *klf2*$^{+/+}$, $n=18$ *klf2a*$^{-/-}$). A *z*-projection of the whole AVC is shown. (**c**) Forced expression of Klf2a in EdCs using the *Tg(fli:gal4FF*^{ubs}*; UAS:klf2a)* line results in fibronectin-positive staining in the atrium and ventricle ($n=12/16$) compared with the AVC-specific staining pattern in controls ($n=12/14$) in at least three independent experiments. A *z*-projection of the whole heart is shown. (**d**) Analysis of *fn1b* morphants showed defective cell clustering in the superior AVC at 48 hpf (control $n=6$, *fn1b*MO $n=5$) and (**e**) reduced numbers of nuclei in the cardiac jelly at 72 hpf (control $n=5$, *fn1b*MO $n=9$). (**f**) Live imaging of valves at 96 hpf showed leaflets fail to form in *fn1b* morphants (controls $n=3$, *fn1b*MO $n=7$). Single frames from the live imaging are shown and red dotted lines outline the superior valve leaflet in each case. The ventricle (v) and atrium (a) are labelled for orientation. (**g**) Single frames from the live imaging of *fn1b*$^{-/-}$ mutant hearts, quantification of the percentage of thick/blocked valves observed in *fn1b* MO, *fn1b*$^{-/-}$ mutant hearts and their respective controls. Student's *t*-test $^{***}P<0.005$, $^{**}P<0.01$, $^{*}P<0.05$. Scale bars, 10 µm.

protrusive 3D structure capable of occluding undesired reversing flow. In the lymphatic system, the initial phases of valve formation have been elegantly described proceeding via the clustering of lymphatic valve progenitors on one side of the collecting vessel, adoption of a cuboidal morphology and the generation of a ring-like constriction within the vessel wall[11]. Valve-forming cells then protrude into the lumen of the lymphatic vessels, which, accompanied by ECM deposition, results in the formation of a valve leaflet[11,19]. In an initially similar manner, we see EdCs clustering on the ventricular side of the superior AVC, but rather than protruding towards the lumen, clustered EdCs then send basal protrusions into the cardiac jelly that they subsequently invade. These cells then proliferate and form a group of cells nestled close to the ventricular inner curvature, before a clear leaflet structure is formed. This significant difference in acquisition of cells into the leaflet in the lymphatic and cardiac systems is likely to be, at least in part, attributable to the different mechanical environments experienced by their progenitors.

During cardiac valve formation in the mouse, EdCs within the AVC undergo an endothelial-to-mesenchymal transition (EMT) and migrate into the cardiac jelly[34], where they proliferate and valve leaflets subsequently elongate. This is in striking contrast to lymphatic valve formation, which proceeds in the absence of an EMT. Interestingly, by embryonic day 10.5 (E10.5) the loss of *klf2* results in AVC cell disorganization and hypocellular cushions, attributed to a defect in the EMT process[35]. We also observe AVC cell disorganization and, in some instances, hypocellular cushions in the absence of *klf2a*. Considering this together with the downregulation of Cdh5 by EdCs inside the cardiac jelly, cell morphology changes and increased proliferation, we hypothesize that the cellular protrusions observed from the clustered cells

from 48 hpf and the subsequent movement of cells into the cardiac is reminiscent of the valve-forming process described in the mouse[35]. Importantly, the presence of an EMT during zebrafish valvulogenesis, although extensively discussed[18,26,28,36], remains to be confirmed and will be an important focus of future studies.

Valve morphogenesis occurs in one of the most hostile mechanical environments in the body: EdCs experience both high flow forces and strong mechanical deformation due to the contraction of the heart and its associated blood flow[1]. Indeed, at embryonic stages, viscosity dominates and the main mechanical forces generated at the heart wall are the tissue strain generated by pressure variations occurring during heart contraction and wall shear stress generated by the flowing blood[1]. These physical features are crucial to understanding how EdCs behave during valve development. Cell clustering close to the ventricular inner curvature in the AVC suggests localized differences in cell tension, which in turn may be responsible for the localized fibrillogenesis observed before multi-layering.

The specific enrichment of *klf2a* expression in the AVC at 48 hpf in response to flow and its regulation of *fibronectin1b* expression provides a mechanism to modulate valve morphogenesis in response to its mechanical environment[25]. For now, we can only speculate on the role of fibronectin in the process of valve formation. One possibility is that fibronectin is necessary for EdCs to acquire a valvular cell fate as it is now becoming clear that forces, ECM and stem cells fate are tightly interdependent[37]. The discovery that the mechanical environment and stretch-sensitive channels are essential determinants of lineage choices in embryonic stem cells further supports this hypothesis[38–40]. Fibronectin could also be involved in promoting the generation of the filopodia observed at the onset of AVC-specific fibronectin accumulation and before the multi-layering process. Indeed, fibronectin-rich nanoenvironments have been demonstrated to be sufficient for orienting cell migration and proliferation[41]. In addition, fibronectin may alter the mechanical properties of the EdCs environment as well as several cellular properties, including the mechanosensitivity of the cells themselves. Fibronectin has been implicated in the endothelial mechanotranduction pathway mediated by *trpv4* (ref. 42), a gene that is also involved in AVV formation[24]. Alternatively, fibrillogenesis could participate in cell-shape changes associated with morphogenesis[43], a feature that has also been attributed to *klf2a* expression[6].

In conclusion, we propose a model for AVV formation in which EdCs arising from the atrium and ventricle make specific contributions to the emerging valve leaflets through coordinated cell-movement events. Enriched mechanosensitive *klf2a* expression on the ventricular side of the AVC subsequently enriches the expression of its downstream target gene *fibronectin1b* to the same cells, establishing an asymmetry within the AVC. Ventricular EdCs extend protrusions towards the fibronectin-rich area and move into the cardiac jelly in a coordinated manner while accompanying morphogenetic changes result in atrial EdCs lining the lumen of the AVC. Subsequent Cdh5-downregulation, increased proliferation and morphological changes within the AVC liken aspects of AVV formation in the zebrafish to that of the mouse and higher vertebrates.

Finally, this study elucidates the impact of mechanical forces on the localized behaviours of valve progenitors during valve morphogenesis. As the origins of most valvulopathies are still unknown[12], our findings highlight the importance of investigating the potential embryonic and mechanical origins of valve defects. The recent demonstration that mitral valve disease can have embryonic origins supports this view[44–46]. Furthermore our work highlights the role of the coordination between morphogenesis and mechanical forces via ECM synthesis during cardiovascular morphogenesis. These findings may also prove meaningful in other biological contexts where mechanical forces and ECM are involved, such as during stem cell niche formation in developing HSCs[47–49], skin stem cells[50] and adult intestinal stem cells[51].

Methods

Zebrafish husbandry, embryo treatments and morpholinos. Animal experiments were approved by the Animal Experimentation Committee of the Institutional Review Board of the IGBMC. Zebrafish lines used in this study were *Tg(fli1a:gal4FF^{ubs}; UAS:kaede)*[52], *Tg(fli1a:nEGFP)^{y7}* (ref. 53), *Tg(kdrl:EGFP)*[54], *Tg(fli:nls-mcherry)*[24], *Tg(myl7:EGFP)*[55], *Tg(fli:gal4FF^{ubs}; UAS:klf2a)* (ref. 6), *Tg(klf2a:H2BEGFP)*[24] and wild-type AB. Zebrafish with a mutant allele of *klf2a* (*Tg(klf2a^{ig4})*) were generated and used in this study. Zebrafish with a mutant allele of *fn1b* (*Tg(fn1b^{Sa553})*) were obtained from the Zebrafish International Resource Center (ZIRC). *Tg(fn1b^{Sa553})* mutants contain a C>T point mutation in exon 1 of the *fn1b* gene leading to a premature stop codon in the predicted translation product. Genotyping was performed by sequencing the PCR product generated with the following primers: forward 5′-AGGGTGAGAGAACCTCATAAAGC-3′, reverse 5′-CTCACTTAAACCGCGAACTGTCC-3′, and sequenced with 5′- TCA GTAAAGAGACTCCTGCTGC-3′. Genotyping was performed on genomic DNA after live imaging. Morpholinos were injected into the yolk at the one-cell stage. All animals were incubated at 28.5 °C for 5 h before treatment with 1-phenyl-2-thiourea (PTU) (Sigma Aldrich) to prevent pigment formation. Morpholinos specific for *tnnt2a* (ref. 30) (5′-CATGTTTGCTCTGATCTGACACGCA-3′) and *fn1b* (5′-AAGTAATAATGTCACCTTGCTCCTC-3′) were obtained from GeneTools. Morpholinos for *klf2a*, *gata1*, *gata2* and *myh6* were described previously[25,56]. Anti-sense MO concentrations ranged from 0.06 to 0.3 mM.

Generation of *klf2a*$^{-/-}$ mutants. We injected a TALEN pair designed to target exon 1 of the *klf2a* gene into single cell wild-type (AB) embryos. We identified the alleles generated and confirmed that potential targeting events could be transmitted through the germline by out-crossing the F0 fish with AB animals and sequencing genomic DNA from pools of 6 F1 embryos. We focused on an INDEL mutation (deletion of 5′-CAGAAGGAA-3′ followed by insertion of 5′-GATGCTGGG AGAG-3′) leading to a premature stop codon in the *klf2a* transcript and raised these F1 animals to adulthood. Studies were performed on F2 fish, and later generations, following out-crossing to transgenic lines of interest. *Klf2a*$^{-/-}$ fish were viable and also kept as homozygous adults. A PCR-based genotyping strategy was established using the following primers to identify the wild-type and mutant alleles (Wild-type: forward 5′-TCGGCGCAGAAGGAAA-3′, reverse 5′-TGT TGAGGTTGTCCATGTTA-3′; mutant: forward 5′-AAGGTCTTCCACCACT CATA-3′, reverse 5′-CCAGCATTTCTCTCCCAGC-3′). Genotyping was performed on genomic DNA from whole embryos after live imaging, or from dissected tails before immunofluorescence analysis, as necessary.

mRNA sequencing of dissected heart samples. Hearts were dissected from *Tg(myl7:EGFP)* embryos at the desired stage and pooled (30 hearts/sample). RNA was extracted using a Nucleospin RNA XS kit (Macherey-Nagel) according to the manufacturer's instructions. After isolation of total cellular RNA, a library of template molecules suitable for high throughput DNA sequencing was created following the Illumina 'mRNA sequencing sample preparation guide' (part #1004898 Rev.D) with some modifications. Briefly, mRNA was purified from 20 ng total RNA using oligo-dT magnetic beads and fragmented using divalent cations at 94 °C for 5 min. The cleaved mRNA fragments were reverse transcribed to cDNA using random primers, then the second strand of the cDNA was synthesized using DNA Polymerase I and RNase H. The next steps of RNA-Seq Library preparation were performed in a fully automated system using SPRIworks Fragment Library System I kit (ref A84801, Beckman Coulter, Inc) with the SPRI-TE instrument (Beckman Coulter, Inc). Briefly, in this system, double-stranded cDNA fragments were blunted, phosphorylated and ligated to indexed adapter dimers, and fragments in the range of ~200–400 bp were size selected. The automated steps were followed by PCR amplification (30 s at 98 °C; (10 s at 98 °C, 30 s at 60 °C and 30 s at 72 °C) × 13 cycles; 5 min at 72 °C), then surplus PCR primers were removed by purification using AMPure XP beads (Agencourt Biosciences Corporation). DNA libraries were checked for quality and quantified using a 2100 Bioanalyzer (Agilent). The libraries were loaded in the flow cell at 6 pM concentration and clusters generated and sequenced in the Illumina Genome Analyzer IIX as single-end 54 base reads. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus[57] and are accessible through GEO Series accession number GSE79585 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE79585).

Bioinformatics and gene ontology analysis. Read quality was assessed with FastQC (S. Andrews, http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Then reads were mapped onto the Zv9 assembly of the zebrafish genome using Tophat v1.4.1 (ref. 58) and the bowtie v0.12.7 aligner. Only uniquely aligned reads were retained for further analyses. Gene expression was quantified using HTSeq

v0.5.3p5 (ref. 59) and gene annotations from Ensembl release 69. Read counts were normalized across libraries with the method proposed by Anders *et al.*[60]. To identify significantly differentially expressed genes, we performed a test for differential expression within the experiment (that is, adjusting for baseline differences between the experiments) using the method proposed by Robinson and Smyth[61] and implemented in the Bioconductor edgeR v3.0.8 package[62]. Adjustment for multiple testing was performed with the Benjamini and Hochberg[63] method. Functional analyses of genes with an adjusted *P* value smaller than 0.05 were performed using DAVID software[64]. Graphics were obtained with the R program (R Core Team, URL http://www.R-project.org/).

qPCR. Products were amplified in a real-time PCR reaction with Light Cycler 480 Real-Time PCR System (Roche) using a UPL Probes Master mix (Roche) according to the manufacturer's instructions. Sequence of primer pairs were as follows: *has2* forward 5′-AGCATCCCTGTTCAACTAACG-3′, reverse 5′-GCTGACCGCTTT ATCACATCT-3′; *klf2a* forward 5′-CCGTCTATTTCCACATTTTCG-3′, reverse 5′-TCCAGTTCATCCTTCCACCT-3′; *fn1b* forward 5′-TGGAAATGTGATGCC ATTGA-3′, reverse 5′-GGCCAATCTGGTAGAACACC-3′; *p53* forward 5′-GAGG TCGGCAAAATCAATTC-3′, reverse 5′-TGGGGGCTGAATAATCAAAT-3′.

***In vivo* imaging.** Zebrafish embryos were staged, anaesthetised with 0.02% tricaine solution and 50 mM BDM, to stop the heart when necessary, and mounted in 0.7% low melting-point agarose (Sigma Aldrich). Confocal imaging was performed on a Leica SP8 confocal microscope. Fast confocal and four-dimensional imaging (to image valve leaflets at 96 hpf and the AVC at 72 hpf) was performed using the resonant scanner mode of the same microscope. Images were acquired with a low-magnification water immersion objective (Leica HCX IRAPO L, 25X, N.A. 0.95). For four-dimensional imaging, time series were acquired at a random time in the cardiac cycle at 35fps for 3 s. The optical plane was moved 2 μm between the *z*-sections until the whole AVC was acquired. Time series of two-dimensional sections were temporally synchronized using Matlab[27]. Blood flow imaging for flow velocity analysis was performed on a Leica DMIRBE inverted microscope using a Photron SA3 high speed CMOS camera (Photron, San Diego, CA) and water immersion objective (Leica ×20, NA 0.7). Image sequences were acquired at a frame rate of 2,000 frames per second.

Photoconversion. Photoconversion was performed using the FRAP module on a SP8 confocal microscope and a Leica HCX IRAPO L, ×25, NA0.95 water immersion objective. *Tg(fli1a:Gal4FF; UAS:Kaede)* embryos were mounted in 0.7% low melting-point agarose supplemented with 50 mM BDM to inhibit heart contraction for the duration of the procedure. A region of interest corresponding to the ventricle, atrium or superior AVC was selected and exposed to 405 nm light (25% laser power). One pre-bleach frame was acquired, followed by 3–5 bleach pulses (3–5 ms each) without acquisition to achieve conversion of the kaede protein to its red form. A *z*-stack of the photoconverted heart was then acquired in the standard confocal mode to record the starting point of each experiment. Embryos were then carefully dissected from the agarose, placed in fish water for 5–10 min until heart contraction resumed and then put at 28.5 °C to develop individually under standard conditions until a time point of interest. The movement of cells within each heart was analysed using Imaris software (Bitplane).

Valve imaging. Embryos were incubated with 4 μM BODIPY-ceramide (Molecular Probes) overnight and then processed as in refs 24,25 to visualize the valve shape.

Flow analysis. Red blood cells were manually tracked through the AVC and their velocity calculated from image sequences acquired at 2,000 frames per second as described previously[24].

Immunofluorescence. Embryos were fixed at the desired stage in 2% paraformaldehyde overnight at 4 °C. BrdU-incorporation studies were performed by incubating embryos in 5 mg ml^{-1} BrdU for the desired length of time (Dietrich *et al.*[6]) before fixation. After washing, embryos were permeabilized in 1× PBS-0.1% Tween-20 containing 0.5% Triton-X 100 for 30 min at room temperature. The pericardial cavity was then carefully pierced with the tip of a forcep to facilitate antibody penetration before blocking in permeabilization buffer supplemented with 5% BSA (anti-fibronectin), 1% BSA and 10% NGS (anti-VECadherin and anti-Alcama) or 1% BSA, 2% NGS and 1% DMSO (anti-BrdU) for 2 h at room temperature. Primary antibodies were added in the relevant blocking solution and incubated between 16 and 48 h at 4 °C. Secondary antibodies were added in blocking solution after thorough washing and incubated overnight at 4 °C. Embryos were thoroughly washed and mounted for imaging on a Leica SP8 confocal. Antibodies were as follows: rabbit anti-fibronectin (F3648, Sigma) 1:100, rabbit anti-VECadherin[65] 1:1,000 (kind gift of the Affolter lab), mouse anti-BrdU (11170376001, Roche Diagnostics) 1:100, mouse anti-Alcama (zn-8, DSHB) 1:500 and goat anti-rabbit and goat anti-mouse Alexa-488 and -594 secondary antibodies (A11034 and A11032, respectively, Life Technologies) were used at 1:500. To directly test the effects of flow on fibronectin synthesis, embryos were incubated in 0.1% tricaine (Ethyl 3-aminobenzoate methanesulfonic acid; Sigma) at pH7 for up to 12 h. They were then rinsed briefly in egg water before being fixed and processed as described above.

Cell proliferation assay. Dechorionated embryos were incubated in fish water containing 5 mg ml^{-1} BrdU between 36 and 48 hpf, 56 and 72 hpf, and 72 and 80 hpf. Incorporation was stopped by washing in fresh fish water and fixation in 4% PFA. Embryos were permeabilized with Proteinase K and DNA denatured with 2 N HCl (method modified from Dietrich *et al.*[6]). BrdU immunolabelling was then performed as described above.

Image analysis. Cell number quantifications and *klf2a:EGFP* signal intensity measurements were made using the Spots tool on Imaris (Bitplane). A single spot was placed at the centre of each nucleus in the AVC, or cardiac jelly, as appropriate. For intensity analysis, the *klf2a:H2BEGFP* reporter line was crossed with the *fli:nlsmCherry* line and the mCherry fluorescence signal was used for normalization[24]. The maximum intensity of each channel was quantified and a ratio generated. These ratios were then averaged across the AVC of individual embryos. Nucleus to nucleus distance analysis was performed using Imaris software and the Measurement Points tool. Nuclei within the AVC were connected to all of their nearest neighbours and the average distances for defined regions were calculated. The extent of cell clustering, in *klf2a* mutants and *fn1b* morphants, was quantified by defining clustered cells as those cells closer than 8 μm to their neighbour (according to our analysis of nucleus-to-nucleus distances in controls; Fig. 2e). The extent of fibronectin staining in the AVC was calculated using the Imaris Surfaces tool to define a volume of fibronectin staining and a volume of the whole AVC. The % of fibronectin coverage was then calculated ([Volume fibronectin staining/ Volume AVC]*100). Hearts were segmented using the surfaces tool and the segmented heart presented, for clarity.

***In situ* hybridization.** *In situ* hybridization was performed as in ref. 66. Anti-sense probes for *fn1b* were generated from a plasmid containing *fn1b* cDNA (obtained from SourceBioscience) amplified using the following primers forward: 5′-ATG ACCCGTGAGTCAGTAA-3′ and reverse (containing T3 sequence): 5′-ATTAA CCCTCACTAAAGGGACTTGGTGCCCTGAGTTCTGAT-3′ and subsequently transcribed using the T3 polymerase.

References

1. Boselli, F., Freund, J. B. & Vermot, J. Blood flow mechanics in cardiovascular development. *Cell. Mol. Life Sci.* **72,** 2545–2559 (2015).
2. Freund, J. B., Goetz, J. G., Hill, K. L. & Vermot, J. Fluid flows and forces in development: functions, features and biophysical principles. *Development* **139,** 1229–1245 (2012).
3. Heisenberg, C. P. & Bellaiche, Y. Forces in tissue morphogenesis and patterning. *Cell* **153,** 948–962 (2013).
4. Guillot, C. & Lecuit, T. Mechanics of epithelial tissue homeostasis and morphogenesis. *Science* **340,** 1185–1189 (2013).
5. Mammoto, T. & Ingber, D. E. Mechanical control of tissue and organ development. *Development* **137,** 1407–1420 (2010).
6. Dietrich, A. C., Lombardo, V. A., Veerkamp, J., Priller, F. & Abdelilah-Seyfried, S. Blood flow and Bmp signaling control endocardial chamber morphogenesis. *Dev. Cell* **30,** 367–377 (2014).
7. Liu, J. *et al.* A dual role for ErbB2 signaling in cardiac trabeculation. *Development* **137,** 3867–3875 (2010).
8. Peshkovsky, C., Totong, R. & Yelon, D. Dependence of cardiac trabeculation on neuregulin signaling and blood flow in zebrafish. *Dev. Dyn.* **240,** 446–456 (2011).
9. Peralta, M. *et al.* Heartbeat-driven pericardiac fluid forces contribute to epicardium morphogenesis. *Curr. Biol.* **23,** 1726–1735 (2013).
10. Hove, J. R. *et al.* Intracardiac fluid forces are an essential epigenetic factor for embryonic cardiogenesis. *Nature* **421,** 172–177 (2003).
11. Sabine, A. *et al.* Mechanotransduction, PROX1, and FOXC2 cooperate to control connexin37 and calcineurin during lymphatic-valve formation. *Dev. Cell.* **22,** 430–445 (2012).
12. MacGrogan, D. How to make a heart valve: from embryonic development to bioengineering of living valve substitutes. *Cold Spring Harb. Perspect. Med.* **4,** a013912 (2014).
13. Wallingford, J. B. Planar cell polarity and the developmental control of cell behavior in vertebrate embryos. *Annu. Rev. Cell. Dev. Biol.* **28,** 627–653 (2012).
14. Back, M., Gasser, T. C., Michel, J.-B. & Caligiuri, G. Biomechanical factors in the biology of aortic wall and aortic valve diseases. *Cardiovasc. Res.* **99,** 232–241 (2013).
15. Sauls, K. *et al.* Developmental basis for filamin-A-associated myxomatous mitral valve disease. *Cardiovasc. Res.* **96,** 109–119 (2012).
16. Richards, J. M., Farrar, E. J., Kornreich, B. G., Moïse, N. S. & Butcher, J. T. The mechanobiology of mitral valve function, degeneration, and repair. *J. Vet. Cardiol.* **14,** 47–58 (2012).

17. Odelin, G. *et al.* Loss of Krox20 results in aortic valve regurgitation and impaired transcriptional activation of fibrillar collagen genes. *Cardiovasc. Res.* **104,** 443–455 (2014).
18. Lagendijk, A. K., Szabo, A., Merks, R. M. & Bakkers, J. Hyaluronan: a critical regulator of endothelial-to-mesenchymal transition during cardiac valve formation. *Trends Cardiovasc. Med.* **23,** 135–142 (2013).
19. Bazigou, E. *et al.* Integrin-alpha9 is required for fibronectin matrix assembly during lymphatic valve morphogenesis. *Dev. Cell* **17,** 175–186 (2009).
20. Boselli, F. & Vermot, J. Live imaging and modeling for shear stress quantification in the embryonic zebrafish heart. *Methods* **94,** 129–134 (2015).
21. Huisken, J. & Stainier, D. Y. Selective plane illumination microscopy techniques in developmental biology. *Development* **136,** 1963–1975 (2009).
22. Lenard, A. *et al.* Endothelial cell self-fusion during vascular pruning. *PLoS Biol.* **13,** e1002126 (2015).
23. Mahou, P., Vermot, J., Beaurepaire, E. & Supatto, W. Multicolor two-photon light-sheet microscopy. *Nat. Methods* **11,** 600–601 (2014).
24. Heckel, E. *et al.* Oscillatory flow modulates mechanosensitive klf2a expression through trpv4 and trpp2 during heart valve development. *Curr. Biol.* **25,** 1354–1361 (2015).
25. Vermot, J. *et al.* Reversing blood flows act through klf2a to ensure normal valvulogenesis in the developing heart. *PLoS Biol.* **7,** e1000246 (2009).
26. Beis, D. *et al.* Genetic and cellular analyses of zebrafish atrioventricular cushion and valve development. *Development* **132,** 4193–4204 (2005).
27. Liebling, M. *et al.* Rapid three-dimensional imaging and analysis of the beating embryonic heart reveals functional changes during development. *Dev. Dyn.* **235,** 2940–2948 (2006).
28. Scherz, P. J., Huisken, J., Sahai-Hernandez, P. & Stainier, D. Y. High-speed imaging of developing heart valves reveals interplay of morphogenesis and function. *Development* **135,** 1179–1187 (2008).
29. Huang, H. T. *et al.* A network of epigenetic regulators guides developmental haematopoiesis *in vivo*. *Nat. Cell. Biol.* **15,** 1516–1525 (2013).
30. Sehnert, A. J. *et al.* Cardiac troponin T is essential in sarcomere assembly and cardiac contractility. *Nat. Genet.* **31,** 106–110 (2002).
31. Kalogirou, S. *et al.* Intracardiac flow dynamics regulate atrioventricular valve morphogenesis. *Cardiovasc. Res.* **104,** 49–60 (2014).
32. Wang, L. *et al.* A blood flow-dependent klf2a-NO signaling cascade is required for stabilization of hematopoietic stem cell programming in zebrafish embryos. *Blood* **118,** 4102–4110 (2011).
33. Nicoli, S. *et al.* MicroRNA-mediated integration of haemodynamics and Vegf signalling during angiogenesis. *Nature* **464,** 1196–1200 (2010).
34. von Gise, A. & Pu, W. T. Endocardial and epicardial epithelial to mesenchymal transitions in heart development and disease. *Circ. Res.* **110,** 1628–1645 (2012).
35. Chiplunkar, A. R. *et al.* Kruppel-like factor 2 is required for normal mouse cardiac development. *PloS One* **8,** e54891 (2013).
36. Lagendijk, A. K., Goumans, M. J., Burkhard, S. B. & Bakkers, J. MicroRNA-23 restricts cardiac valve formation by inhibiting Has2 and extracellular hyaluronic acid production. *Circ. Res.* **109,** 649–657 (2011).
37. Bellas, E. & Chen, C. S. Forms, forces, and stem cell fate. *Curr. Opin. Cell Biol.* **31,** 92–97 (2014).
38. Engler, A. J., Sen, S., Sweeney, H. L. & Discher, D. E. Matrix elasticity directs stem cell lineage specification. *Cell* **126,** 677–689 (2006).
39. McBeath, R., Pirone, D. M., Nelson, C. M., Bhadriraju, K. & Chen, C. S. Cell shape, cytoskeletal tension, and RhoA regulate stem cell lineage commitment. *Dev. Cell* **6,** 483–495 (2004).
40. Trappmann, B. *et al.* Extracellular-matrix tethering regulates stem-cell fate. *Nat. Mater.* **11,** 642–649 (2012).
41. Albuschies, J. & Vogel, V. The role of filopodia in the recognition of nanotopographies. *Sci. Rep.* **3,** 1658 (2013).
42. Thodeti, C. K. *et al.* TRPV4 channels mediate cyclic strain-induced endothelial cell reorientation through integrin-to-integrin signaling. *Circ. Res.* **104,** 1123–1130 (2009).
43. Compagnon, J. *et al.* The notochord breaks bilateral symmetry by controlling cell shapes in the zebrafish laterality organ. *Dev. Cell* **31,** 774–783 (2014).
44. Dina, C. *et al.* Genetic association analyses highlight biological pathways underlying mitral valve prolapse. *Nat. Genet.* **47,** 1206–1211 (2015).
45. Durst, R. *et al.* Mutations in DCHS1 cause mitral valve prolapse. *Nature* **525,** 109–113 (2015).
46. Levine, R. A. *et al.* Mitral valve disease-morphology and mechanisms. *Nat. Rev. Cardiol.* **12,** 689–710 (2015).
47. Jing, L. *et al.* Adenosine signaling promotes hematopoietic stem and progenitor cell emergence. *J. Exp. Med.* **212,** 649–663 (2015).
48. Kim, P. G. *et al.* Flow-induced protein kinase A-CREB pathway acts via BMP signaling to promote HSC emergence. *J. Exp. Med.* **212,** 633–648 (2015).
49. Diaz, M. F. *et al.* Biomechanical forces promote blood development through prostaglandin E2 and the cAMP-PKA signaling axis. *J. Exp. Med.* **212,** 665–680 (2015).
50. Hsu, Y. C., Li, L. & Fuchs, E. Emerging interactions between skin stem cells and their niches. *Nat. Med.* **20,** 847–856 (2014).
51. Shyer, A. E., Huycke, T. R., Lee, C., Mahadevan, L. & Tabin, C. J. Bending gradients: how the intestinal stem cell gets its home. *Cell* **161,** 569–580 (2015).
52. Herwig, L. *et al.* Distinct cellular mechanisms of blood vessel fusion in the zebrafish embryo. *Curr. Biol.* **21,** 1942–1948 (2011).
53. Roman, B. L. *et al.* Disruption of acvrl1 increases endothelial cell number in zebrafish cranial vessels. *Development* **129,** 3009–3019 (2002).
54. Jin, S. W., Beis, D., Mitchell, T., Chen, J. N. & Stainier, D. Y. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development* **132,** 5199–5209 (2005).
55. Huang, C. J., Tu, C. T., Hsiao, C. D., Hsieh, F. J. & Tsai, H. J. Germ-line transmission of a myocardium-specific GFP transgene reveals critical regulatory elements in the cardiac myosin light chain 2 promoter of zebrafish. *Dev. Dyn.* **228,** 30–40 (2003).
56. Galloway, J. L., Wingert, R. A., Thisse, C., Thisse, B. & Zon, L. I. Loss of gata1 but not gata2 converts erythropoiesis to myelopoiesis in zebrafish embryos. *Dev. Cell* **8,** 109–116 (2005).
57. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30,** 207–210 (2002).
58. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* **25,** 1105–1111 (2009).
59. Anders, S. HTSeq: Analysing high-throughput sequencing data with Python. Available at: http://www-huber.embl.de/users/anders/HTSeq (2014).
60. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11,** R25 (2010).
61. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40,** 4288–4297 (2012).
62. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).
63. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* **57,** 289–300 (1995).
64. Huang, da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).
65. Blum, Y. *et al.* Complex cell rearrangements during intersegmental vessel sprouting and vessel fusion in the zebrafish embryo. *Dev. Biol.* **316,** 312–322 (2008).
66. Thisse, C. & Thisse, B. High-resolution *in situ* hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3,** 59–69 (2008).

Acknowledgements

We thank D. Riveline, J. Pestel, D. Stainier and the Vermot laboratory for discussion and thoughtful comments on the manuscript. We thank K. Yaniv, H.G. Belting and M. Affolter for providing the fish stocks, antibodies and protocols for immunohistochemistry. M. Charpentier is acknowledged for help with talen construction. We thank the IGBMC fish facility (S. Geschier and S. Gredler) and the IGBMC imaging center, in particular B. Gurchenkov, P. Kessler, M. Koch and D. Hentsch. We thank M. Philipps, B. Jost, C. Keime and A. Velt in the IGBMC Microarrays deep sequencing platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009), for performing mRNA sequencing experiments and for helpful advice on analysis. This work was supported by HFSP, INSERM, AFM, FRM, the seventh framework program (MC-IRG256549) and by the grant ANR-10-LABX-0030-INRT, a French State fund managed by the Agence Nationale de la Recherche under the frame program Investissements d'Avenir labeled ANR-10-IDEX-0002-02. E.S. was supported by the FRM and Post Doctorat IDEX. J.P.C. was supported by the grant ANR-II-INBS-0014.

Author contributions

E.S. and J.V. designed the experiments. E.S. performed the experiments. J.P.C. generated the *klf2a* talen sequences and C.R. generated the *klf2a* mutants and designed the genotyping strategy. S.R. and N.F. provided technical help in generating, maintaining and analysing the *klf2a* and *fn1b* mutant line. E.S. and J.V. analysed the data and wrote the paper.

Additional information

Accession codes: The gene expression data have been deposited in the NCBI Gene Expression Omnibus (GEO) database under accession code GSE79585.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Steed, E. *et al. klf2a* couples mechanotransduction and zebrafish valve morphogenesis through Fibronectin synthesis. *Nat. Commun.* 7:11646 doi: 10.1038/ncomms11646 (2016).

This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/

Université de Strasbourg

École Doctorale des Sciences de la Vie et de la Santé STRASBOURG

igbmc Institut de génétique et de biologie moléculaire et cellulaire

Nathalie Faggianelli-Conrozier

Deciphering the roles of *Klf2a*, *Klf2b* and *Egr1* transcription factors in heart valve development using zebrafish as model organism

Abstract

Cardiac valves are necessary for maintaining a unidirectional blood flow in the cardiovascular system of vertebrates. Their efficient gating function requires a highly controlled developmental program. However, this program may be impaired and thus leading to defective valves. In fact, congenital heart valve diseases represent the most common form of birth defects. Therefore, cardiac valve development studies constitute a challenging research field. In this thesis, we used the zebrafish as a model organism for studying the formation of atrioventricular valves. To date, it is known that mechanical forces generated by blood flow constitute key modulators dictating valve formation. In particular, they initiate valvulogenesis by restricting the expression of the transcription factor *Klf2*a in a subset of endocardial cells of the atrio-ventricular canal. Our work demonstrated the activation of another transcription factor, *Egr1*, in this same region and within the same time window. We aimed at deciphering the mechanosentitive gene network involving *klf2a*, its paralog *klf2b* as well as *egr1*, by combining genome-wide analysis of gene expression and chromatin accessibility with live imaging. We addressed the potential interactions of these factors and studied their downstream signalling pathways. Finally, we demonstrated that *egr1*, klf2a/klf2b modulates valve morphogenesis by specifically controlling *flt1*, *has2* and *wnt9b* expression. Moreover, we showed that *klf2b* could regulate *egr1*. Together, these results uncover a novel mechanosensitive axis during cardiac valve development.

Résumé

La circulation du flux sanguin à sens unique dans le système cardiovasculaire des vertébrés est assurée par les valves cardiaques. Leur formation est très contrôlée au cours du développement embryonnaire. Cependant, il arrive que celle-ci soit défectueuse, et donc à l'origine de maladies cardiaques congénitales. Ces maladies représentent une des causes majeures de décès à la naissance. L'étude de la formation des valves cardiaques constitue donc un champ de recherche majeur. Dans cette thèse, nous avons utilisé le poisson zèbre, comme animal d'étude modèle pour étudier la formation des valves atrio-ventriculaires. Les forces mécaniques générées par le flux sanguin constituent un signal modulant le programme génétique valvulaire. Elles initient la formation des valves en contraignant l'expression du facteur de transcription, Klf2a, à un groupe de cellules endothéliales du canal atrio-ventriculaire. Nos travaux ont démontré l'activation d'un autre facteur, Egr1, dans cette même région dans le même lapse de temps. Notre étude a cherché à élucider le réseau génétique impliquant *klf2a*, son paralogue *klf2b*, et *egr1* en combinant une analyse pangénomique de l'expression génique et des sites accessibles de la chromatine avec une approche d'imagerie haute résolution *in vivo*. Nous avons déterminé les interactions entre ces facteurs et les réseaux qu'ils régulent. Cette étude a finalement démontré qu'*egr1*, *klf2a/klf2b* modulent la morphogénèse des valves cardiaques en contrôlant en particulier *flt1*, *has2* et *wnt9b*. De plus, nos travaux tendent à montrer que *klf2b* régulerait *egr1*. Cette étude a donc permis de révéler un nouvel axe de régulation génétique mécano-dépendant nécessaire à la bonne formation des valves cardiaques.